

APPENDIX D

POINT-BY-POINT RESPONSES TO QUESTIONS RAISED BY BLS REVIEWER ON DECEMBER 7, 2006

The following questions/comments were made regarding Part B (pages 12-15) of the supporting statement for the Florida Agricultural Worker Survey (FAWS):

Question 1: In reference to the table below (page 12), the reviewer asked “*Are the number of farms and workers here, the number sampled or the number of respondents?*”

1. Description of Universe and Sample

a) Universe

Entity	Universe	Sample
Crop Production Areas	9	9
Farms	8,088	231
Crop Workers (estimated)	60,000	1,850

Response to Question 1: 231 is the estimated number of farms that will need to be approached (sampled) and invited to participate in the survey. Participation occurs when the operator or manager or the farm establishment allows interviewers to approach his/her workers in order to invite those workers to participate in the survey. 1,850 is the target number of workers to be interviewed. The estimated number of Crop Workers to be sampled is 2,055.

Assuming that 1) 80 percent of farms will cooperate, 2) 90 percent of randomly selected workers will agree to be interviewed, and 3) about ten workers will be interviewed per farm, approximately 231 farms will need to be approached and 2,055 workers will need to be invited to be interviewed to obtain 1,850 farm worker interviews on 185 participating farms.

Question 2: In reference to the random selection of employers in each of the three commodities (page 12), the reviewer commented “*I could find no mention of how the sample is allocated among the three commodities.*”

Response to Question 2: As noted above, approximately 1,850 workers will be interviewed on 185 farms. The distribution of the 185 farms by crop category will be proportional to the number of Florida farms by crop category as reported in the 2002 Census of Agriculture. The 2002 Census of Agriculture population of producers with proposed sample sizes are as follows:

Commodity	Florida Farms ^a	Florida Farms with ≥ 500 acres ^a	Proposed Employer Sample Size
Citrus	7,653	254	125
Tomatoes	218	23	35
Strawberries	217	17 ^b	25

^a U.S. Department of Agriculture, NASS, 2002 *Census of Agriculture*.

^b Data unavailable for farms with ≥ 500 acres; number refers to farms with ≥ 100 acres.

Although there are nearly the same number of tomato and strawberry farms in Florida, tomato farms will be oversampled to account for their larger farm size and greater geographical dispersion.

Question 3: In reference to the selection of workers (page 12), the reviewer commented “*The selection procedure should be explained in this document, rather than referring to another document or at least the other document should be included in the package.*”

Response to Question 3: The selection procedures are discussed in new Appendix D “FAWS INTERVIEWER’S INSTRUCTIONS FOR CONTACT AND SELECTION OF GROWERS AND WORKERS,” which is attached here and will be attached to the final package.

Question 4: In reference to the number of interviewees per region (page 13), the reviewer asked “*Is this selected employers or selected workers.*”

Response to Question 4: Interviewees denotes farm worker respondents.

Question 5: In reference to interviewer sampling instructions (page 13), the reviewer commented “*The instructions should be described here.*”

Response to Question 5: The interviewers sampling instructions are discussed in new Appendix D “FAWS INTERVIEWER’S INSTRUCTIONS FOR CONTACT AND SELECTION OF GROWERS AND WORKERS,” which is attached here and will be attached to the final package.

Question 6: In reference to inflating worker data (page 13), the reviewer commented “*Neither the weighting or the estimation procedures are explicitly stated....in each region.*”

Response to Question 6: The survey uses a stratified multistage sample and corresponding complex sample estimation methods to generate unbiased population estimates for the FAWS population and measures of the precision of those estimates. The weighting scheme is designed to allow population estimates to be made for either a

specific crop's workforce (since most of the presentation of the data will be in terms of descriptive statistics and modeling) or for the entire FAWS population if desired.

Within the complex sample, there are quotas for interviews by crop. For each crop, a multistage subsample is drawn that is proportional to payroll size. Regions have interview allocations proportional to total farm payroll in each crop, with farms selected with probabilities proportional to payroll and the number of workers interviewed is proportional to payroll (or in some cases proportional to the square root of payroll). This PPS sample is designed to be self weighting within crop, such that each worker has an equal chance of selection. However, data limitations make this sample design difficult to achieve in practice. For example, the payroll data used to sample is from previous years and current employment data is not available until after the interviews are completed. Thus, post-sampling weight adjustments are made to correct for any inaccuracies or systematic departures from the sampling design, ensuring that the interviews correctly represent the labor force at the time of interview.

