MATHEMATICA
Policy Research, Inc.

P.O. Box 2393
Princeton, NJ 08543-2393
Telephone (609) 799-3535
Fax (609) 799-0005
www.mathematica-mpr.com

MEMORANDUM

TO: Mike Jones, Wayne Gordon, Daniel Carroll

FROM: Steve Williams, Frank Potter, Dan Kasprzyk

DATE: 12/30/2009
NAWS - 2

SUBJECT: National Agricultural Workers Survey: Current Analysis Weights

---

1. EXECUTIVE SUMMARY

The Employment and Training Administration (ETA) was requested to provide for an independent review and assessment of the revised equations and the calculation of the sampling weights for the National Agricultural Workers Survey (NAWS). The revisions were implemented in response to concerns raised by the Bureau of Labor Statistics (BLS). This assessment included

1. an appraisal of the modified equations for the weights and the computation of the weights,
2. an independent analysis of the potential bias in survey estimates based on the old weights that were produced by ETA and the NAWS contractor, and
3. examination of the current (modified) weights for potential bias.

Mathematica conducted this review and the following summarized our assessment.

- **Equations and Accuracy of the Computations:** By concurrent review of the equations in the text of the OMB statement and the program code used in the computation of the weights, it was determined that the modified equations in the text did address the concerns raised by BLS. The program code was checked and tested to confirm that the equations were accurately implemented. Computations made by Mathematica matched the weights computed by the NAWS contractor.
- **Accuracy of Published Results:** For the variables that we used, we have found no significant or substantive differences in the point estimates and sampling errors by using the analysis weights developed following the methods in the OMB statement for the years 2001 to 2005 as well as for the period 2006 to 2008.
- **Impact of Methodology Change:** The concern about quota sampling that was previously noted was addressed in proposed methodology by completing the allocated sample for the final employer in the sample for each sample county.

- **Potential for Bias in Estimates Computed using the Current (Modified) Weights:** A number of factors indicate that survey estimates generated using the current (modified) weights have a potential for bias. The factors that indicate the potential for bias include (a) the omission of the first stage weights in the computation of the sampling weights, (b) the use of and variation in the post-sampling weights, (c) the use of a global nonresponse factor rather than adjustments at the sampling stages, and (d) the lack of written procedures for use in the selection of the workers at the grower. These elements are discussed below.

- **The Sampling Weights Omit a Stage of Sampling:** Sampling weights omit the first stage of sampling, which is the selection of the panel of 90 FLAs with unequal probabilities. These weights should be accounted for in the sampling weights. The omission of the first-stage weights means the valid inference population is reduced to the 90 FLAs, which has quite different characteristics than the actual target population (because the FLAs were selected with unequal probabilities). The first stage weights for the current sample of 90 FLAs range from 1.0 for large FLAs to approximately 20.0 or larger for small FLAs. The text for the OMB statement did not state that these first stage weights were not used and the text did not provide a rationale for ignoring the first stage weights.

- **Sampling and Analysis Weights Adjustments:** Post-sampling weights (usually referred to as post-stratification adjustments) are normally used to adjust probability based survey estimates to known control totals (that is to fine-tune probability-based estimates) and the range of adjustment values is generally between 0.90 and 1.10. In the NAWS weighting methodology, these post-stratification adjustments are sometimes very large or very small. Between 2001 and 2005 (cycles 38 to 52), the post-stratification factors for the year weights range from less than 0.20 to more than 60, when the first-stage weights are omitted. For one recent year (2007, cycles 56, 57 and 58), the post-stratification factors for the year weights range from less than 0.50 to more than 300, when the first-stage weights are omitted. When the first-stage weights and comparable factors are included, the range is from less than 0.50 to almost 50.0. Post-sampling weights are relied on to account for nonresponse of counties and FLAs as well as the first stage sampling weights. These Post-sampling weights dominate the weight computations in each cycle and year.

- **Global Nonresponse Adjustment:** At the region level, a nonresponse adjustment factor is used but nonresponse adjustment may have greater potential for reducing non-response bias if more sensitive procedures (such as the weighting class procedures or response propensity models) were developed for smaller units like FLAs, counties, or employers. However, limited information exists at all levels. While implicit adjustments for nonresponse are implemented for employees and employers, county and FLA level nonresponse adjustments should be considered rather than assuming that global adjustments through "post-sampling" weight adjustments are sufficient for compensating for nonresponse.

- **Design Effects (Deff) from Clustering and Unequal Weights:** Large design effects (Deffs) exist due in-part to clusters of sample cases but are also caused by severe unequal weighting related to selecting employers with equal probabilities. Using probability proportional to size selection should be considered for employers, if data are available for the individual employer. Note that the use of first-stage sampling weights, which are not reflected in the past Deffs, would further impact precision. Because the first stage sample of 90 FLAs and the subsequent subsample of 30 for each cycle are selected pps WOR, the product of the weights can result in wide variation in the weights. If this stage of subsampling continues to be used, equal probabilities should be used either to select the 90 FLAs or the cycle subsample of 30.

- **Selection of Workers:** The methodology for selecting workers is not clearly documented, and the interviewer manual does not provide instructions or forms to be used in selecting workers at the final stage. In the OMB statement, Appendix A: Contacting and Selecting Farm Workers, instructions are provided on how to chose the workers (see Attachment A). While these instructions provide a brief description of systematic sampling, no instructions are provided on how to select the initial worker nor how handle situations when the number of workers is not an even multiple of the sample size. This lack of instructions and forms may result in non-random sampling of agricultural workers at the final stage.

- **Accuracy and Clarity of Part B Text in the OMB document:** Most of the Part B section of the OMB statement is clear and accurate and the program code is consistent with the text. However, the presentation of the weighting methodology is incomplete and needs clarification and the text is not consistent with the program code. Understanding the weighting procedures and the equations in the OMB text relied on a review of the program code that was used for computing the weights. Since the program code generally would not be available to reviewers or users, the text itself would not present a clear description of the methods used to compute the weights. For example, the text does not clearly state that the first stage weights are not used and no rationale is given for not using these weights. In some portion of the text the use of subscripts was inadequate for clarity. Also, the text does not show the link between the post-sampling weights and the worker-level sampling weights that is needed in the final analyses. Following the guidance of the *Federal Committee on Statistical Methodology Working Paper 32* and the *Principles and Practices of Statistical Agencies,* improved and more thorough documentation of survey and statistical procedures is warranted.

