National Agricultural Workers Survey (NAWS)

Summary of Nonresponse and
Design Studies

Submitted: September 30, 2019

Submitted to:
Daniel Carroll, USDA/DOL/OPDR

U.S. Department of Labor

Employment and Training Administration

Office of Policy Development and Research

Room N-5637

200 Constitution Ave., N.W.

Washington, D.C. 2021

Submitted by:

Susan Gabbard

JBS International, Inc.

555 Airport Boulevard Suite 400

Burlingame, CA 94010

This report was prepared for the U.S. Department of Labor, Employment and Training Administration, Office of Policy Development and Research by JBS International, Inc., under contract 47QRAA18D00AE; Task Contract 1630DC-19-F-00019. Since contractors conducting research and evaluation projects under government sponsorship are encouraged to express their own judgment freely, this report does not necessarily represent official opinion or policy of the U.S. Department of Labor.

Table of Contents

[Introduction 4](#_Toc20733484)

[COMPLETED STUDIES 4](#_Toc20733485)

[Nonresponse Study 1 – NAWS Item Nonresponse Rates 4](#_Toc20733486)

[Analysis 4](#_Toc20733487)

[Results 5](#_Toc20733488)

[Nonresponse Study 2 – NAWS Unit (Employer) Nonresponse 5](#_Toc20733489)

[Analysis 5](#_Toc20733490)

[Results 6](#_Toc20733491)

[Nonresponse Study 3 – Follow Up with Employers Who Were Not Successfully Screened During the Initial NAWS Data Collection 6](#_Toc20733492)

[Analysis 6](#_Toc20733493)

[Results 7](#_Toc20733494)

[Nonresponse Study 4 – Examining Employer Eligibility Over Time and NAWS Response Rates 7](#_Toc20733495)

[Analysis 7](#_Toc20733496)

[Results 7](#_Toc20733497)

[Design Study A – Efficiency of the NAWS Sampling Design 7](#_Toc20733498)

[Analysis 8](#_Toc20733499)

[Results 8](#_Toc20733500)

[Design Study B – Optimal Interview Allocations for NAWS Sampling 8](#_Toc20733501)

[Analysis 8](#_Toc20733502)

[Results 9](#_Toc20733503)

[in progress or potential future studies 9](#_Toc20733504)

[Nonresponse Study 5 – Comparison of the Characteristics of Respondents to National Data 9](#_Toc20733505)

[Nonresponse Study 6 – Comparing Worker Data of Employers Who Change Response States 9](#_Toc20733506)

[Design Study C – Extending the Optimal Allocation Study 10](#_Toc20733507)

[Appendix A: Nonresponse Study 1 – NAWS Item Nonresponse Rates 11](#_Toc20733508)

[Introduction 12](#_Toc20733509)

[Sample 12](#_Toc20733510)

[Data Preparation 12](#_Toc20733511)

[Analysis 12](#_Toc20733512)

[Results 12](#_Toc20733513)

[Conclusions 16](#_Toc20733514)

[Appendix B: Nonresponse Study 2 – NAWS Unit (Employer) Nonresponse 17](#_Toc20733515)

[Introduction 18](#_Toc20733516)

[Sample 18](#_Toc20733517)

[Data Preparation 18](#_Toc20733518)

[Analysis 19](#_Toc20733519)

[Results 20](#_Toc20733520)

[Conclusions 27](#_Toc20733521)

[Appendix C: Nonresponse Study 3 – Follow Up with Employers Who Were Not Successfully Screened During the Initial NAWS Data Collection 28](#_Toc20733522)

[Introduction 29](#_Toc20733523)

[NRFU screening 29](#_Toc20733524)

[Sampling Universe 30](#_Toc20733525)

[Nonresponse follow-up attempts 30](#_Toc20733526)

[Data Preparation and the Analytic Sample 31](#_Toc20733527)

[Analysis 31](#_Toc20733528)

[Results 32](#_Toc20733529)

[Conclusions 33](#_Toc20733530)

[Appendix D: Nonresponse Study 4 – Examining Employer Eligibility Over Time and NAWS Response Rates 36](#_Toc20733531)

[Appendix E: Design Study A – Efficiency of the NAWS Sampling Design 41](#_Toc20733532)

[Method 42](#_Toc20733533)

[Results 44](#_Toc20733534)

[Conclusions 44](#_Toc20733535)

[Appendix F: Design Study B – Optimal Interview Allocations for NAWS Sampling 58](#_Toc20733536)

[Method 59](#_Toc20733537)

[Results 60](#_Toc20733538)

# Introduction

JBS has undertaken a series of small studies to examine possible nonresponse bias in the National Agricultural Workers Survey (NAWS) and to assess the efficiency of the survey’s design, including four nonresponse studies that assessed employer and item response rates, and two design studies that assessed and potentially improved the survey’s design. This document summarizes each of these studies. The full report for each study is attached in Appendices A–F. The final section of this document describes studies that are in progress or are potential future studies.

 The nonresponse studies included:

* Nonresponse Study 1 examined the nonresponse rates among items of the NAWS questionnaire.
* Nonresponse Study 2 examined nonresponse bias by comparing employers who allowed interviews, eligible employers who refused to allow interviews, and employers whose eligibility could not be determined.
* Nonresponse Study 3 attempted further contact with employers who were not successfully screened during the regular data collection cycle to determine whether further contact attempts would result in finding eligible employers who might improve the NAWS response rate.
* Nonresponse Study 4 is a Markov chain analysis that incorporated prior data to examine whether employer’s eligibility (i.e., eligible, ineligible, or unable to be determined) impacts response rates.

The design studies covered the following:

* Design Study A examined NAWS’s sampling design efficiency by using a series of nested ANOVA to look for interactions between levels of sampling and key survey variables.
* Design Study B examined the tradeoffs in the efficiency of interview allocations. Each of these studies are summarized below.

# COMPLETED STUDIES

# Nonresponse Study 1 – NAWS Item Nonresponse Rates

This study examined nonresponse for questionnaire items. Calculating item nonresponse is one of the survey standard outlines in *OMB’s Standard and Guidelines for Statistical Surveys (2006)*. The full report for the item nonresponse study can be found in Appendix A.

## Analysis

Item nonresponse was examined for 86 items on the 2011–2016 NAWS questionnaire that covered all sections answered by the respondents, except for items in the household and work grid. Of the 58 items asked of all respondents, the denominator of the nonresponse rates was the count of respondents. For the 28 items asked only if certain criteria were fulfilled (i.e., having a skip pattern), the denominator was the number of respondents who met the criteria for being asked the question. For both kinds of items, the number of valid responses was the numerator.

## Results

For the 58 items asked of all respondents, across fiscal years 2011 to 2016, the average nonresponse rate was less than 0.5 percent. Certain items had higher nonresponse rates than others. For example, the item “When was the last time your parents did hired farm work in the U.S.A?” had a nonresponse rate of up to 3.5 percent.

For the 28 items with skip patterns, across the years, the average nonresponse rate was less than two percent. The item “Does this employer keep in contact with you about future employment before leaving at the end of the season?” had the highest annual nonresponse rate of up to 9.4 percent.

Overall, the NAWS items showed very low item nonresponse with most items exceeding 95 percent valid answers and a few items having 90–94 percent valid responses. For items with less than 70 percent valid responses, the Office of Management and Budget (OMB) requires additional analysis of item nonresponse. No additional analysis was undertaken since all items exceeded the OMB criteria of 70 percent.

# Nonresponse Study 2 – NAWS Unit (Employer) Nonresponse

This study assessed nonresponse bias by comparing information in the sampling frame on eligible respondents and nonrespondents. While the sampling data is somewhat sparse for nonrespondents, three pieces of information are useful: geographic location, North American Industry Classification System (NAICS) code, and the source used to obtain employer names. The NAWS uses three sources to acquire employer names: a) the Bureau of Labor Statistics (BLS) Quarterly Census of Employment and Wages (QCEW) microdata on employers paying unemployment insurance (UI) taxes, b) marketing lists, and c) internet searches and contacts with knowledgeable local individuals. Geographic area and source lists are available for all employers, while NAICS codes are available for all employers who pay UI taxes, marketing list employers, and some additional employers.  The full report can be found in Appendix B.

## Analysis

This study examined three characteristics (source of the employer list, NAICS, and geography) and made three comparisons:

1. Employers allowing interviews compared to sampled employers that refused or were unable to be screened (i.e., excluding employers who are ineligible).
2. Employers allowing interviews compared to eligible employers who refused.
3. Employers who are eligible compared to employers whose eligibility could not be determined.

Nonresponse bias was calculated using the bias calculation formula from OMB’s *Standard and Guidelines for Statistical Surveys* (2006). The formula defines bias for an estimate, $B\left(\overbar{y}\_{r}\right) $, as the following:

$$B\left(\overbar{y}\_{r}\right)=\overbar{y}\_{r}- \overbar{y}\_{t}= \left(\frac{n\_{nr}}{n}\right) (\overbar{y}\_{r}- \overbar{y}\_{nr})$$

where:

*yt* = the mean based on all sample cases;

*yr*= the mean based only on respondent cases;

$\overbar{y} $ = the mean based only on the nonrespondent cases;

*n*  = the number of cases in the sample; and

*nnr*= the number of nonresponding cases.

## Results

The results show that nonresponse rate for the sources was 83–89 percent, 55–57 percent, and 61–75 percent for comparisons A, B, and C, respectively. Furthermore, there was low bias (2–4 percent) across the three comparisons and sources. There were more variations in nonresponse rates for NAICS, but bias remained low at less than 10 percent. The nonresponse rate for the six regions of NAWS was 80–87 percent, 49–62 percent, and 57–66 percent for comparison A, B, and C, respectively, and bias was low at less than seven percent across the three comparisons. The nonresponse rate for the 12 regions of NAWS was 75–87 percent, 46–63 percent, and 54–66 percent for comparison A, B, and C, respectively, and bias was similarly low (less than 7 percent).