Other small deviations from the sampling plan make it necessary to implement post-sampling weights. These deviations include discrepancies between the number of interviews allocated and completed, and the unequal probabilities of finding part-time versus full-time workers (probabilities which can only be established after interviews are conducted). Post-sampling weights, therefore, are used to adjust the relative value of each interview so that correct estimates can be obtained from the sample.

The post-sampling weighting scheme is composed of several components, which are multiplied together. The first component of weight (week) reflects the probability of finding respondents who have workweeks of differing lengths (part-time versus full-time) and is the inverse of the number of days worked divided by the average work week. This gives higher weights to part-time workers who have a lower probability of being sampled (defined as the number of days of the interviewing week that they were available). The next two components of weight (region, crop) reflect the relative importance of a region and a time of year. The worker component of weight is proportional to the amount of payroll within the given crop using the 2007 employment figures. In some instances, it may be important to combine information from the three crop subsamples. To combine worker interviews from the three crops, workers are given a crop weight which represents the value of the payroll of the crop in which they worked relative to the total of the three crops' payroll. So, if citrus makes up 50 percent of payroll, the sum of the weights of citrus workers would make up 50 percent of the weight of the combined interviews. The last component (season), accounts for the different probabilities of workers who work different amounts of time during the year. This component of weight is calculated from the work grid and is the inverse of number of months during the sampling period that the worker worked divided by the total month duration of the interviewing season (approximately four months). This compensates for the fact that short-term workers are less likely to be sampled than workers who are employed throughout the whole interviewing season (four months.)

Estimation will be done using well established procedures and statistical software for working with complex samples. These methods assure the production of unbiased population estimates as well as their sampling variability or precision.

Question 7: In reference to the discussion on non-response adjustment (page 15), the reviewer commented *“Although this subsection is entitled nonresponse adjustment, only response rates are discussed, not a method for adjusting for nonresponse.”*

Response to Question 7: Non-response in the FAWS can come in two forms. Employers can refuse to participate or workers can refuse to participate. Since the FAWS uses similar methods to the NAWS, it can be assumed that the participation rates in the FAWS will be no worse than those in the NAWS. In fact, employers may be more likely to participate in a survey sponsored by USDA and the University of Florida (UF) than the Department of Labor (DOL) as both UF and USDA have strong grower outreach through the Extension Service whereas agricultural employers often perceive contact with DOL as related to enforcement.

One way that the FAWS will adjust for non-response is by setting allocation targets that exceed the number of interviews required. The usual way that this is done is for the allocation to exceed the required number of responses by the inverse of the expected response rate. This assists with non-response that equally affects all cells of the design.

Another aspect of non-response may be different rates of response in different crops or geographic areas. This kind of non-response is handled by post-sampling adjustment of weights. As discussed earlier, the post-sampling adjusted weights ensure that each interview represents the correct proportion of the population. If response is lower in one region, for example, each interview will receive a higher weight so that the sum of weights for that region represents the correct share of the total weights.

The critical non-response issue is whether or not the non-respondents differ in any key attributes from the respondents. The existing NAWS data are the only available benchmark data on which to gauge potential differences for farm workers. However, the NAWS sample for the three Florida commodities is necessarily small, consequently providing limited information to identify distinctions between FAWS respondents and non-respondents.

Question 8: In reference to reliability (page 15), the reviewer commented *“Neither the degree of accuracy, nor procedures for computing variance estimates and inference methods are detailed.”*

Response to Question 8: Due to the use of a complex sample, both the estimates and the variability of the estimates will be calculated using statistical software that accounts for complex samples. Special procedures are needed since, unless the complex nature of the sample is considered, the variability and precision of the estimates will usually be

incorrect. To evaluate the reliability and precision of the estimates, we will consider both their standard errors and their coefficients of variation.

One measure of the effect of the complex sample is to calculate the design effect, which is the ratio of the squares of the standard errors of an estimate calculated in two ways: taking the complex design of the sample into account (numerator) and treating the sample as though it were a simple random sample. This measure is generally calculated for important estimates. Design effects greater than one show the loss in precision from using a complex sample. Design corrected standard errors can be calculated as well.