- **Quality Control of Survey Procedures and Reduction of Nonsampling Error.** While not an aspect of the review of the NAWS estimation system, additional attention ought to be paid to ensure that grower and worker samples are drawn correctly. This suggests that more emphasis on a strong, explicitly defined quality control program for the field implementation is necessary.

2. STUDY OVERVIEW

2.1 Objectives

As part of ongoing efforts to assess the accuracy of the information it reports, the U.S. Department of Labor, Employment and Training Administration (ETA) requested a review of the National Agricultural Workers Survey (NAWS) sampling weights. The review is to investigate the extent of potential biases in the calculation of the weights and determine the effect of any biases on previous estimates produced by ETA and the NAWS contractor. A central issue is to ensure that the weights account for each of the relevant stages of the sample design and all of the relevant sources of variability in the selection probabilities, and thus reflects the probabilities of selection of the observed units. Weights to be addressed are:

1. The weights previously used by the contractor;
2. The weights being proposed for future use by the contractor, and
3. The corrected inverse-probability weights intended to account for all relevant features of the sample design

A related objective is to review concerns listed by the BLS regarding the calculation and implementation of the survey's sampling weights and ETA's response to those concerns.

To assess the extent of the potential bias in previous survey estimates, estimates were computed based on the old and new weights and compared. These old and new weights were computed following the procedures described by the NAWS contractor and programs provided by the contractor. An additional weight was developed to address concerns about the exclusion of the first stage weights. An analysis of micro data collected over several years was conducted to evaluate the statistical significance and practical significance of differences among the estimates computed.

Practical significance was defined as a difference between a new and old estimate of five percentage points or greater. To determine the extent of potential for bias, a list of key variables were identified. These variables are:

- Agricultural worker hourly wage (WAGET1)
- Age
- Spanish as a primary language
- Place of birth
- Legal status
- Farm work weeks
- Family income
- Highest grade completed
- Years since first arriving to the United States
- Years doing farm work
- Number of children in the household under age 18

Several documents relating to the NAWS design and related methodology were reviewed. These include the text of the OMB clearance document (both the current [January 26, 2009] version and a prior version), the description of the NAWS and the statistical methods that are posted on the ETA website, the NAWS Codebook for Public Access Data. Most of the documents suggested for review were those directly related to the survey, but others were also mentioned.

2.2 Study Methodology

Consistent with the study objectives, the review involved:

- Conference calls were held with ETA, JBS, and BLS staff on methods and issues relating to the study
- Various documents were reviewed, including the text in Part B of the OMB statement, past NAWS reports and other related documents, and program code used by JBS to compute weights and estimates
- Sets of weights were calculated using the JBS program code for the weights. Because the weights were computed using the JBS program code, these weights were a "somewhat independent" assessment of the weighting procedures and the same basic methods were used, for example, first stage weights were not used.
- Point estimates for eight key variables were computed to compare the published results with new methodology being proposed using weights computed assuming a

probability proportional to size (pps) without replacement (WOR) sample. The standard errors (SEs) were computed assuming a pps with replacement (WR) sample because of study time constraints. There was no attempt made to compute the joint inclusion probabilities, which are needed to correctly calculate the SEs assuming pps WOR sampling.

- The comparison of the "new weights" with and without the first stage weights

3. NATIONAL AGRICULTURAL WORKERS SURVEY (NAWS) SAMPLING DESIGN

3.1 Introduction

The National Agricultural Workers Survey (NAWS) is an employment-based survey of randomly sampled hired crop workers. The U.S. Department of Labor (DOL) sponsors the survey and collaborates with other Federal agencies to meet the nation's need for farm worker statistics. As a result, the NAWS is the nation's primary information source for demographic information regarding the employment, health and living conditions of hired crop workers. NAWS findings serve many purposes, including informing debates on immigration policy, contributing to formulas on farm worker population size or program funding allocations, and providing data to support farm worker policy, and program planning, design, and evaluation.

The goal of the NAWS statistical methods is to produce statistics for the hired crop work force. The NAWS survey population includes all field workers employed in crop agriculture in the conterminous United States (U.S.). The mobility of a large segment of the hired crop workforce and the temporal nature of agricultural work pose unique challenges to obtaining a nationally representative random sample of migrant and seasonal crop farm workers.

As a result of these objectives and issues, the NAWS uses a complex sampling design that includes both stratification and clustering. The NAWS is an establishment survey, sampling workers at their places of employment since a household survey would be infeasible. In the document "Statistical Methods of the National Agriculture Workers Survey (available on the ETA website *www.doleta.gov/agworker/statmethods.cfm*), the survey design was developed to achieve nearly equal weights for a nationally representative sample of individual workers.

3.2 Details of the NAWS Sampling Design

The NAWS uses stratified multi-stage sampling to account for seasonal and regional fluctuations in the level of farm employment. The stratification consists of 12 geographic regions. Three surveys per year are conducted; each of these "cycles" is based on a stand-alone sample selected from a standing roster of 90 randomly selected, multi-county areas. Each of the 12 strata is represented in each cycle. The cross of 3 cycles X 12 strata are considered as 36 strata for analyses. The county, or multi-county, units (farm labor areas-FLAs) are considered the primary sampling units (PSUs) for analyses. Farm employers located within PSUs are considered the secondary level and workers employed by farmers/growers are the tertiary level of sampling units. In fact, therefore, the initial stage of selecting the stand-alone panel of 90 FLAs, and selection of counties within sample FLAs are ignored in this characterization. The number of interviews allocated to each location is proportional to the crop activity at that time of the year. Interview allocation is thus proportional to stratum size.

In each sample county, a simple random sample of agricultural employers is drawn from a list compiled from public agency records, mostly unemployment insurance records for the larger employers. NAWS interviewers then contact the sample growers or farm labor contractors, arrange access to the work site, and draw a random sample of workers at the work site. Thus, the sample includes only farm workers actively employed at the time of the interview. Operationally in the newly proposed methodology, interviewers conduct interviews until county allocations are satisfied, interviewing the full allocation for the final employer in the sample. The full allocation to the final sample county in the FLA is not necessarily completed— the sampling stops when the FLA allocation of interviews is obtained. More specifically, data collection ceases when the quota of interviews is complete, as opposed to when all growers in the sample of growers have been contacted. That is, sample counties faces a similar issue as employers did before changing the methodology now requiring the full allocation to the employer be completed for the last employer. Obviously, the numbers of counties and employers in the sample are not known in advance of the field interviewing, and so, are random variables.