JBS also conducted regression analysis to determine the association between employer characteristics (source, NAICS, and geography) for the three comparisons. The results show that in comparisons A and C, employers selected from the InfoUSA source were significantly less likely to participate compared to BLS sourced employers. In all three comparisons, employers with NAICS 1114 (Greenhouse, Nursery, and Floriculture Production) had the highest likelihood of participating in NAWS, compared to NAICS 1119 (Other Crop Farming). In terms of geography, in all three comparisons, employers in four of the six regions (East, Southeast, Midwest, and Northwest), and 10 of the 12 regions, were significantly more likely to participate in NAWS compared to California.

Overall, the results showed that although unit nonresponse rates were high, there was little nonresponse bias between responding and nonresponding employers overall and across NAICS, sampling regions, and list source.

# Nonresponse Study 3 – Follow Up with Employers Who Were Not Successfully Screened During the Initial NAWS Data Collection

This study is a repeat of the 2009 study where JBS attempted further contact with employers who were not successfully screened during the data collection. NAWS staff made additional attempts via mail and telephone calls to contact 779 unscreened employers from the Fall 2017 data collection cycle to determine their eligibility status. The goal of this additional nonresponse follow-up (NRFU) was to determine whether further contact attempts or lengthening the onsite data collection period would improve the employer response rate. The full report can be found in Appendix C.

## Analysis

The NRFU resulted in 30 percent of previously unscreened respondents responding and being screened. Mail-only response was 17 percent and mail plus telephone response was 35 percent. Respondents in the NRFU study were coded as Eligible, Ineligible, or Unscreened using coding categories similar to the Fall 2017 sample’s response codes. Response rates were calculated for the Fall 2017 Cycle with and without the additional NRFU data using the formula for the unweighted response rate (RRU) from OMB’s *Standard and Guidelines for Statistical Surveys* (2006).

## Results

The initial response rate for the Fall 2017 cycle was 25 percent. The lower bound on the NRFU response rate was calculated assuming all eligible NRFU respondents refused to allow interviews, the response rate then decreased to 20 percent. The upper bound was calculated assuming all NRFU respondents agreed to interviews and was 37 percent. Assuming the same proportion of the NRFU respondents allowed interviews as the original Fall 2017 responding employers, the response rate was 27 percent. The additional NRFU did not substantially change the response rate. This result was the same as the 2009 study. That is, additional effort at finding and screening employers provides information on more employers but does not improve the NAWS response rate.

# Nonresponse Study 4 – Examining Employer Eligibility Over Time and NAWS Response Rates

In 2019, JBS repeated the Markov analysis that was completed in 2007. A small number of agricultural employers appear on the survey’s sampling list in multiple administrations of the survey. Attempts to contact these employers may have had different outcomes at different time periods. This study used Markov chain analysis to incorporate information from prior data periods about employers’ states – whether eligible, ineligible, or unable to be determined – and looked at the impact on response rates. The full report can be found in Appendix D.

## Analysis

The analysis used contact data on agricultural employers contacted from FY 2006–2017 (cycles 53–88). Some employers were contacted in as many as six different cycles, for a total of 34,774 contacts. Each contact was coded in response category 1–8 (Yes, Yes but, Qualified refusal, Don’t know, Incomplete, Not in sample, Skipped, Office codes), and the probability of an employer moving from one of these categories in a particular cycle, to each of the eight possible categories in the next cycle, was found. The study also calculated the expected percentage of employers in each of the eight categories after a large number of cycles.

## Results

The results of the analysis showed a five-percentage point gain in the CASRO response rate from 15 to 20 percent. The overall expected response rate after a large number of cycles is 20 percent.

# Design Study A – Efficiency of the NAWS Sampling Design

To better understand the study’s design effects, JBS’s statistical team at Portland State University conducted analyses that used multivariate analysis to identify whether the survey’s sampling design was efficient. The study used a series of nested ANOVA’s to identify whether there were significant interactions between the various levels of sampling and key survey variables. An efficient design has homogeneous strata and heterogeneous clusters. If this were true, then the analysis should show that the homogeneous strata vary from each other significantly and that the heterogeneous clusters do not vary significantly. The full report on this study can be found in Appendix E.

## Analysis

The ANOVA analyses looked at nine different key variables and their relations to the sampling levels. The key variables were hourly or hourly-equivalent wage, employed by a farm labor contractor, indigenous, unauthorized, number of farm employers, paid hourly or by the piece, number of farm work days, and number of children in the household. The sampling levels were fiscal year, cycle, region, farm labor area, county, Zip Code region, and agricultural employer. The analysis was conducted separately on 2011–2012 data and on 2013–2014 data. A third set of ANOVA analyses was conducted using the combined 2011–2014 data. This analysis examined the current NAWS weight and a proposed change of the employer weight to include a more complex employer nonresponse calculation.

## Results

The analyses on the 2011–2012 data and on 2013–2014 data showed similar findings:

* Region (or the cycle/region interaction) is consistently found to be a significant effect in all except two variables. This suggests that, for most variables, the stratification by geographic location divides farm workers into heterogeneous groups and is, therefore, an effective design strategy.
* Clustering at the county level and employer level are also consistently significant effects, indicating farm workers within one cluster of employers (or county) are significantly different than farm workers in another cluster of employers (or county) for a particular combination of higher-level clustering and stratification. This is not an optimal design element, but likely necessary for efficient data collection.
* Outside of Region, County, and Employer, there is little consistency in significant effects across the sampling level variables.

The analysis on 2011–2014 data using the first more complex employer nonresponse calculation showed little difference between weights.

# Design Study B – Optimal Interview Allocations for NAWS Sampling

The purpose of this study was to see how interview allocations would change if they were optimized for statistical efficiency and/or cost reduction. The current interview allocation is proportional to the distribution of crop workers across geographic areas. The result is that crop worker allocations are concentrated in a small number of sampling regions with large numbers of crop workers, resulting in small allocations and potentially larger variances for estimates in the other regions. The NAWS statisticians calculated optimal interview allocations for each of the three cycles and 12 sampling regions used to stratify the NAWS sample. The goal was to gain more information about how to reduce interviewing costs and improve the precision of point estimates. The full report can be found in Appendix F.

## Analysis

The optimal allocations were calculated for nine variables that are considered key findings from the NAWS:

* The worker’s hourly wage or hourly equivalent wage if a piece rate worker;
* Number of farm employers in the past 12 months;
* Number of farm work days in the past 12 months;
* Number of children in the household;
* The employer was an agricultural producer and not a labor contractor;
* The worker lacked work authorization;
* The worker had only one farm employer;
* The worker was paid an hourly wage as opposed to a piece rate or salary; and
* The number of children in household was three or fewer.

Two types of allocations were calculated. The *optimal allocation* achieved both statistical and cost efficiency. The *Neyman allocation* was a special case of optimal allocation that assumed the cost of each stratum was approximately equal and thus calculated statistical efficiency only.

## Results

The results show that both optimal allocation and Neyman allocation would increase interview allocations in the larger crop labor region in all cycles. Regions with currently small interview allocations would have even smaller allocations if allocations were optimized for statistical and/or cost efficiency.

# in progress or potential future studies

# Nonresponse Study 5 – Comparison of the Characteristics of Respondents to National Data

JBS anticipated comparing the characteristics of respondents to national data on NAICS and geographic distribution separately, and where sample size and data allowed, on NAICS and geographic region combined. While there are no exact matches to the NAWS employer universe in a single Federal data source, it was expected that some comparisons could be made.

The first anticipated comparison was between NAWS NAICS 1151 employers allowing interviews with QCEW data on NAICS 1151 employers. However, the vast majority of the 1151 employers on the list come from the UI microdata which is used to generate the QCEW results. This portion of the study was redundant with analysis done in Nonresponse Study 2.

The second anticipated comparison was between NAICS 111 employers allowing interviews with the 2017 CoA data on farms with hired farm labor. The initial attempt to compare the data sources revealed that a direct comparison is not straightforward. One concern was the difference in the definitions of a hired farm worker between the CoA and the NAWS, particularly the possibility that in some regions the CoA data may include large numbers of family workers that are not eligible for the NAWS. Further examination is planned, including an examination of USDA’s Farm Costs and Returns Survey to better understand family labor on farms.

# Nonresponse Study 6 – Comparing Worker Data of Employers Who Change Response States

This goal of this study is to gain insight into whether workers from eligible growers who refuse to participate in the NAWS are different than workers from employers who consent to participate. While it is not possible to interview workers whose employers refuse, the Markov analysis done in Nonresponse Study 4 allows NAWS staff to look at farm worker data from agricultural employers who were in the survey multiple times and at least once allowed interviews. This study will compare workers with agricultural employers who change categories from allowing interviews to refusing to participate (and vice versa) as well as workers whose employers always allow interviews. The analysis will focus on the key variables used in Design Study B above.

NAWS staff will first identify two groups of agricultural employers: 1) those that have consented at every contact, and 2) those that have sometimes consented and sometimes refused. After reviewing the data, the second group may be further subdivided into those that initially refused and then participated and those that participated and later refused. The analysis will compare groups by analyzing survey responses of the farm workers, using *t*-tests and ANOVA for numerical items and chi-squared tests for categorical items.

