APPENDIX F

**POINT-BY-POINT RESPONSES TO
QUESTIONS RAISED BY BLS REVIEWER ON FEBRUARY 1, 2007**

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 allocation of employers and workers across commodities to the sample, the reviewer noted *"This (how employers are allocated) is not explained. Also, although the above table indicates how the sample of <u>employers</u> is allocated among the commodities, it is not clear how the sample of <u>employees</u> is allocated among the commodities."*

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.

Response to Question 1: The statistical methodology of Part B has been formalized. Allocations of employers and workers to the sample across strata are based on Cochran's (1977 p. 317) result for an optimum sampling rate for each stratum for a given expected cost. The allocations follow from equation (2) on page 14 of the revised document. The allocations are dependent on the employer inclusion probabilities, the within employer variance of earnings, and the relative size of employers.

_____

Question 2: In reference to the employer and worker selection procedures discussed in Appendix D "FAWS INTERVIEWER'S INSTRUCTIONS FOR CONTACT AND SELECTION OF GROWERS AND WORKERS," the reviewer commented *"The new Appendix D seems to describe a quota sample not a probability sample. That is, it does not seem possible to correctly calculate the probability of the employees in a particular sampled farm being interviewed because the quota may be met before the farm is reached*

on the list. Probability sampling is the generally accepted procedure for conducting a survey."

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.

Response to Question 2: As noted in the response to question 1, the methodology for part B of the Supporting Statement has been revised to formally set forth the statistical basis for the sampling procedure. In addition, Appendix D (attached to the revised package) has been revised to be in agreement with the sampling methodology. The employer sample is to be drawn using maximum entropy sampling (Supporting Statement, B.2.Selection of Employers, pp. 12-13); interviewers are to contact *only* the employers drawn in the random sample (Appendix D, I.A.1). Sampling of workers at selected employers is based on a systematic sample based on the sampling frame constructed at the employer's site (Supporting Statement, B.2.Selection of Workers, pp. 13-14). Interviewers are to contact *only* those workers drawn in the systematic sample (Appendix D, II.C).

_____

Question 3: The reviewer noted a lack of clarity in the supporting statement regarding the sampling weights as well as inconsistencies between the supporting statement and the interviewer instructions, *"For all these reasons it is necessary to explain how the base weights, that is the weights before the "post-sampling weighting scheme," are calculated."*

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.

Response to Question 3: The revision of the statistical methodology spells out in detail the sampling and weighting procedure. This is best illustrated by the estimation of total earnings in equation (8) of the supporting statement. Although this is developed at length in the revised supporting statement, the first line of the equation is repeated here to clarify the weights.

$$\hat{Y}_h = \frac{1}{n_h} \sum_{i=1}^{n_h} \frac{M_{hi}}{z_{hi} m_{hi}} \sum_{j=1}^{m_{hi}} y_{hij}$$

The sampling rate for workers in the hth stratum is $f_{0h} = \dfrac{\pi_{hi}}{k_{hi}}$ where the numerator is the inclusion probability for the ith employer and the denominator denotes that every k^{th} worker is selected in a systematic sample, the inverse of the worker sampling rate for the ith employer. The inclusion probability is defined as $\pi_{hi} = n_h z_{hi}$ where n_h is the employer sample size in stratum h, z_{hi} is defined as $z_{hi} = \sqrt{M'_{hi}} \Big/ \displaystyle\sum_{i=1}^{N_h} \sqrt{M'_{hi}}$ and M'_{hi} is the estimated number of employees for the ith employer based on pre-survey QCEW data. Consequently, employers are selected with probability proportional to the square root of estimated size. As shown in the estimate of earnings equation above, employer weights are the inverse of the inclusion probability. The k_{hi} defines the within employer sampling rate, setting the sample size at m_{hi} with M_{hi} employees found at employer i. The within employer worker weight is the inverse of the within employer sampling rate.

The Interviewer Instructions (Appendix D) no longer refer to a maximum number of workers to be interviewed at employers. The number of workers to be interviewed is determined by the *a priori* worker sampling rate given to the interviewer for each employer and the number of workers reported by each employer interviewed. The instructions are consistent with the probability based sample developed in the methodology.

_____

Question 4: Regarding post-sampling weights adjustments for variations in length of work week for workers, the reviewer asked, *"Is the average work week calculated separately by crop and/or region, or is there simply one average work week that applies to the whole 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).

Response to Question 4: The work week post-sample weight is calculated separately for each worker sampled. The adjustment is shown in detail in section B.3.Days per Week Weight Adjustment and B.3.Nonresponse Adjustment starting on page 32 of the supporting statement.

_____

Question 5: Regarding post-sample weights for crop and region, the reviewer asked, *"Is this region and crop or region and time of year?"*

The next two components of weight (region, crop) reflect the relative importance of a region and a time of year.

3

Response to Question 5: The revised post-sample weight adjustments are only for variations in days worked per week. Since this is a survey designed to be conducted at peak employment for a single state and workers in three specific commodities, the region and crop/time of year post-sample adjustments done in the NAWS are unnecessary for the current survey.

Question 6: Regarding the adjustment of post-sampling weights to agree with 2007 employment figures, the reviewer commented, *"I believe this is better understood as a benchmark factor and is typically the last stage of weighting adjustment, which it is not here."*

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.)

Response to Question 6: A formal section on nonresponse adjustment has been added (B.3.Nonresponse Adjustment starting on page 32 of supporting statement). The nonresponse adjustment is a calibration approach adjusting the weights at the final step to the benchmark data in the QCEW as specified in equation (62).

Question 7: Regarding the discussion of nonresponse adjustment, the reviewer commented, *"I would suggest putting this paragraph in the response to the previous question, since that is where post-sampling adjustments are discussed."*

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.

Response to Question 7: As noted in the response to question 6, post-sample weight adjustment is integrated with the formal nonresponse adjustment in the revised

supporting statement. See B.3.Days per Week Weight Adjustment and B.3.Nonresponse Adjustment, page 32-35 of the revised supporting statement.

_____

Question 8: Regarding the degree of accuracy of the estimates, the reviewer commented, *"The degree of accuracy is still not mentioned at all. This is quite important since it is necessary to know if the proposed sample sizes will yield estimates with sufficient accuracy. Also, the procedures for computing variance estimates are not described in sufficient detail."*

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."*

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.

Response to Question 8: The revised supporting statement develops the formal population estimates for the mean of earnings, the variance of the estimates, and procedures for the estimate of the variance given the sample data in B.2.Estimation Procedures - Average Earnings. The design-based standard error of mean earnings for the sample sizes of 244 (184) employers and 1,805 (1,624) workers (expected respondents) is 1.8% of the population mean as indicated in B.2.Accuracy in the revised supporting statement. A 95% confidence interval for the population mean that is no more than 3.6% from the sample mean provides sufficient accuracy for the purpose of the survey.

_____