In conference calls with JBS staff, the JBS staff reported that the field interviewers are trained to select a random sample of agricultural workers at each site. In addition to the training, they reported that a statistician conducts on-site review of methods used by the field interviewers to select the workers. The explicit documentation of methods used to select the worker sample and the quality assurance review by JBS staff is not available.

3.3 Stratification

The geographic strata comprise the regions shown in Table 1. Twelve strata are typically used (some cycles may use more or fewer strata, depending on anticipated worker counts. In any event, the strata are obtained by combining some of the USDA 17 agricultural regions. By definition, strata are constructed as a partitioning of the primary sampling units (PSUs), so that each unit belongs to one and only one stratum. In the OMB statement, the development of the analysis strata for the variance estimation procedure is described. The text on page 30 reads

For the NAWS, the STRATA are defined as the cycle/region combinations used for the first level of sampling and coded in a variable called dmaregn.

Using cycles as strata in the statistical analyses software is inappropriate because the cycles are selected from the same set of PSUs. Finally, a constant adjustment (the region level post-sampling weight) is used at the stratum level to adjust for missing FLAs in the sample of FLAs (for example 25 percent were missing in cycle 59 conducted in 2008). The reasons for missing

FLAs is unclear but reportedly varies, some may be missing because no employers agreed to participate, some because of no crop activity at that time of year, some may simply be missed in error.

4. RESULTS

4.1 Analysis Weights

The tables comparing estimated averages using old and proposed new weights in Table 2., present the estimates, sampling errors, and design effects (DEFFs) for both sets of weights. First the calculation of both sets of weights were verified through our calculations as described approximately in the OMB Part B and by reviewing the program code JBS used to select the samples and to compute the weights. In this table, no changes were made in basic methodology (that is the first stage sampling weights were not used). No significant or practical differences were found. For some estimates of the sampling errors and of DEFFs exceeded the 5 percent threshold for practical significance, but because these were estimates of the sampling errors and of DEFFs, we deemed this as not a serious problem.

Table 3 similarly compares the two types of weights except for percentage estimates. The practical differences are not large and only then for relatively minor categories.

Table 4 using with and without using first stage probabilities also shows no differences at least for estimating means for the key variables. This is difficult to explain, because of the substantial differences among first-stage selection probabilities within a stratum. This could be caused by the post-sampling adjustment factor, which seems to have substantial effect on the weights.

In Tables 5 and 6, we show the estimates, the standard error (SE), and design effects (Deff) for selected variables for the years 2001 to 2005. Again, no significant or practical differences were found in the survey estimates, although the standard error and the design effects do vary.

4.2 Sampling Weights

Survey data typically require some weighting of individual responses, even when the sample is based on equal probability selection. The design of NAWS has been described as having an objective toward a self weighting sample, but this was found infeasible in practice. Hence, sampling weights are very important for obtaining accurate survey results. These weights must reflect the sample design features including selection probabilities at each stage of sampling.

The sampling design is a stratified, multistage design. The stratification is geographic with one or more states contained in each stratum (usually 12 strata but sometimes 14—obtained by grouping 17 USDA crop areas). A hierarchy of 4 types of units is used in the multi-stage design: FLAs, counties, employers, and workers. The 498 FLAs are multi-county farm labor areas designed to have similar farm labor usage and size. Given this design, the sampling weights are constructed as the reciprocal of the product of the selection probabilities at each of the 4 stages (as described in the text of the OMB statement).

A typical feature of such a design is to select units with probabilities proportional to size (estimated number of final units) at each stage up until the final units (workers in this case). Equal number of final stage units are selected from the next to last stage using simple random sampling; resulting in a near self-weighting sample of final stage units (equal probability of being in the survey).

Two noteworthy departures of NAWS methodology to the process just described are:

1. the sample of 90 first stage units (FLAs) is sub-sampled to yield a cycle sample of approximately 30 FLAs, and
2. the next-to-last stage (the employer sample) is selected with simple random rather than probability proportional to size.

Sub-sampling of the primary sampling units (FLAs) would normally use pps at one of the selection steps and equal probability selection at the other. One result of using pps selection at both stages, the sampling weights become very unequal with the final sample containing a very disproportionate number of large FLAs for a cycle sample. If the inclusion probabilities reflecting both sampling steps are reflected in estimates, they will be essentially unbiased but have large sampling errors.

NAWS methodology, however, ignores the probabilities of selection at the first sampling step in calculating sampling weights, resulting in a potential for a reduction of sampling error but introducing serious potential for bias (which post-sampling weights cannot eliminate because they adjust all sample FLAs in a stratum by the same factor). Basically, ignoring the initial selection probabilities causes the sampling weights to produce inferences to a smaller population (the 90 FLAs), which is composed of a population in which large FLAs are more frequent than in the true population of interest.

The second departure, regarding the use of equal probability rather than pps, does not introduce bias but is a major contributor to the relatively large design effect of the survey (increased sampling error). Equal probability of selection used at the next to last stage requires that all workers for the selected employer be included in the sample. This is not an ideal solution because the unequal weighting is replaced by large clustering effect, which again increases Deff. Because employer information from unemployment insurance files gives some indication of employer size at least for larger employers, we suggest this decision implemented at request of NIOSH beginning in 1999 be reconsidered.

The effects of the current NAWS procedures on the sampling weights are shown in Table 7. In the top panel of this table, the basic sampling weights are computed for the 2006 and 2008 NAWS samples. The basic weight is described in the text as

Sampling weights are calculated as the inverse of the probability of being selected:

$$Wt_i = 1 / \text{prob},$$

where prob = workprob*growprob*counprob*flaprob,

$$\text{with} \quad workprob = \frac{\text{number of workers interviewed at the farm location}}{\text{total number of workers at that location}},$$

$$\text{growprob} = \frac{\text{number of growers interviewed in the county}}{\text{total number of qualified growers in that county}},$$

Because the county and the FLA are selected with pps and WOR, the probability of selections for these stages (counprob and flaporb) is based on the algorithm described in the text.

In Table 7, the sum of the basic sampling weights is nearly 1 million larger when the first stage weights are used for 2006 and more than 1.1 million larger for 2008. For both years and regardless of whether the first stage weights are used, the ranges for the sampling weights are very large. For 2006, the basic weights range from 7.5 to 15,930 without the first stage weight and range from 51 to 16,240 with the first stage weight. For 2008, the range of the basic weights is even larger; from 13 to 124,978 without the first stage weight (the ratio of the largest to the smallest weight is 9,600) and range from 34 to 124,978 with the first stage weight (the ratio is only 3,600).