# Design Study C – Extending the Optimal Allocation Study

After reviewing the results of the optimal allocation study, ETA asked that the study be extended to looking at the variables used and the numbers of years of data used for NFJP population estimates. A goal of the NAWS is to provide accurate regional estimates for crop workers for calculating three factors that are part of the NFJP population estimate – calculations of NFJP eligibility, time in residence, and annual employment. This study will follow the same methods described for Design Study B above.

# Appendix A: Nonresponse Study 1 – NAWS Item Nonresponse Rates

**NAWS Item Nonresponse Rates**

## Introduction

Item nonresponse looks at nonresponse for questionnaire items. Calculating item nonresponse is one of OMB’s survey standards outlines in ***OMB’s Standard and Guidelines for Statistical Surveys (2006)***. Item nonresponse is calculated as the percent of respondents for whom no valid response was recorded. If that rate is above 30 percent for an item, OMB standards call for additional analysis to identify further the implications of that bias.

## Sample

The sample for the item-nonresponse study consisted of 12,602 agricultural worker interviews from NAWS fiscal years 2011–2016 (cycles 68-85).

## Data Preparation

Prior to analysis, several steps were taken to prepare the data. First, new dichotomous variables were created for items that were “Mark all that apply” so that 1=Answered and 9=Missing. Second, all variables were examined to determine which values are considered truly missing. Items that were truly missing (i.e., “Not answered”) were recoded to -1 to separate it from non-missing value (i.e., “Not applicable” and “Don’t know”). Finally, a dataset was created that indicates the number of respondents with valid responses and the number missing for each item and fiscal year.

## Analysis

Item nonresponse rates were analyzed by fiscal year and whether the item depended on the answer of a previous item. The numerator for each nonresponse rate consisted of the number of agricultural workers who did not answer the item. For items that did not depend on how a previous item was answered (i.e., no skip pattern; farmworker was required to answer all of these items), the denominator was the total sample size for that item. For items that depended on the answer to a previous item (i.e., skip pattern), the denominator was the number of agricultural workers eligible to respond to that item.

## Results

Table 4 shows the item nonresponse rate for items that do not have a skip pattern and the average nonresponse rate for each fiscal year. Overall, the average nonresponse rate in each fiscal year was less than 0.5 percent. The nonresponse rates for all items are low in all fiscal years, from 0 percent to 3.5 percent. Item 15 (“When was the last time your parents did hired farm work in the U.S.A?”) had the highest nonresponse rate, ranging from 1.1 in fiscal years 2014 and 2016 to 3.5 percent in fiscal year 2011.

Table 4. Item Nonresponse for Items Without Skip Patterns, by Fiscal Year.

|  |  |
| --- | --- |
|  | **Fiscal year** |
|  | **2011** | **2012** | **2013** | **2014** | **2015** | **2016** |
| **Item 1** | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% |
| **Item 2** | 0.1% | 0.3% | 0.1% | 0.2% | 0.3% | 0.2% |
| **Item 3** | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% |
| **Item 4** | 0.1% | 0.1% | 0.0% | 0.0% | 0.0% | 0.0% |
| **Item 5** | 0.1% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% |
| **Item 6** | 0.1% | 0.5% | 0.6% | 0.1% | 0.1% | 0.3% |
| **Item 7** | 0.0% | 0.1% | 0.1% | 0.0% | 0.0% | 0.0% |
| **Item 8** | 0.4% | 0.1% | 0.2% | 0.2% | 0.3% | 0.2% |
| **Item 9** | 0.9% | 0.7% | 1.5% | 0.4% | 2.1% | 1.4% |
| **Item 10** | 2.6% | 0.7% | 0.8% | 0.5% | 0.5% | 0.6% |
| **Item 11** | 3.2% | 0.4% | 0.6% | 0.6% | 0.5% | 0.4% |
| **Item 12** | 0.4% | 0.1% | 0.2% | 0.1% | 0.2% | 0.1% |
| **Item 13** | 0.6% | 0.1% | 0.1% | 0.3% | 0.4% | 0.6% |
| **Item 14** | 0.7% | 0.5% | 0.3% | 0.3% | 0.5% | 0.8% |
| **Item 15** | 2.8% | 3.5% | 2.1% | 1.1% | 1.4% | 1.1% |
| **Item 16** | 0.0% | 0.1% | 0.0% | 0.0% | 0.0% | 0.0% |
| **Item 17** | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% |
| **Item 18** | 0.4% | 0.8% | 0.4% | 0.0% | 0.0% | 0.2% |
| **Item 19** | 0.0% | 0.1% | 0.3% | 0.1% | 0.6% | 0.3% |
| **Item 20** | 0.1% | 0.1% | 0.2% | 0.1% | 0.4% | 0.2% |
| **Item 21** | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% |
| **Item 22** | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% |
| **Item 23** | 0.1% | 0.4% | 0.5% | 0.9% | 0.8% | 1.1% |
| **Item 24** | 0.1% | 0.2% | 0.1% | 0.2% | 0.1% | 0.2% |
| **Item 25** | 0.0% | 0.2% | 0.1% | 0.2% | 0.1% | 0.0% |
| **Item 26** | 0.0% | 0.1% | 0.0% | 0.1% | 0.1% | 0.2% |
| **Item 27** | 0.1% | 0.1% | 0.2% | 0.2% | 0.1% | 0.1% |
| **Item 28** | 0.1% | 0.1% | 0.1% | 0.2% | 0.2% | 0.2% |
| **Item 29** | 0.3% | 0.1% | 0.3% | 0.2% | 0.3% | 0.2% |
| **Item 30** | 0.2% | 0.2% | 0.4% | 0.5% | 0.3% | 0.1% |

Table 4. Item Nonresponse for Items without Skip Patterns, by Fiscal Year (Cont.)

|  |  |
| --- | --- |
|  | **Fiscal year** |
|  | **2011** | **2012** | **2013** | **2014** | **2015** | **2016** |
| **Item 31** | 0.1% | 0.1% | 0.4% | 0.1% | 0.0% | 0.1% |
| **Item 32** | 0.1% | 0.1% | 0.1% | 0.1% | 0.2% | 0.1% |
| **Item 33** | 0.0% | 0.1% | 0.1% | 0.2% | 0.2% | 0.1% |
| **Item 34** | 0.3% | 0.6% | 0.6% | 0.6% | 0.5% | 0.4% |
| **Item 35** | 0.3% | 0.3% | 0.1% | 0.2% | 0.9% | 0.4% |
| **Item 36** | 0.4% | 0.2% | 0.2% | 0.3% | 0.1% | 0.1% |
| **Item 37** | 0.1% | 0.1% | 0.1% | 0.3% | 0.2% | 0.2% |
| **Item 38** | 0.4% | 0.5% | 0.2% | 0.7% | 0.8% | 1.0% |
| **Item 39** | 1.1% | 0.4% | 0.7% | 0.8% | 0.8% | 1.4% |
| **Item 40** | 1.7% | 1.9% | 2.4% | 2.2% | 2.3% | 2.3% |
| **Item 41** | 0.7% | 1.3% | 0.6% | 0.5% | 0.5% | 0.4% |
| **Item 42** | 0.3% | 0.5% | 0.7% | 0.1% | 0.3% | 0.3% |
| **Item 43** | 0.2% | 0.4% | 0.6% | 0.3% | 0.4% | 0.3% |
| **Item 44** | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% |
| **Item 45** | 0.8% | 0.7% | 0.5% | 0.4% | 0.2% | 1.2% |
| **Item 46** | 1.2% | 0.8% | 0.8% | 0.6% | 0.2% | 1.2% |
| **Item 47** | 0.8% | 0.8% | 0.6% | 0.6% | 0.3% | 1.0% |
| **Item 48** | 0.1% | 0.1% | 0.2% | 0.0% | 0.1% | 0.7% |
| **Item 49** | 0.1% | 0.0% | 0.0% | 0.0% | 0.1% | 0.0% |
| **Item 50** | 0.5% | 0.5% | 0.4% | 0.2% | 0.2% | 0.1% |
| **Item 51** | 0.2% | 0.1% | 0.1% | 0.2% | 0.0% | 0.0% |
| **Item 52** | 0.3% | 0.3% | 0.2% | 0.2% | 0.5% | 0.4% |
| **Item 53** | 0.5% | 0.5% | 0.1% | 0.7% | 0.3% | 0.4% |
| **Item 54** | 0.6% | 1.6% | 0.8% | 0.5% | 0.6% | 0.6% |
| **Item 55** | 0.1% | 0.1% | 0.1% | 0.2% | 0.1% | 0.0% |
| **Item 56** | 0.1% | 0.1% | 0.1% | 0.2% | 0.1% | 0.0% |
| **Item 57** | 0.1% | 0.1% | 0.1% | 0.3% | 0.1% | 0.0% |
| **Item 58** | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% |
|  |  |  |  |  |  |  |
| **Average for Items 1–58** | 0.4% | 0.4% | 0.3% | 0.3% | 0.3% | 0.4% |

Table 5 shows the item nonresponse rate for items that have skip patterns and the average nonresponse rate for each fiscal year. Overall, the average nonresponse rate in each fiscal year was less than 2 percent. Fiscal years 2013 and 2014 had the lowest average nonresponse rates (1.1 percent), while fiscal year 2012 had the highest average nonresponse rate (1.7 percent). The nonresponse rate was low in all fiscal years, from 0 percent to 9.4 percent. Item 24 (“Does this employer keep in contact with you about future employment before leaving at the end of the season?”) had the overall highest nonresponse rate, ranging from 5.1 percent in fiscal year 2016 to 9.4 percent in fiscal year 2013. Items that also had higher nonresponse rates relative to other items with skip patterns are Item 4 (2.7–5.1 percent, “And in your home country, do you own or are you buying any of the following items?”); Item 22 (0–6.0 percent, “Are you paid as an individual or by the crew?”); Item 23 (2.2–8.3 percent, “How and when do you receive the money bonus?”); and Item 25 (3.3–7.1 percent, “Do you pay a fee to the grower/contractor ‘raiteros’ for rides to work?).

All items included in the item nonresponse study had nonresponse rates lower than 30 percent. So, none of the NAWS variables met the OMB criteria for further analysis of bias.

Table 5. Item Nonresponse for Items with Skip Patterns, by Fiscal Year (Cont.)

|  |  |
| --- | --- |
|  | **Fiscal year** |
|  | **2011** | **2012** | **2013** | **2014** | **2015** | **2016** |
| **Item 1** | 1.7% | 2.3% | 0.6% | 1.4% | 1.5% | 1.1% |
| **Item 2** | 0.2% | 0.3% | 0.3% | 0.1% | 0.0% | 0.1% |
| **Item 3** | 0.4% | 0.5% | 0.5% | 0.5% | 0.0% | 0.3% |
| **Item 4** | 2.7% | 4.8% | 2.9% | 3.9% | 5.1% | 4.9% |
| **Item 5** | 1.1% | 1.5% | 1.0% | 0.2% | 0.8% | 0.3% |
| **Item 6** | 0.6% | 1.6% | 0.3% | 0.7% | 0.9% | 0.9% |
| **Item 7** | 0.6% | 0.8% | 1.1% | 0.4% | 0.8% | 0.5% |
| **Item 8** | 0.3% | 0.3% | 0.7% | 0.2% | 0.1% | 0.3% |
| **Item 9** | 0.5% | 0.7% | 1.1% | 0.5% | 0.2% | 0.5% |
| **Item 10** | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.1% |
| **Item 11** | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.1% |
| **Item 12** | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.2% |
| **Item 13** | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.2% |
| **Item 14** | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.1% |
| **Item 15** | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.1% |
| **Item 16** | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.1% |
| **Item 17** | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.1% |
| **Item 18** | 0.1% | 0.0% | 0.0% | 0.3% | 0.0% | 0.2% |
| **Item 19** | 0.1% | 0.0% | 0.0% | 0.4% | 0.0% | 0.2% |
| **Item 20** | 1.1% | 1.7% | 1.0% | 1.6% | 1.7% | 1.4% |
| **Item 21** | 0.9% | 1.3% | 1.2% | 1.1% | 1.0% | 1.0% |
| **Item 22** | 3.9% | 5.7% | 0.0% | 3.2% | 2.9% | 6.0% |
| **Item 23** | 3.8% | 8.3% | 5.0% | 3.7% | 3.8% | 2.2% |
| **Item 24** | 7.7% | 7.1% | 9.4% | 8.0% | 7.6% | 5.1% |
| **Item 25** | 4.4% | 6.2% | 3.6% | 3.3% | 7.1% | 4.8% |
| **Item 26** | 0.8% | 1.2% | 0.2% | 0.3% | 0.3% | 0.6% |
| **Item 27** | 0.9% | 2.8% | 1.5% | 0.2% | 0.2% | 0.9% |
| **Item 28** | 0.6% | 0.2% | 0.7% | 0.2% | 0.3% | 0.5% |
|  |  |  |  |  |  |  |
| **Average for Items 1–28** | 1.2% | 1.7% | 1.1% | 1.1% | 1.2% | 1.2% |

## Conclusions

The purpose of this study was to determine the number of agricultural workers who did not answer each item (item nonresponse). The item nonresponse study indicated that, on average, there was less than 0.5 percent nonresponse rate for items without a skip pattern and less than 2 percent average nonresponse rate for items with a skip pattern. Since the average and individual item nonresponse rates are less than 30 percent, further analysis of bias is not necessary.

# Appendix B: Nonresponse Study 2 – NAWS Unit (Employer) Nonresponse

**NAWS Unit (Employer) Nonresponse**

## Introduction

The first analysis will assess NAWS nonresponse bias by comparing information in the sampling frame on eligible respondents and nonrespondents. This study was described in Part B of the OMB submission as follows: “While the sampling data is somewhat sparse for nonrespondents, three pieces of information are useful: geographic location, NAICS code, and the source used to obtain employer names. The NAWS will use three sources of employer names: a) the BLS UI list, b) marketing lists, and c) internet searches and contacts with knowledgeable local individuals. Geographic area and source lists are available for all employers, while NAICS codes are available for all employers who pay UI taxes, marketing list employers, and some additional employers.”

Using all three variables (source, NAICS, and geography), we made the following comparisons:

1. Employers allowing interviews compared to sampled employers that refused or were unable to be screened (i.e., excluding the ineligible),
2. Employers allowing interviews compared to eligible employers that refused, and
3. Eligible employers compared to unscreened sample members (employers whose eligibility could not be determined).

Nonresponse bias was calculated using the bias calculation formula from OMB’s Standard and Guidelines for Statistical Surveys (2006).

## Sample

The sample for the unit nonresponse study consisted of 26,151 agricultural employers from NAWS fiscal years 2013–2017 (cycles 74–88) that were contacted by JBS. The source lists of employers were obtained primarily from BLS and were supplemented with data from a commercial list (InfoUSA), as well as other sources (e.g., consultant and internet searches).

## Data Preparation

Prior to analysis, several steps were taken to prepare the data. Response codes were collapsed to allow for easier interpretation. All response codes, 1–46 and 97–99 for fiscal years 2013–2016 and 1–18 for fiscal year 2017 were recoded into response categories (1=Interviewed, 2=Refused, 3=Eligibility unknown, 4=Not in sample/Not eligible, 5=Cannot assign category or no need to contact). All sources were collapsed into 1=BLS, 2=InfoUSA, and 3=Other.

Data cleaning and recoding were conducted for incorrect or missing NAICS codes. One hundred thirty-four employers from the BLS source did not have the expected NAICS codes 1111 (Oilseed and Grain Farming), 1112 (Vegetable and Melon Farming), 1113 (Fruit and Tree Nut Farming), 1114 (Greenhouse, Nursery, and Floriculture Production), 1119 (Other Crop Farming), or 1151 (Support Activities for Crop Production). The NAICS codes for these employers were examined across quarters and years to find the correct NAICS codes. For example, an employer’s first quarter NAICS code might be 112, but other quarters are 111 or 1151. Of the 134 employers, 44 were recoded to NAICS 111 or 1151. Specifically, 1 employer was recoded to NAICS 1111, 7 were recoded to NAICS 1112, 9 were recoded to NAICS 1113, 8 were recoded to NAICS 1114, 7 were recoded to NAICS 1119, and 12 were recoded to NAICS 1151. The remaining 89 employers were unable to be recoded because all quarters and years show they have NAICS 1121–1129. One employer was recoded to missing due to inconsistent NAICS code information.

An additional 585 employers from the InfoUSA source did not have the expected NAICS codes 1111–1119 or 1151. All of these employers have secondary NAICS codes that are 1111–1119 or 1151. Specifically, 34 employers were recoded to NAICS 1111, 1 was recoded to NAICS 1112, 120 were recoded to NAICS 1113, 123 were recoded to NAICS 1114, 141 were recoded to NAICS 1119, and 166 were recoded to NAICS 1151.

Of the 1643 employers that do not have NAICS codes, 385 have SIC codes. The SIC codes for the 385 employers were converted to the appropriate NAICS codes using NAICS Identification Tools (https://www.naics.com/search/). Of the 385 employers, 31 employers were recoded to NAICS 1111, 53 were recoded to NAICS 1112, 141 were recoded to NAICS 1113, 60 were recoded to NAICS 1114, 79 were recoded to NAICS 1119, 20 were recoded to NAICS 1151, and 1 was recoded to NAICS 1122.

Due to the small number of employers (N = 102 across fiscal years 2013–2017) with NAICS 1121–1129, they were collapsed into NAICS 112.

## Analysis

Nonresponse rate, differences between respondents and nonrespondents, and nonresponse bias analyses were conducted to examine the differences between respondents and nonrespondents across different characteristics. The following characteristics were examined: source of the employer list (BLS, InfoUSA, or Other), NAICS (1111, 1112, 1113, 1114, 1119, 1151, and 112), and geography (divided into six and 12 regions). Using all three characteristics, the following comparisons were made:

1. Employers allowing interviews were compared to sampled employers that refused or were unable to be screened (i.e., excluding ineligible employers).
2. Employers allowing interviews compared to eligible employers that refused.
3. Eligible employers compared to employers whose eligibility could not be determined.

Nonresponse bias was calculated using the bias calculation formula from OMB’s *Standard and Guidelines for Statistical Surveys* (2006):

$$B\left(\overbar{Y}\_{r}\right)= \overbar{Y}\_{r}- \overbar{Y}\_{t}=\left(\frac{n\_{nr}}{n}\right)\left(\overbar{Y}\_{r}- \overbar{Y}\_{nr}\right)$$

where:

$\overbar{Y}\_{t}$ = the mean based on all sample cases;

$\overbar{Y}\_{r}$ = the mean based only on respondent cases;

$\overbar{Y}\_{nr}$ = the mean based only on nonrespondent cases;

$n$ = the number of cases in the sample; and

$n\_{nr}$ = the number of nonrespondent cases.

The formula provides a measure of nonresponse bias, which depends on the nonresponse rate and the difference between the means of respondents and nonrespondents on the three key variables. The smaller each of these components are, the smaller is the nonresponse bias.

In addition to nonresponse bias, logistics regressions were conducted to examine the effects of each characteristics while holding other characteristics constant.