When point estimates and standard errors are computed using these basic weights, the design effects (a measure of the loss in precision caused by unequal weighting clustering and other factors) range in 2006 from 5.0 to around 25 or 36 (depending on the use of the first stage weights). For 2008, the design effects range from 2.4 to 162 when the first stage weights are not used and 8 to 104 when the first stage weights are used. A design effect of 162 with a sample size of 2,182 interviews implies that the effective sample size for this estimate is approximately 13.5 interviews (2,182 interviews divided by 162). The size of these design effects imply that some estimates may be unreliable. For comparison, values for the Deff of more than 2 or 3 are considered large for many surveys.

On Table 8, we show a summary of the year weight (labeled as PWTYCRD in the text of the OMB statement) and post-sampling adjustment factor (labeled as PWTYCR in the text) for 2006 and 2008 NAWS. The year weight is normalized to sum to the sample size for the year, but the variation shown in the basic weights is still apparent in the year weight. The 2006 year weights range from 0.0183 to 21.3 when the first stage weights are not used (the ratio of the largest to the smallest weight is more than 1,160, which substantially less than that for the basic weight) and for the 2008 NAWS from 0.0263 to 21.2 when the first stage weights are used. The design effect from unequal weights and the design effects computed for the key estimates are also substantially lower than those using the basic weight. The reduction in these design effects are caused by the post-sampling weights.

4.3 Post-Sampling Weights

Post-sampling weights referred to in NAWS documents is essentially what is often used to adjust sampling weights to agree with external totals such as the USDA based estimate of total number of crop workers; this step is often referred to as post-stratification. This adjustment in this case, the estimate of number of crop workers in the cycle and region are adjusted from full time workers to number of workers on the basis of average workday adjustments. The post-sampling weights adjustment is the ratio of the distribution of the estimate of number of crop workers in the region for a cycle based on USDA data to the weighted distribution of the estimated number of crop workers in the region for a cycle based on normalized survey weights.

Typically, post-stratification adjustments are relatively small, often only cosmetic, because sampling weights alone should produce good estimates of these control totals. However, because the initial sampling probabilities relating to selecting the panel of 90 FLAs is omitted from the NAWS sampling weights, estimates are driven by the adjustment factors. This region-level adjustment for NAWS is relied on to account for missing sample FLAs and counties and omission of a major sampling probability in the sampling weights. This use of post-stratification adjustments raises serious concern for bias in results, not addressed by proposed new methodology described on OMB Part B. Utilizing the missing sampling probabilities could relieve much of the concern. Also, either initial or final FLA selection should be equal probability sampling.

The effects of the post-sampling weights are shown in Table 8. As noted previously, normalizing the weights to the sample size results in a reduction of the absolute size of the weights, and will not change the design effect from unequal weights. Therefore, the range of weights is still very large from 0.018 to 21.33 and 0.026 to 21.21 for the 2006 weights and 0.0115 to 29.62 and .0137 to 25.70 for the 2008 weights when the first stage weight is or is not ignored. The post-sampling weight adjustments decrease the magnitude of the design effects from unequal weighting (as can be seen by comparing Deff values in Table 7 and Table 8). The design effect from unequal weighting after post-sampling adjustments is 35 percent smaller than the Deff of the basic sampling weights in 2006 (Deffs of 4.29 and 2.77, respectively) and 82 percent smaller for 2008 weights when the first stage weights are ignored (Deffs of 28.9 and 4.94, respectively). When the first stage weights are used, the design effect from unequal weighting after post-sampling adjustments is a 18.5 percent smaller in 2006 (Deffs of 3.36 and 2.74, respectively) and a 75 percent smaller for 2008 weights (Deffs of 18.0 and 4.5, respectively). This implies that the post-sampling weights are performing a substantial smoothing of the weights at a region level and decreasing the effect of the outlier weights. This weight trimming is not explicitly controlled for in the NAWS computational methods.

In the bottom half of Table 8, we present the adjustment factors for the year weight. These factors are essentially a combined adjustment that includes a post-stratification adjustment and the nonresponse adjustments. A post-stratification adjustment is used to adjust probability based survey estimates to known control totals (that is to fine-tune probability-based estimates) and the range of values is generally between 0.90 and 1.10. For nonresponse adjustment factors, a response rate near 75 percent results in an adjustment factor near 1.33 and an adjustment factor of 2.0 implies a response rate of 50 percent or less. For the NAWS, the average post-sampling weight adjustment is 7 and 8.8 for the year weights without the first stage weights (for 2006 and 2008 NAWS, respectively) and between 4 and 5 when the first stage weights are used. These adjustments range from 0.58 to 66 for the 2006 NAWS and 0.18 to 107 for the 2008 NAWS when the first weight is not used and from 0.47 to 16 for the 2006 NAWS and 0.22 to 107 for the 2008 NAWS when the first weight is not used. It is our opinion that the magnitude and the variation in these post-sampling factors dominate the weight computations.

cc: Todd Anderson

Attachment A: From the document 1205-0453_NAWS_Supporting_Statement_Part B_1 26 09.docx

Appendix A: Contacting and Selecting Farm Workers

E. HOW TO CHOSE ELIGIBLE WORKERS FOR THE STUDY

Random Selection

As a sample of workers from a Grower/Employer is needed, the workers are to be chosen at random. All eligible workers of the Grower/Employer must have an equal chance of being chosen.

Workers in different areas (locations)

In the fields, it is common that people who have similar characteristics such as gender, age, birth place, type of work, ethnicity, and etc. tend to group together. If this is the case, you should randomly choose a proportional number of workers from each group or the sampling would not be a good representation. **For example:** for a certain Grower/Employer you have 2 crews of employees. One crew is comprised of single, males with an average age of 24 yrs old, and from Oaxaca with about 6 months of residency in the United States. In contrast, the second crew is comprised of single females. If you choose from only the first crew you will not have a good representation of that grower's employees.

Selecting workers located in different areas

If the Grower/Employer informs you that his employees are distributed over two fields (in the same county) use the proportional formula (below #4) to calculate how many from each field you need to interview. The same proportional formula should be used if you locate workers in different residencies. **For example**, if the workers live in two different labor camps or housing then find out how many live in each dwelling and calculate proportionately how many you should interview from each dwelling.

Proportional selection of workers

When you find that workers are divided into different areas, randomly sampling from each group will be necessary to maintain equal likelihood of selection for everyone. The following formula serves as a guide to calculate the number of workers that should be selected when you find that workers are divided into different areas. In this example, there are 3 fields and you are allowed to conduct 12 interviews for this grower.

a	**b**	**c**
Number of workers per location	**Number of workers per location ÷ Total of workers**	**%X# total of interviews = 12**
Field A = 20	20 ÷ 30 = 66.6%	.666 **x** 12 = 08 interviews
Field B = 05	05 ÷ 30 = 16.6%	.166 **x** 12 = 02 interviews
Field C = 05	05 ÷ 30 = 16.6%	.166 **x** 12 = 02 interviews
Workers total = 30		**Total = 12 interviews**

Once you have determined the number of workers to be selected, identify the correct sampling interval. For example, If five workers are to be selected from a crew of 15, then select workers in intervals of three –every third worker. Count off the workers in order, e.g., from right to left or front to back, and select every third worker.

TABLE 1

NAWS SAMPLING REGIONS
(14 REGIONS USED FOR SOME CYCLES)

NAWS Sampling Regions	USDA Regions (Code & Name)		States in USDA Region
AP12	AP1	Appalachian I	NC, VA
	AP2	Appalachian II	KY, TN, WV
CBNP	CB1	Corn Belt I	IL, IN, OH
	CB2	Corn Belt II	IA, MO
	NP	Northern Plains	KS, NE, ND, SD
CA	CA	California	CA
DLSE	DL	Delta	AR, LA, MS
	SE	Southeast I	AL, GA, SC
FL	FL	Florida	FL
LK	LK	Lake	MI, MN, WI
MN12	MN1	Mountain I	ID, MT, WY
	MN2	Mountain II	CO, NV, UT
MN3	MN3	Mountain III	AZ, NM
NE1	NE1	Northeast I	CT, ME, MA, NH, NY, RI, VT
NE2	NE2	Northeast II	DE, MD, NJ, PA
PC	PC	Pacific	OR, WA
SP	SP	Southern Plains	OK, TX

Source: Part B of NAWS OMB statement, January 26, 2009

TABLE 2

IMPACT OF USING CURRENT AND PROPOSED NEW WEIGHTS FOR ESTIMATED SURVEY MEANS

Estimate	Using Current (Old) Weights	Using Proposed (Revised) Weights	Difference (Percent)
Average Family Income	9.46	9.48	0.15
Standard Error	0.20	0.19	-3.16
DEFF	12.06	11.38	-5.62
Number of children with parent	2.26	2.25	-0.18
Standard Error	0.05	0.05	-0.63
DEFF	2.76	2.74	-0.63
Average Highest Grade	7.57	7.56	-0.12
Standard Error	0.26	0.25	-2.50
DEFF	14.17	13.37	-5.65
Number of children under 18 with Household	0.73	0.71	-2.11
Standard Error	0.06	0.06	-3.24
DEFF	10.38	9.89	-4.77
Average Age	34.96	35.00	0.10
Standard Error	0.89	0.89	0.40
DEFF	16.66	16.71	0.30
Average Work Days/year	188.7	188.6	-0.05
Standard Error	7.98	7.99	0.14
DEFF	24.86	24.89	0.10
Average US stay, years	11.49	11.52	0.28
Standard Error	0.97	0.97	0.97
DEFF	21.79	21.98	0.97
Average Wage	7.98	8.05	0.86
Standard Error	0.11	0.12	9.88
DEFF	7.23	8.10	11.95

TABLE 3

IMPACT OF USING CURRENT AND PROPOSED NEW WEIGHTS FOR ESTIMATED SURVEY PERCENTAGES

	Using Current (Old) Weights	Using Proposed (Revised) Weights	Difference (Percent)
Country of Birth (major categories)			
United States	22.02	21.89	-0.61
Standard Error	4.02	3.93	-2.11
DEFF	35.24	33.92	-3.74
Puerto Rico	0.59	0.56	-4.94
Standard Error	0.25	0.24	-2.47
DEFF	3.92	3.92	0.04
MEXICO	74.16	74.49	0.45
Standard Error	4.34	4.22	-2.88
DEFF	36.88	35.08	-4.86
CENTRAL AMERICA	2.65	2.48	-6.16
Standard Error	0.85	0.76	-10.55
DEFF	10.51	8.95	-14.87
Current Status			
Citizen	25.18	24.97	-0.81
Standard Error	3.99	3.91	-1.98
DEFF	31.43	30.37	-3.40
Green Card	21.00	21.12	0.60
Standard Error	2.12	2.13	0.67
DEFF	10.08	10.17	0.90
Other Work Authorization	0.81	0.75	-6.51
Standard Error	0.22	0.20	-8.87
DEFF	2.22	1.97	-11.22
Unauthorized	53.02	53.15	0.25
Standard Error	4.18	4.16	-0.65
DEFF	26.19	25.86	-1.27

TABLE 4

IMPACT OF USING INITIAL SELECTION PROBABILITY IN PROPOSED NEW WEIGHTS FOR ESTIMATED SURVEY MEANS PERCENTAGES

Estimate	Without First Stage Weight	With First Stage Weight	Difference (Percent) [1]
Average Family Income	10.1	10.2	0.25
Standard Error	0.16	0.17	4.48
DEFF	10.2	11.0	7.35
Average Highest Grade	7.77	7.85	1.05
Standard Error	0.27	0.27	0.06
DEFF	25.2	25.2	-0.05
Average Number Kids in HH	0.70	0.70	-0.37
Standard Error	0.05	0.04	-8.44
DEFF	8.2	6.9	-15.72
Average Age	35.1	35.1	0.16
Standard Error	0.66	0.66	-0.39
DEFF	12.8	12.8	-0.62
Average Work Days per Year	194.4	194.9	0.30
Standard Error	5.87	5.54	-5.66
DEFF	18.5	16.63	-9.89
Average US stay (in years)	12.62	12.56	-0.46
Standard Error	0.86	0.82	-4.07
DEFF	24.0	22.3	-7.25
Average Wage	8.64	8.69	0.52
Standard Error	0.14	0.14	-0.48
DEFF	13.8	13.7	-0.56