## Results

Table 1 shows the distribution of the entire sample by source, NAICS, and geography. The majority (69 percent) of the employers were obtained from the BLS list. Almost one-third of the employers had NAICS 1119 (Other Crop Farming), followed by almost a quarter (22 percent) of employers with NAICS 1113 (Fruit and Tree Nut Farming). The majority (31 percent) of the employers were from California.

Table 2 shows the nonresponse rate and bias for the three comparisons, for source and NAICS. The nonresponse rate for source was 83–89, 55–57, and 61–75 percent for comparisons A, B, and C, respectively. There was low bias (2–4 percent) across the three comparisons and sources. There were more variations in nonresponse rates for NAICS; 70–95, 47–81, and 43–74 percent for comparison A, B, and C, respectively. Despite these larger variations and higher nonresponse, the bias remained low (0–10 percent). NAICS 1114 (Greenhouse, Nursery, and Floriculture Production) had the largest differences between respondents and nonrespondents (15 percent) and the largest bias of 10 percent for comparison A, but its nonresponse rate was the lowest (70 percent).

Table 3 shows the nonresponse rate and bias for the three comparisons, for geography (six regions and 12 regions). The nonresponse rate for the six regions was 80–87, 49–62, and 57–66 percent for comparison A, B, and C, respectively. The bias was low (1–7 percent) across the three comparisons. California has one of the highest nonresponse rates and one of the highest biases. The nonresponse rates for the 12 regions were 75–87, 46–63, and 54–66 percent for comparison A, B, and C, respectively. Similar to the 6-region analysis, the 12 regions also had low bias (0–7 percent) across the three comparisons, and California had one of the highest nonresponse rates and bias.

Table 1. Distribution of Source, NAICS, and Geography (Entire Sample).

|  |  |  |
| --- | --- | --- |
|  | **Sample size** | **Percent** |
| **Source** |  |  |
| BLS | 17981 | 69% |
| InfoUSA | 6897 | 26% |
| Other | 1271 | 5% |
| **NAICS** |  |  |
| 1111 (Oilseed and Grain Farming) | 2326 | 9% |
| 1112 (Vegetable and Melon Farming) | 1851 | 8% |
| 1113 (Fruit and Tree Nut Farming) | 5373 | 22% |
| 1114 (Greenhouse, Nursery, and Floriculture Production) | 3240 | 13% |
| 1119 (Other Crop Farming) | 7379 | 30% |
| 1151 (Support Activities for Crop Production) | 4235 | 17% |
| 112 (Cattle Ranching and Farming, Hog and Pig Farming, Poultry and Egg Production, Sheep and Goat Farming, Aquaculture, or Other Animal Production) | 102 | <1% |
| **Region 6** |  |  |
| East | 3829 | 15% |
| Southeast | 3255 | 12% |
| Midwest | 4664 | 18% |
| Southwest | 2978 | 11% |
| Northwest | 3365 | 13% |
| California | 8059 | 31% |
| **Region 12** |  |  |
| AP12 | 1723 | 7% |
| CA | 8059 | 31% |
| CBNP | 2977 | 11% |
| DLSE | 1691 | 6% |
| FL | 1564 | 6% |
| LK | 1687 | 6% |
| MN12 | 1302 | 5% |
| MN3 | 986 | 4% |
| NE1 | 968 | 4% |
| NE2 | 1138 | 4% |
| PC | 2063 | 8% |
| SP | 1992 | 8% |

AP12 = KY, NC, TN, VA, WV. CA = CA only. CBNP = IA, IL, IN, KS, MO, ND, NE, OH, SD. DLSE = AL, AR, GA, LA, MS, SC. FL = FL only. LK = MI, MN, WI. MN12 = CO, ID, MT, NV, UT, WY. MN3 = AZ, NM. NE1 = CT, MA, ME, NH, NY, RI, VT. NE2 = DE, DC, MD, NJ, PA. PC = OR, WA. SP = OK, TX.

East = AP12, NE1, NE2. Southeast = DLSE, FL. Midwest = CBNP, LK. Southwest = MN3, SP. Northwest = MN12, PC. California = California only.

Table 2. Unit Nonresponse Rate and Bias by Source and NAICS.

|  |  |  |  |
| --- | --- | --- | --- |
|  | **A. Nonresponse among all eligible and unscreened employers**  | **B. Nonresponse rate among eligible employers** | **C. Eligibility Rate**  |
| **Variable** | **Nonresponse rate** | **Difference between respondents and nonrespondents** | **Bias1** | **Nonresponse rate** | **Difference between respondents and nonrespondents** | **Bias** | **Nonresponse rate** | **Difference between respondents and nonrespondents** | **Bias1** |
| **Source** |  |  |  |  |  |  |  |  |  |
| BLS | 83% | 4% | 3% | 57% | -2% | -1% | 61% | 7% | 4% |
| InfoUSA | 86% | -3% | -2% | 55% | 1% | 1% | 69% | -5% | -3% |
| Other | 89% | -2% | -1% | 55% | 0% | 0% | 75% | -2% | -2% |
| **NAICS** |  |  |  |  |  |  |  |  |  |
| 111 or 1151 (vs 112) | 84% | 0% | 0% | 57% | 0% | 0% | 62% | 0% | 0% |
| 1111 | 89% | -4% | -3% | 67% | -3% | -2% | 68% | -2% | -1% |
| 1112 | 82% | 1% | 1% | 56% | 0% | 0% | 58% | 2% | 1% |
| 1113 | 84% | -1% | -1% | 56% | 1% | 1% | 64% | -2% | -1% |
| 1114 | 70% | 15% | 10% | 47% | 9% | 4% | 43% | 12% | 5% |
| 1119 | 86% | -5% | -5% | 60% | -3% | -2% | 66% | -4% | -3% |
| 1151 | 89% | -6% | -5% | 63% | -4% | -2% | 69% | -5% | -3% |
| 112 | 95% | 0% | 0% | 81% | 0% | 0% | 74% | 0% | 0% |

Comparison A = Employers allowing interviews compared to sampled employers that refused or unable to be screen (i.e., excluding the ineligible).

Comparison B = Employers allowing interviews compared to eligible employers who refused.

Comparison C = Eligible employers compared to employers whose eligibility could not be determined).

NAICS 1111 = Oilseed and Grain Farming. NAICS 1112 = Vegetable and Melon Farming. NAICS 1113 = Fruit and Tree Nut Farming. NAICS 1114 = Greenhouse, Nursery, and Floriculture Production. NAICS 1119 = Other Crop Farming. NAICS 1151 = Support Activities for Crop Production. NAICS 112 = Cattle Ranching and Farming, Hog and Pig Farming, Poultry and Egg Production, Sheep and Goat Farming, Aquaculture, or Other Animal Production.

**1**Bias = $\left(\frac{n\_{nr}}{n}\right)\left(\overbar{Y}\_{r}- \overbar{Y}\_{nr}\right)$

Table 3. Unit Nonresponse Rate and Bias by Geography.

|  |  |  |  |
| --- | --- | --- | --- |
|  | **A. Nonresponse among all eligible and unscreened employers** | **B. Nonresponse rate among eligible employers** | **C. Eligibility Rate**  |
| **Variable** | **Nonresponse rate** | **Difference between respondents and nonrespondents** | **Bias1** | **Nonresponse rate** | **Difference between respondents and nonrespondents** | **Bias** | **Nonresponse rate** | **Difference between respondents and nonrespondents** | **Bias1** |
| **Region 6** |  |  |  |  |  |  |  |  |  |
| East | 80% | 4% | 3% | 49% | 4% | 2% | 60% | 1% | 1% |
| Southeast | 80% | 3% | 3% | 49% | 4% | 2% | 61% | 1% | 1% |
| Midwest | 85% | -1% | -1% | 58% | -1% | 0% | 65% | -1% | -1% |
| Southwest | 86% | -2% | -2% | 60% | -2% | -1% | 66% | -1% | -1% |
| Northwest | 80% | 4% | 3% | 53% | 2% | 1% | 57% | 3% | 2% |
| California | 87% | -8% | -7% | 62% | -8% | -5% | 65% | -3% | -2% |
| **Region 12** |  |  |  |  |  |  |  |  |  |
| AP12 | 84% | 0% | 0% | 53% | 1% | 0% | 65% | -1% | 0% |
| CA | 87% | -8% | -7% | 62% | -8% | -5% | 65% | -3% | -2% |
| CBNP | 86% | -2% | -1% | 61% | -2% | -1% | 65% | -1% | -1% |
| DLSE | 81% | 2% | 1% | 46% | 3% | 1% | 64% | 0% | 0% |
| FL | 80% | 2% | 1% | 51% | 2% | 1% | 59% | 1% | 1% |
| LK | 83% | 0% | 0% | 52% | 1% | 1% | 64% | 0% | 0% |
| MN12 | 82% | 1% | 1% | 54% | 0% | 0% | 60% | 1% | 0% |
| MN3 | 84% | 0% | 0% | 55% | 0% | 0% | 65% | 0% | 0% |
| NE1 | 75% | 2% | 1% | 46% | 2% | 1% | 54% | 1% | 1% |
| NE2 | 77% | 2% | 1% | 46% | 2% | 1% | 57% | 1% | 1% |
| PC | 79% | 3% | 2% | 53% | 2% | 1% | 56% | 3% | 2% |
| SP | 87% | -2% | -2% | 63% | -2% | -1% | 66% | -1% | -1% |

Comparison A = Employers allowing interviews compared to sampled employers that refused or unable to be screen (i.e., excluding the ineligible).