1 Precentages are based on unrounded estimates.

TABLE 5

IMPACT OF USING INITIAL SELECTION PROBABILITY IN PROPOSED NEW WEIGHTS FOR ESTIMATED SURVEY MEANS

	2001			2002			2003			2004			2005		
Estimate	Without First Stage Weight	With First Stage Weight	Difference (Percent) [1]	Without First Stage Weight	With First Stage Weight	Difference (Percent) [1]	Without First Stage Weight	With First Stage Weight	Difference (Percent) [1]	Without First Stage Weight	With First Stage Weight	Difference (Percent) [1]	Without First Stage Weight	With First Stage Weight	Difference (Percent) [1]
Average Highest Grade															
Mean	7.09	7.20	1.5	7.53	7.62	1.2	7.32	7.44	1.6	7.72	7.90	2.3	7.5	7.5	1.0
SE	0.20	0.22	13.9	0.31	0.33	6.6	0.24	0.26	10.4	0.23	0.26	14.2	0.31	0.35	11.9
DEFF	10.28	13.04	26.8	23.33	26.10	11.9	14.68	17.44	18.7	10.93	13.88	27.0	9.92	11.97	20.7
Number of children under 18 with Household															
Mean	0.63	0.63	0.1	0.67	0.67	0.3	0.76	0.77	1.4	0.74	0.71	-3.2	0.72	0.70	-3.1
SE	0.07	0.07	1.7	0.05	0.06	9.4	0.07	0.08	25.5	0.06	0.06	2.5	0.08	0.08	-4.1
DEFF	10.71	11.10	3.6	6.47	7.65	18.3	9.73	14.17	45.7	7.98	8.81	10.4	11.17	10.64	-4.8
Average Age															
Mean	32.9	32.9	-0.1	33.4	33.4	0.1	33.6	33.3	-0.8	34.3	34.1	-0.4	35.3	35.3	0.2
SE	0.89	0.86	-3.2	0.59	0.62	6.0	0.60	0.57	-5.1	0.61	0.66	8.0	0.98	0.99	0.8
DEFF	15.50	14.79	-4.6	7.78	8.65	11.3	8.55	7.78	-9.0	7.24	8.30	14.6	12.40	12.31	-0.7
Average Work Days/year															
Mean	174.1	173.1	-0.6	175.2	175.3	0.1	169.1	167.1	-1.2	183.1	183.9	0.4	183.6	183.2	-0.2
SE	6.71	6.97	3.9	6.44	6.37	-1.1	6.70	6.51	-2.8	5.15	5.77	12.1	8.51	8.50	-0.1
DEFF	16.34	17.41	6.6	13.83	13.50	-2.4	16.48	15.50	-5.9	8.78	11.01	25.4	17.52	17.16	-2.0
Average US stay, years															
Mean	9.45	9.50	0.5	10.31	10.32	0.1	10.39	10.28	-1.0	11.04	10.89	-1.4	11.5	11.6	1.4
SE	0.86	0.89	4.1	0.65	0.69	5.9	0.52	0.56	7.0	0.68	0.67	-1.8	0.96	1.01	5.6
DEFF	18.11	19.40	7.1	10.77	11.86	10.1	7.58	8.71	15.0	10.06	9.82	-2.4	13.18	14.85	12.6
Average Wage															
Mean	7.21	7.24	0.4	7.33	7.38	0.7	7.50	7.54	0.6	7.79	7.83	0.5	7.90	7.95	0.7
SE	0.12	0.13	8.9	0.16	0.15	-1.6	0.17	0.18	4.4	0.14	0.15	2.3	0.16	0.21	32.1
DEFF	9.35	10.74	14.9	17.83	17.39	-2.5	17.48	19.21	9.9	12.16	12.68	4.3	8.39	13.73	63.5

1 Precentages are based on unrounded estimates.

TABLE 3
IMPACT OF USING INITIAL SELECTION PROBABILITY IN PROPOSED NEW WEIGHTS FOR ESTIMATED SURVEY MEANS PERCENTAGES

	2001			2002			2003			2004			2005		
	Without First Stage Weight	With First Stage Weight	Difference (Percent) [1]	Without First Stage Weight	With First Stage Weight	Difference (Percent) [1]	Without First Stage Weight	With First Stage Weight	Difference (Percent) [1]	Without First Stage Weight	With First Stage Weight	Difference (Percent) [1]	Without First Stage Weight	With First Stage Weight	Difference (Percent) [1]
Country of Birth (major categories)															
United States															
Mean	19.9	21.5	8.2	26.4	27.0	2.4	24.3	25.5	4.9	23.4	25.5	8.9	22.2	23.0	3.2
SE	3.43	3.75	9.3	4.55	4.75	4.3	3.59	3.88	8.3	3.77	4.12	9.4	4.40	4.75	8.0
DEFF	22.98	25.89	12.7	35.90	38.48	7.2	25.0	28.4	13.7	24.13	27.25	12.9	24.87	28.34	14.0
Puerto Rico															
Mean	0.60	0.59	-0.4	0.30	0.36	19.4	0.62	0.92	48.1	1.14	1.47	28.5	0.29	0.29	-0.1
SE	0.31	0.28	-8.0	0.16	0.18	13.3	0.56	0.86	52.9	0.57	0.85	50.4	0.20	0.21	5.8
DEFF	5.00	4.25	-15.0	2.67	2.87	7.6	18.5	29.3	58.4	8.67	15.30	76.5	3.06	3.42	12.0
Mexico															
Mean	76.8	75.4	-1.8	71.2	70.3	-1.2	71.9	70.5	-1.9	69.5	67.2	-3.4	74.2	73.2	-1.3
SE	3.56	3.85	8.0	4.66	4.86	4.3	3.71	4.25	14.7	4.56	5.02	10.0	4.53	4.87	7.7
DEFF	22.1	24.8	12.1	35.6	38.0	6.9	24.4	31.1	27.8	30.0	34.8	16.2	23.8	27.0	13.3
Central America															
Mean	2.0	1.7	-15.5	1.8	1.8	-3.4	2.92	2.77	-5.2	5.2	5.2	-0.5	2.7	2.9	8.5
SE	0.85	0.68	-20.1	0.61	0.58	-5.6	1.50	1.47	-1.5	1.96	2.12	8.5	1.05	1.09	3.8
DEFF	11.60	8.73	-24.7	7.05	6.50	-7.7	28.3	28.9	2.1	23.60	27.94	18.4	9.40	9.36	-0.4
Current Status															
Citizen															
Mean	22.8	24.4	7.0	29.2	29.8	2.1	28.7	30.1	4.9	26.4	28.7	8.7	25.0	25.7	2.7
SE	3.57	3.86	8.0	4.39	4.58	4.5	3.71	4.23	14.0	3.76	4.24	12.8	4.42	4.75	7.6
DEFF	22.20	24.74	11.4	31.02	33.47	7.9	23.85	30.14	26.4	22.10	26.70	20.8	23.05	26.25	13.9
Green Card															
Mean	21.6	21.3	-1.4	20.5	20.2	-1.2	21.7	21.2	-2.4	24.5	23.3	-4.8	21.3	21.5	0.8
SE	2.17	2.38	9.9	2.00	2.09	4.6	2.32	2.42	4.4	3.25	3.35	2.9	2.27	2.39	5.4
DEFF	8.52	10.40	22.0	8.18	9.03	10.4	11.25	12.47	10.9	17.35	19.02	9.6	6.78	7.49	10.4
Other Work Authorization															
Mean	0.52	0.52	-0.6	0.62	0.58	-6.7	1.42	1.30	-8.3	2.09	2.19	4.8	0.83	0.76	-8.7
SE	0.16	0.17	3.3	0.21	0.19	-10.5	0.59	0.57	-2.7	0.80	0.97	21.2	0.25	0.24	-5.0
DEFF	1.5	1.6	7.4	2.4	2.1	-14.2	8.8	9.1	3.1	9.5	13.4	40.4	1.7	1.6	-1.1
Unauthorized															
Mean	55.0	53.7	-2.4	49.7	49.4	-0.7	48.2	47.4	-1.6	47.0	45.8	-2.6	52.8	52.1	-1.4
SE	3.78	3.97	4.9	3.63	3.68	1.5	3.04	3.40	11.9	3.32	3.52	6.1	4.76	5.09	7.0
DEFF	17.74	19.42	9.5	17.58	18.10	3.0	13.14	16.49	25.5	13.43	15.17	13.0	20.15	23.05	14.4

[1] Precentages are based on unrounded estimates.

Note: The standard error of an estimate is denoted by SE and the design effect from all sources is denoted by DEFF.

TABLE 7

COMPARISION OF BASIC SAMPLING WEIGHT FOR 2006 and 2008 NAWS
WITH AND WITHOUT FIRST STAGE WEIGHTS

	Basic Weight (NEW_WT) [1]	
Measure	Without First Stage Weight	With First Stage Weight
2006 (1,519 Respondents)		
Sum of Weights (Estimate of Number of Worker)	1.49 million	2.32 million
Average Weight	978	1,525
Minimum Size	7.5	51
Maximum Size	15,930	16,240
Standard Deviation	1,772	2,344
Coefficient of Variation (Percent) [2]	181	154
Design Effect From Unequal Weights [3]	4.29	3.36
Design Effects for Key Estimates [4]	5.04 to 36.07	5.88 to 24.72
2008 (2,182 Respondents)		
Sum of Weights (Estimate of Number of Worker)	3.60 million	4.73 million
Average Weight	1,651	2,167
Minimum Size	13	35
Maximum Size	124,978	124,978
Standard Deviation	8,724	8,934
Coefficient of Variation (Percent) [2]	528	412
Design Effect From Unequal Weights [3]	28.91	18.00
Design Effects for Key Estimates [4]	2.41 to 161.89	8.10 to 104.31

[1] Based on the weights for 1,519 respondents from the 2006 survey and 2,182 from the 2008 survey before normalization to sum to number of interviews.

[2] Coefficient of variation is the standard deviation of the weights divided by the average weight.

[3] Design effect from unequal weights is an estimate of the effect on the sampling variance caused by the variation of weights.

[4] The design effect is a measure of the increase in the variance of an estimate caused by the sampling design and reflects the effects of the variation in the weights and also the clustering the sample within counties and employers and stratification.

TABLE 8

COMPARISION OF YEAR WEIGHT AND YEAR WEIGHT POST-SAMPLING ADJUSTMENT
FOR 2006 and 2008 NAWS WITH AND WITHOUT FIRST STAGE WEIGHTS

Measure	2006 (1,519 Respondents) Without First Stage Weight	2006 (1,519 Respondents) With First Stage Weight	2008 (2,182 Respondents) Without First Stage Weight	2008 (2,182 Respondents) With First Stage Weight
Year Weight (PWTYCRD)				
Sum of Weights (Normalized to Sample Size)	1,519	1,519	2,182	2,182
Average Weight	1.145	1.145	0.786	0.786
Minimum Size	0.0183	0.0263	0.0115	0.0137
Maximum Size	21.33	21.21	29.62	25.70
Standard Deviation	1.523	1.509	1.562	1.477
Coefficient of Variation (Percent) [2]	133	131	199	187
Design Effect From Unequal Weights [3]	2.77	2.74	4.94	4.50
Design Effects for Key Estimates [4]	3.72 to 14.02	3.38 to 13.70	12.75 to 28.47	11.29 to 22.63
Year Weight Post-sampling Adjustment (PWTYCR)				
Average Adjustment	7.05	4.17	8.81	4.91
Median Adjustment	3.22	2.81	4.64	3.87
Minimum Adjustment	0.58	0.47	0.18	0.22
Maximum Adjustment	65.5	15.5	106.6	56.69
Standard Deviation	11.3	3.4	13.2	5.1
Coefficient of Variation (Percent) [2]	161.2	82.4	149.6	103.6

[1] Based on the weights for 1,519 respondents from the 2006 survey and 2,182 from the 2008 survey before normalization to sum to number of interviews.

[2] Coefficient of variation is the standard deviation of the weights divided by the average weight.

[3] Design effect from unequal weights is an estimate of the effect on the sampling variance caused by the variation of weights.

[4] The design effect is a measure of the increase in the variance of an estimate caused by the sampling design and reflects the effects of the variation in the weights and also the clustering the sample within counties and employers and stratification.

The following are brief biographical summaries of the authors of the memo.

Daniel Kasprzyk (Ph.D., Mathematical Statistics, George Washington University) is Vice President and Managing Director of Surveys and Statistics. He is responsible for overseeing the statistical staff and the Washington DC survey research staff in Mathematica's Survey and Information Services Division. He is project director for statistical consultation projects that assist the National Center for Education Statistics, the Energy Information Administration, and the Internal Revenue Service. Dr. Kasprzyk has over 25 years experience developing and managing large-scale sample surveys and methodological research associated with federal survey programs.