Comparison B = Employers allowing interviews compared to eligible employers who refused.

Comparison C = Eligible employers compared to employers whose eligibility could not be determined.

AP12 = KY, NC, TN, VA, WV. CA = CA only. CBNP = IA, IL, IN, KS, MO, ND, NE, OH, SD. DLSE = AL, AR, GA, LA, MS, SC. FL = FL only. LK = MI, MN, WI. MN12 = CO, ID, MT, NV, UT, WY. MN3 = AZ, NM. NE1 = CT, MA, ME, NH, NY, RI, VT. NE2 = DE, DC, MD, NJ, PA. PC = OR, WA. SP = OK, TX. East = AP12, NE1, NE2. Southeast = DLSE, FL. Midwest = CBNP, LK. Southwest = MN3, SP. Northwest = MN12, PC. California = California only.

**1**Bias = $\left(\frac{n\_{nr}}{n}\right)\left(\overbar{Y}\_{r}- \overbar{Y}\_{nr}\right)$

Tables 4 and 5 show the regression results with the 6 regions and 12 regions, respectively. The estimate, standard error, statistical significance (*p*-values), and odds ratio are presented. The odds ratio shows the likelihood of employers participating in the NAWS compared to the reference category (BLS, NAICS 1119, and California) when holding all other variables constant.

In comparisons A and C, employers selected from the InfoUSA source are significantly less likely to participate compared to BLS and while holding all other variables constant (NAICS and region). There were no significant differences between the three sources in comparison B.

In all three comparisons, employers with NAICS 1114 (Greenhouse, Nursery, and Floriculture Production) had the highest likelihood of participating in NAWS, compared to NAICS 1119 (Other Crop Farming). For example, in comparison A, employers with NAICS 1114 were 2.5 times more likely to be interviewed than those with NAICS 1119 (Table 4).

In all three comparisons, employers in four of the six regions (East, Southeast, Midwest, and Northwest) were significantly more likely to participate in NAWS compared to California. Employers in in the Southwest region also had higher odds of participating than employers in California, but that was only significant in comparisons A. In terms of the 12 regions, employers in 10 of the regions had significantly higher odds of participating compared to California. Region SP was not statistically different compared to California. Florida in comparison C was also not significantly different compared to California.

Table 4. Regression with Source, NAICS, and Six Regions.

|  |  |  |  |
| --- | --- | --- | --- |
|  | **A. Nonresponse among all eligible and unscreened employers** | **B. Nonresponse rate among eligible employers** | **C. Eligibility Rate** |
|  | **B** | **Std** **Error** | **Sig** | **Odds radio** | **B** | **Std error** | **Sig** | **Odds radio** | **B** | **Std error** | **Sig** | **Odds radio** |
| **Source** |  |  |  |  |  |  |  |  |  |  |  |  |
| BLS1 | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- |
| InfoUSA | -0.32 | 0.07 | <.0001 | 0.73 | 0.03 | 0.09 | 0.7016 | 1.04 | -0.46 | 0.06 | <.0001 | 0.63 |
| Other | 0.37 | 0.67 | 0.5838 | 1.45 | 0.46 | 1.10 | 0.6752 | 1.59 | -0.04 | 0.68 | 0.9551 | 0.96 |
| **NAICS** |  |  |  |  |  |  |  |  |  |  |  |  |
| 11191 | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- |
| 1111 | -0.39 | 0.11 | 0.0003 | 0.68 | -0.26 | 0.13 | 0.0396 | 0.77 | -0.26 | 0.07 | 0.0005 | 0.77 |
| 1112 | 0.31 | 0.10 | 0.0011 | 1.37 | 0.22 | 0.11 | 0.0548 | 1.24 | 0.19 | 0.07 | 0.0087 | 1.21 |
| 1113 | 0.26 | 0.08 | 0.0007 | 1.30 | 0.38 | 0.09 | <.0001 | 1.47 | -0.01 | 0.06 | 0.8867 | 0.99 |
| 1114 | 0.93 | 0.07 | <.0001 | 2.53 | 0.55 | 0.09 | <.0001 | 1.73 | 0.82 | 0.06 | <.0001 | 2.26 |
| 1151 | -0.16 | 0.09 | 0.0652 | 0.85 | 0.05 | 0.10 | 0.666 | 1.05 | -0.22 | 0.06 | 0.0002 | 0.80 |
| 112 | -1.11 | 0.59 | 0.0588 | 0.33 | -1.00 | 0.63 | 0.1129 | 0.37 | -0.51 | 0.29 | 0.079 | 0.60 |
| **Region**  |  |  |  |  |  |  |  |  |  |  |  |  |
| California1 | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- |
| East | 0.64 | 0.08 | <.0001 | 1.90 | 0.62 | 0.10 | <.0001 | 1.86 | 0.31 | 0.07 | <.0001 | 1.37 |
| Southeast | 0.49 | 0.08 | <.0001 | 1.63 | 0.54 | 0.10 | <.0001 | 1.72 | 0.16 | 0.06 | 0.0116 | 1.17 |
| Midwest | 0.39 | 0.09 | <.0001 | 1.48 | 0.31 | 0.10 | 0.0023 | 1.37 | 0.22 | 0.07 | 0.001 | 1.25 |
| Southwest | 0.19 | 0.09 | 0.0306 | 1.21 | 0.19 | 0.10 | 0.0731 | 1.21 | 0.06 | 0.06 | 0.3115 | 1.07 |
| Northwest | 0.56 | 0.07 | <.0001 | 1.76 | 0.41 | 0.09 | <.0001 | 1.51 | 0.38 | 0.06 | <.0001 | 1.46 |

1 Reference category.

Table 5. Regression with Source, NAICS, and 12 Regions.

|  |  |  |  |
| --- | --- | --- | --- |
|  | **A. Nonresponse among all eligible and unscreened employers** | **B. Nonresponse rate among eligible employers** | **C. Eligibility Rate** |
|  | **B** | **Std** **Error** | **Sig** | **Odds radio** | **B** | **Std error** | **Sig** | **Odds radio** | **B** | **Std error** | **Sig** | **Odds radio** |
| **Source** |  |  |  |  |  |  |  |  |  |  |  |  |
| BLS1 | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- |
| InfoUSA | -0.38 | 0.08 | <.0001 | 0.68 | -0.02 | 0.09 | 0.7965 | 0.98 | -0.50 | 0.06 | <.0001 | 0.61 |
| Other | 0.29 | 0.72 | 0.6940 | 1.33 | 0.53 | 1.14 | 0.6419 | 1.70 | -0.07 | 0.71 | 0.9185 | 0.93 |
| **NAICS** |  |  |  |  |  |  |  |  |  |  |  |  |
| 11191 | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- |
| 1111 | -0.40 | 0.11 | 0.0002 | 0.67 | -0.24 | 0.13 | 0.0558 | 0.78 | -0.27 | 0.07 | 0.0003 | 0.76 |
| 1112 | 0.28 | 0.10 | 0.0042 | 1.32 | 0.20 | 0.12 | 0.0881 | 1.22 | 0.17 | 0.07 | 0.0243 | 1.18 |
| 1113 | 0.26 | 0.08 | 0.0012 | 1.30 | 0.39 | 0.09 | <.0001 | 1.48 | -0.02 | 0.06 | 0.7084 | 0.98 |
| 1114 | 0.96 | 0.08 | <.0001 | 2.61 | 0.59 | 0.09 | <.0001 | 1.80 | 0.83 | 0.06 | <.0001 | 2.29 |
| 1151 | -0.17 | 0.09 | 0.0483 | 0.84 | 0.05 | 0.11 | 0.6688 | 1.05 | -0.23 | 0.06 | 0.0001 | 0.79 |
| 112 | -1.14 | 0.59 | 0.052 | 0.32 | -1.06 | 0.64 | 0.0984 | 0.35 | -0.53 | 0.29 | 0.0656 | 0.59 |
| **Region**  |  |  |  |  |  |  |  |  |  |  |  |  |
| CA1 | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- |
| AP12 | 0.45 | 0.11 | <.0001 | 1.56 | 0.48 | 0.14 | 0.0004 | 1.61 | 0.15 | 0.09 | 0.0819 | 1.16 |
| CBNP | 0.35 | 0.10 | 0.0008 | 1.41 | 0.23 | 0.12 | 0.0517 | 1.26 | 0.23 | 0.08 | 0.0031 | 1.26 |
| DLSE | 0.75 | 0.10 | <.0001 | 2.11 | 0.79 | 0.13 | <.0001 | 2.21 | 0.27 | 0.08 | 0.0011 | 1.31 |
| FL | 0.28 | 0.10 | 0.0041 | 1.33 | 0.33 | 0.12 | 0.0062 | 1.39 | 0.06 | 0.08 | 0.4795 | 1.06 |
| LK | 0.52 | 0.12 | <.0001 | 1.68 | 0.49 | 0.15 | 0.0008 | 1.63 | 0.24 | 0.09 | 0.01 | 1.27 |
| MN12 | 0.61 | 0.12 | <.0001 | 1.84 | 0.48 | 0.14 | 0.0008 | 1.61 | 0.36 | 0.09 | <.0001 | 1.43 |
| MN3 | 0.48 | 0.13 | 0.0002 | 1.61 | 0.45 | 0.16 | 0.0044 | 1.56 | 0.20 | 0.10 | 0.0383 | 1.22 |
| NE1 | 1.07 | 0.14 | <.0001 | 2.92 | 0.91 | 0.18 | <.0001 | 2.48 | 0.68 | 0.12 | <.0001 | 1.98 |
| NE2 | 0.64 | 0.12 | <.0001 | 1.90 | 0.64 | 0.15 | <.0001 | 1.90 | 0.30 | 0.10 | 0.0027 | 1.35 |
| PC | 0.55 | 0.08 | <.0001 | 1.74 | 0.39 | 0.10 | <.0001 | 1.48 | 0.39 | 0.06 | <.0001 | 1.48 |
| SP | 0.05 | 0.11 | 0.6372 | 1.05 | 0.06 | 0.12 | 0.6316 | 1.06 | 0.00 | 0.08 | 0.9542 | 1.00 |

1 Reference category.

## Conclusions

The purpose of the unit nonresponse study was to examine nonresponse among agricultural employers sampled by the NAWS to determine whether there was any systematic bias between employers granting permission for their workers to be interviewed compared with nonrespondents, those refusing interviews, those unable to be screened, or those for which survey staff could not determine eligibility. Potential bias was examined across source, NAICS code, and geographic region.

The results of the unit nonresponse study indicated that nonresponse rates were between 43 and 95 percent, depending on whether respondents were compared to those refusing and unable to be screened (70–95 percent), only those unable to be screened (46–81 percent), or those for which eligibility could not be determined (43–75 percent). There were small variations in nonresponse rates between BLS, InfoUSA, and other sources, and between the 6 or 12 regions. There were larger variations in nonresponse rates between the NAICS codes. Although nonresponse rates are high, the bias was less than 10 percent, which indicates that there were small differences between respondents and nonrespondents across the three sources, seven NAICS codes, and 6 or 12 geographic locations.

# Appendix C: Nonresponse Study 3 – Follow Up with Employers Who Were Not Successfully Screened During the Initial NAWS Data Collection

Follow Up with Employers Who Were Not Successfully Screened During the Initial NAWS Data Collection

## Introduction

* As part of efforts to understand agricultural employer Nonresponse in the National Agricultural Workers Survey, JBS conducted a nonresponse follow-up (NRFU) study. The additional contact was focused on identifying whether nonresponding employers were eligible or ineligible. The goal was to see if further efforts could improve the NAWS response rate.
* JBS used both mail and telephone attempts to contact nonresponding agricultural employers whose survey eligibility was unable to be determined by interviewer contacts during the Fall 2017 interview cycle (October 2017–February 2018). The additional contact efforts were carried out from May 2018 through August 2018. The gap between the end of the cycle and the follow-up time period was deliberate. Contacting employers at another time in the agricultural cycle was done to see if employers that might have been too busy to respond during the fall would respond if contacted during another season.

## NRFU screening

The additional NRFU focused on screening nonresponding employers for eligibility using the questions similar to those that interviewers would ask employers to determine eligibility when carrying out the survey. Employers were asked whether there had been employees actively working during the time period when NAWS interviewers were on site (the reference period) and whether these workers had been doing qualifying tasks on qualifying crops. Employers were also asked how many qualifying workers they had during the reference period.

Interviewers obtain this screening information during their initial contacts with the employers and ask these questions as part of a conversation, probing when needed for additional information. For the NRFU study, the eligibility contacts were standardized and distilled to a set of four questions that were included in a) a telephone script for contacting employers by phone, and b) a letter from the survey director that asked the employers to return their answers to the questions by mail. Both the script and the letter included the same explanation of the survey and the reasons for contacting the employers. The letter also included a JBS contact name and telephone number that the employer could call with questions or concerns about responding. The letter text can be found at the end of this appendix.

The Fall 2017 contact attempts with a nonresponding employer generally happened during a single interviewer trip. In a few counties with large interview allocations, interviewers made more than one trip and continued to contact nonresponding employers from earlier trips. To standardize the reference period, these employers were asked for information about their operations during the first interviewer trip to their county.

## Sampling Universe

The universe for the Nonresponse Follow-up (NRFU) Study included 779 growers in 67 counties across 23 states. The list included each employer’s name and contact information along with documentation of the contact attempts made by the NAWS interviewer.

* The criteria used to select nonresponding agricultural employers for the NRFU study were 1) the employers had been randomly selected for inclusion in the NAWS employer sampling list for Fall 2017, and 2) response codes indicated that the agricultural employers eligibility had not been determined because: a) the employer had not responded to the NAWS interviewers’ outreach during Fall 2017; b) the employer outreach was incomplete and eligibility for survey inclusion had not been determined; or c) the response code for that employer was missing in the interviewer documentation.

## Nonresponse follow-up attempts

The nonresponse follow-up was carried out in four waves that included different combinations of mail and telephone contact attempts. Table 1 shows each wave and the number of employers who responded, the number of employers who did not respond at each wave, and the response rate.

The initial mailing consisted of 779 agricultural employers. A second mailing was sent to 448 NRFU sample members who had not responded to the first letter, who were not in the wave receiving only one mailing plus phone calls, and whose first letters had not been returned as undeliverable. Of the 779 agricultural employers, a random subset of 268 employers also received phone contact attempts. One group of 127 received only the first mailing and one or more telephone calls, while the remaining 141 received both mailings and one or more telephone calls.

Data collection began April 1, 2018 with the first mailing. The first set of NFRU calls took place from May 22 to June 14. The second mailing was sent out beginning June 14 and the second set of telephone calls were conducted from July 16 to August 17, 2018.

Table 1. Results of each wave of nonresponse follow up.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Mailing** | **N** | **Response** | **Non Response** | **Response Rate** | **Cumulative response** | **Cumulative response rate** |
| **First mailing\*** | 779 | 83 | 696 | 11% | 83 | 11% |
| **Second mailing** | 448 | 53 | 395 | 12% | 136 | 17% |
|  **Total for both rounds of**  **mail response** | **779** | **136** |  | **17%** |   |   |
| **Phone follow up after mailing(s)** |   |   |   |   |   |   |
| **One mailing and phone follow up** | 127 | 51 | 76 | 40% |  51  | 40% |
| **Two mailings and phone follow up** | 141 | 44 | 97 | 31% |  95  | 35% |
|  **Total phone response** | **268** | **95** |  | **35%** |   |   |
| **Mail and Phone Response** **Combined** | **779** | **231** | **548** | **30%** |   |   |

*\*Nonresponse to the first mailing included 121 letters that were returned as undeliverable.*

A total of 231 follow-up screenings were completed. One hundred and thirty-six employers responded to the mailings and returned their screening information by mail. Another 95 screenings were completed by telephone. The response rate was 30 percent at the completion of data collection ($\frac{231}{779}=0.30$).

## Data Preparation and the Analytic Sample

Prior to analysis, steps were taken to prepare the NRFU data and the Fall 2017 data for analysis. The first step was to review the responses to the screening questions and determine if, at the end of the NRFU data collection, the contacted employers could be classified as eligible, ineligible and/or eligibility unknown. This was done by examining the responses to the four eligibility questions. For example, responses to the question about whether the employer had active workers during the reference period included answers such as “Yes” or “Yes, they’re year-round.” There were coded as eligible provided the crops and tasks were qualifying. Responses such as “No, don’t hire workers” or “No…do not farm or hire farmworkers” were coded as ineligible.

Seven agricultural employers were found to have been ineligible for the NRFU study. Further cleaning of the response data at the end of FY2018 resulted in one agricultural employer being removed from the sample because the employer had been interviewed in Fall 2017. Six additional agricultural employers were removed from analysis because they were not contacted during Fall 2017 and should not have been included in the survey. The final sample size for analysis was 772 agricultural employers receiving additional NRFU; 229 of them were successfully screened. The final samples consisted of 1,721 agricultural employers in Fall 2017 and 772 agricultural employers in the NRFU study.

Finally, the 2018 NRFU data was merged with the full Fall 2017 NAWS employer sample to calculate employer response rates. A new set of response codes was created that updated the fall 2017 response codes using the NRFU data. For example, if an agricultural employer’s eligibility was unknown in Fall 2017 but found to be eligible in the NRFU data, the final response code was eligible.

## Analysis

Table 2 shows the response codes of the NRFU sample after coding the screened NRFU employers. An important issue for the analysis was that the employers in the NRFU sample did not have the opportunity to agree or refuse to participate in the NAWS survey because the survey period had passed. Employers contacted in Fall 2017 had response codes that included whether eligible employers had agreed or refused to participate in the NAWS.

To address this issue, the response rate was calculated for four scenarios:

1. The initial Fall 2017 cycle without the NRFU data.
2. All eligible agricultural employers screened during the NRFU that refused to allow interviews.
3. A proportion of eligible agricultural employers screened during the NRFU that allowed interviews
4. All eligible agricultural employers screened during the NRFU that allowed interviews.

Based on Table 2, the number for Scenario 3 is based on the number of interviews (N = 142) and refusals (N = 191) in Fall 2017, and number of eligible agricultural employers in the NRFU sample allowing interviews (N = 123). The estimated number of interviews in the NRFU sample was 52 [142/(142+191)\*123 = 52.45].

Table 2. Result of Nonresponse Follow Up with a Sample

of Employers of Unknown Eligibility.

|  |
| --- |
| **Eligibility after the nonresponse follow up** |
| **Response Code** | **Employers** |
| **Eligible** | 123 |
| **Ineligible** | 106 |
| **Eligibility Unknown** | 543 |
| **Total** | 772 |

Response rates were calculated using the formula for the unweighted response rate from the Office of Management and Budgets’ *Standards and Guidelines for Statistical Surveys*.[[1]](#footnote-1)

$$RRU= \frac{C}{C+R+NC+O+e(U) }$$

Where:
*C* = number of completed cases or sufficient partials;
*R* = number of refused cases;
*NC* = number of noncontacted sample units known to be eligible;
*O* = number of eligible sample units not responding for reasons other than refusal;
*U* = number of sample units of unknown eligibility, not completed; and
*e* = estimated proportion of sample units of unknown eligibility that are eligible.