Dr. Kasprzyk played a significant role in the development of the Survey of Income and Program Participation (SIPP). In the Income Survey Development Program that preceded the SIPP, Dr. Kasprzyk planned the pilot studies, researched the use of administrative records as sampling frames, and directed extramural research in the areas of imputation, longitudinal imputation, and statistical matching. In the early years of the SIPP, he was program manager for the SIPP methodological research program, an effort to improve SIPP questionnaire design, data collection, survey design and estimation. He also served as liaison to the federal statistical agencies and the academic research community in matters relating to survey research and content. After the SIPP had been in the field for several years, he managed the Census Bureau committees involved with survey operations, research, data users and data products, and he supervised the documentation and dissemination of the SIPP longitudinal public-use research file. He organized and was liaison to the American Statistical Association/Section on Survey Research Methods Working Group on the Technical Aspects of the SIPP. Dr. Kasprzyk also initiated the National Academy of Sciences evaluation of the SIPP and served as the Census Bureau's liaison to the expert panel.

In his most recent government position as Program Director at the National Center for Education Statistics (NCES), he was responsible for the Schools and Staffing Survey system, a system of sample surveys of schools, principals, teachers, and school districts. He also directed other NCES sample survey programs and projects, including the National Household Education Survey, the National Study of Postsecondary Faculty, and the NCES Fast Response Survey System, private school frame development, and developed and implemented methodological studies to assist NCES in understanding and improving its data.

Dr. Kasprzyk has experience with a broad array of surveys and nonsampling error issues in surveys. He contributed to the design and implementation of technical documentation, including the SIPP Quality Profile, and two editions of the Schools and Staffing Survey Quality Profile. He is currently directing the preparation of a quality profile for the NCES' National Household Education Survey and is a senior advisor on the preparation of a "design and methodology" report for the American Community Survey staff. He served on the National Science Foundation's Board of Overseers for the Panel Study of Income Dynamics and as a reviewer of the grant that funded the Health and Retirement Study. He was chief editor of *Panel Surveys* (published by John Wiley and Sons, 1989) and was a twenty year member of the Office of Management and Budget's Federal Committee on Statistical Methodology, where his last project addressed sources of error in surveys (chair of committee and editor of *Statistical Policy*

Working Paper 31: Measuring and Reporting Sources of Error in Surveys). He serves as an Associate Editor for the *Journal of Official Statistics* and *Survey Methodology*. Dr. Kasprzyk is a Fellow of the American Statistical Association (ASA), an appointee to the American Statistical Association's Census Bureau Advisory Committee, an elected Vice–President of the American Statistical Association, and has been an officer in several sections of the ASA as well as the Washington Statistical Society.

Frank Potter, Ph.D. (Biostatistics, University of North Carolina at Chapel Hill), a senior fellow at Mathematica Policy Research (Mathematica), specializes in the design and implementation of probability surveys of people, program participants, and health professionals and the implementation of such statistical tasks as weight adjustment, imputation of missing data procedures, and data analysis. Dr. Potter is the Senior Statistician on Feeding America's 2009 Hunger in America study to examine service adequacy and reasons for going to emergency food providers. He is responsible for sample selection and weighting for this multi-stage sample of more than 60,000 clients selected from 185 food banks and their more than 30,000 emergency food service providers. Dr. Potter was also the Senior Statistician on Feeding America's 2006 Hunger in America. Dr. Potter is currently a senior statistical consultant for the evaluation of the Social Security Administration's (SSA's) Ticket to Work program, a program to increase access to, and the quality of, rehabilitation and employment services for those receiving disability benefits. He worked on the sample design and the selection of adult beneficiaries and program participants using the SSA data bases and the weighting, imputation, and estimation activities for the initial survey years and been a senior statistical consultant for the subsequent survey years.

For Mathematica's evaluation of the State Children's Health Insurance Program (SCHIP), he was lead statistician and directed the selection (using both single stage and multistage designs) of more than 25,000 children who were currently or previously enrolled in this program in 10 states. Dr. Potter was also the senior statistician for the design and implementation of SSA's National Survey of Children and Families; a multistage survey to collect information on 9,000 children with disabilities who were receiving or have applied for Supplemental Security Income. He directed the sample selection and the weighting, estimation, and imputation activities. Previously, Dr. Potter managed the sampling, weighting, imputation, and documentation activities for the 1995 National Survey of Family Growth, Cycle V (NSFG), a national survey of women ages 15 to 44 to obtain information on fertility and family-planning practices (sponsored by the National Center for Health Statistics). He developed and implemented sample selection procedures for a sample of 14,000 women from the National Health Interview Survey. He directed the development of the variance estimation procedures, the imputation procedures, and the weight computation activities.

Dr Potter has served as a referee for various journals including Journal of Official Statistics, Journal of General Internal Medicine, American Journal of Epidemiology, Health Care Financing Review, and Medical Care as well as a reviewer of grant proposals submitted to National Science Foundation.

Stephen R. Williams is a senior sampling statistician with over 35 years experience in government and private sector research. His technical and managerial experience spans his positions as Assistant Director of the Center for Research in Statistics at the Research Triangle Institute, Senior Economist at Southern Research Institute, Mathematical Statistician at the U.S. Department of Agriculture (USDA), and most recently as Senior Statistician for Mathematical Policy Research. Recent federal clients include the Environmental Protection Agency, Centers for Disease Control, and the Department of Housing and Urban Development. Recent non-profit clients include the Electric Power Research Institute and the Center for the Studying Health System Change. Mr. Williams provided sampling and statistical support on a national survey of air quality in Canada and on health and risk-behavior studies in the U.S. (including a national survey of young adults and a study of HIV infection in the general population in the National Household Seroprevalence Survey). Mr. Williams has reviewed USDA survey and estimation methodology in response to GAO request: "Statistical Review of Survey Methodology and Estimation of the Statistics Unit of Economics, Statistics, Cooperative Service" and more recently reviewed OMB clearance applications by Regional Educational Laboratories to the U.S. Department of Education.

His focus has been in project management and in the statistical tasks of these research projects. These efforts have permitted him to apply concepts in sampling theory learned at USDA and in the graduate programs at Iowa State University and Universities of Florida and Alabama and to keep abreast of and develop new methods. He had a primary role in developing a method for sampling multiple domains with specified precision and a method for random-digit-dialing with known probabilities in telephone surveys (both presented at ASA conferences). He has in-depth experience with these techniques as well as others in design optimization and data imputation that have been used routinely in the numerous and varied projects in his background.