## Results

Table 3 shows the response rates for the four scenarios. Of the 1,721 agricultural employers in the Fall 2017 cycle, 142 allowed interviews, resulting in a response rate of 25 percent. If all 123 eligible agricultural employers in the NRFU refused to allow interviews, the resulting response rate would be 20 percent (a 5% decrease compared to Fall 2017). On the other hand, if all 123 eligible NRFU respondents agreed to allow interviews, the response rate would be 37 percent. If the proportion of the NRFU respondents allowing interviews was the same as the Fall 2017 eligible employers, there would be an additions 52 employers allowing interviews and the response rate would be 27 percent.

Table 3. Number of Agricultural Employers from Fall 2017 NAWS Sample.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Category** | **Fall 2017****All NRFU Coded as Eligibility Unknown** | **All Eligible NRFU Refusing Interviews** | **NRFU Share of** **Refusals Same** **as Fall 2017** | **All Eligible NRFU** **Allowing Interviews** |
| **Interviewed** | 142 | 142 | 194 | 265 |
| **Refused** | 191 | 314 | 262 | 191 |
| **Ineligible** | 570 | 676 | 676 | 676 |
| **Eligibility Unknown** | 818 | 589 | 589 | 589 |
| **Total** | 1721 | 1721 | 1721 | 1721 |
|  |   |   |   |   |
| **Response Rate** | 25% | 20% | 27% | 37% |

## Conclusions

The purpose of this follow-up study was to determine whether additional effort to contact nonresponding agricultural employers would improve the NAWS response rate. NAWS interviewers contacted 779 nonresponding employers via mail and follow-up telephone calls. At the end of the NRFU data collection period, 231 agricultural employers responded. Employers responded to both the mail and telephone contact attempts. With a mail-only response rate of 17 percent over both mailings and a 35 percent response in the sample who received mailings combined with telephone follow up, it increased the overall NRFU response rate to 30 percent.

Response rates were calculated for four possible scenarios: 1) the initial Fall 2017 cycle without additional effort to contact nonresponding agricultural employers, 2) all eligible agricultural employers who responded to the NRFU refused to allow interviews, 3) a proportion of eligible agricultural employers who responded to the NRFU allowed interviews, and 4) a similar proportion of eligible agricultural employers who responded to the NRFU allowed interviews compared to the Fall 2017 eligible respondents.

The results show that a likely lower bound on the response rate is 20 percent when all NRFU eligible refuse, and conversely the upper bound is 37 percent when all are assumed to allow interviews. A more probable response rate is that the proportion of eligible NRFU employers allowing interviews would be similar to that of eligible employers in Fall 2017 for a response rate of 27 percent.

Comparing the probable NRFU adjusted response rate of 27 percent to the actual Fall 2017 response rate of 25 percent shows that the NRFU had only a small impact on the NAWS response rate. This result is the same as was found in a similar study done in 2009. That is, further efforts of contacting nonresponding employers does not substantially affect the response rates.

The persistence of nonresponding employers is likely an artifact of the NAWS employer list construction. The main component of the NAWS employer sampling frame is the BLS list of employers participating in the Federal unemployment insurance (UI) system. In most states, only large employers participate in the UI system. To overcome this bias, JBS enriches the sampling frame with administrative and commercial lists of employers as well as through internet searches. The quality of these lists is not as high as the UI list, resulting in large numbers of potentially eligible employers that are unable to be contacted.

**Sample of the Nonresponse Mail-out Survey**

Date

«NawsId» / «ListOrder»

«TradeName»

«Address»

«City», «State» «ZipCode»

To Whom It May Concern:

JBS International is conducting a private follow-up to verify the accuracy of our field operations. A field representative from the National Agricultural Workers Survey, attempted to contact you last fall about our survey.

The main objective of the survey sponsored by the U.S. Department of Labor is to identify trends in the make-up of the hired farm workforce. The information obtained helps agricultural employers and grower organizations stay informed about the characteristics of the hired farm workforce and helps public and private agencies better plan programs for farm workers.

JBS International, Inc. is a private research firm that provides professional, technical, and management services for policy analysis and program evaluation to government agencies, education agencies, and the private sector. JBS International, Inc. has no connection to any union organization.

One of our representatives was in your area last fall/spring and unsuccessfully attempted to contact you. We are trying to assess why our representative was unsuccessful. This helps us improve our records, monitor our field representatives and provide more accurate survey results. To do this, we are asking you to answer the following questions:

Did you have employees working on crops, plants, vines, or trees (and their fruits or seeds) during **October 2017 thru February 2018?** Please select one: **YES** or **NO**

1. If **YES:** Were they performing activities related to growing, harvesting or on-farm processing of your raw product during the week of: **«Reference Period»** Please select one: **YES** or **NO**
2. What types of crops, plants, vines, or trees (and their fruits or seeds) were they primarily working on:

1. Approximately **how many workers** did you have in these activities during this time period?

Your cooperation would be greatly appreciated. Please return your response either by mail in the postage-paid envelope, by email to Susan Gabbard at sgabbard@jbsinternational.com, or by fax to (650) 348-0260.

If you have any questions, I would be happy to speak with you. I can be reached between 8:00 a.m. and 5:00 p.m. Pacific Time at the following toll-free number (866) YES-NAWS or (866) 937-6297.

Sincerely,



Susan M. Gabbard, Ph.D.
Project Director

# Appendix D: Nonresponse Study 4 – Examining Employer Eligibility Over Time and NAWS Response Rates

**Examining Employer Eligibility Over Time and NAWS Response Rates**

A small number of agricultural employers appear on the survey’s sampling list in multiple administrations of the survey. Attempts to contact these employers may have had different outcomes at different time periods. This study used Markov chain analysis to incorporate information from prior data periods about employers’ states – whether eligible, ineligible, or unable to be determined – and looked at the impact on response rates. The information used for this analysis contains the responses from 34,774 growers collected during cycles 53 through 88 (FY 2006–2017). Each contact was coded with a number from 1 to 8. The following table shows the distribution of codes at the time of first contact.

**Initial state**

| **Cat** | **Description** | **Count** | **Percent** |
| --- | --- | --- | --- |
| **1** | Yes | 3,442 | 9.9 |
| **2** | Yes but | 850 | 2.4 |
| **3** | Qualified refusal | 2,991 | 8.6 |
| **4** | Don’t know | 10,611 | 30.5 |
| **5** | Incomplete | 2,763 | 7.9 |
| **6** | Not in sample | 11,396 | 32.8 |
| **7** | Skipped | 481 | 1.4 |
| **8** | Office codes | 2,240 | 6.4 |
| **Total** |  | 34,774 | 100.0 |

The following table shows all 64 possible transition probabilities.

**Transition probabilities**

|  |  | **From this state** |
| --- | --- | --- |
|  |  | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** |
| **To this State** | **1** | 0.420 | 0.174 | 0.119 | 0.074 | 0.125 | 0.073 | 0.108 | 0.142 |
| **2** | 0.062 | 0.034 | 0.027 | 0.019 | 0.026 | 0.014 | 0.022 | 0.030 |
| **3** | 0.084 | 0.134 | 0.233 | 0.082 | 0.149 | 0.074 | 0.103 | 0.070 |
| **4** | 0.119 | 0.242 | 0.204 | 0.412 | 0.228 | 0.269 | 0.341 | 0.320 |
| **5** | 0.083 | 0.126 | 0.138 | 0.090 | 0.189 | 0.070 | 0.103 | 0.062 |
| **6** | 0.160 | 0.224 | 0.196 | 0.239 | 0.221 | 0.419 | 0.220 | 0.218 |
| **7** | 0.008 | 0.008 | 0.017 | 0.016 | 0.015 | 0.009 | 0.036 | 0.011 |
| **8** | 0.065 | 0.058 | 0.067 | 0.069 | 0.046 | 0.072 | 0.067 | 0.148 |
|  | **Tot** | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |

The (left stochastic) transition probability matrix can be used to predict the distribution of codes after a certain number of contacts. If the initial distribution of contacts is contained in the column vector **x0** and the transition probability matrix is called **P**, then the predicted distribution after *n* steps is **xn** = **Pnx0**.

For example, we would predict the distribution of codes at the time of second contact (i.e., after one step) to be:



Similarly, the predicted distribution of codes at the third contact (second step) would be:



By taking the limit of **P***n* as *n* approaches infinity, the transition matrix converges to:

|  |  |  |
| --- | --- | --- |
|  |  | **From this state** |
|  |  | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** |
| **To this State** | **1** | 0.140 | 0.139 | 0.140 | 0.140 | 0.139 | 0.139 | 0.139 | 0.140 |
| **2** | 0.026 | 0.026 | 0.026 | 0.026 | 0.026 | 0.026 | 0.026 | 0.026 |
| **3** | 0.103 | 0.103 | 0.103 | 0.103 | 0.103 | 0.103 | 0.103 | 0.103 |
| **4** | 0.282 | 0.281 | 0.282 | 0.282 | 0.281 | 0.281 | 0.281 | 0.282 |
| **5** | 0.097 | 0.097 | 0.097 | 0.097 | 0.097 | 0.097 | 0.097 | 0.097 |
| **6** | 0.268 | 0.268 | 0.268 | 0.268 | 0.267 | 0.268 | 0.268 | 0.268 |
| **7** | 0.013 | 0.013 | 0.013 | 0.013 | 0.013 | 0.013 | 0.013 | 0.013 |
| **8** | 0.072 | 0.072 | 0.072 | 0.072 | 0.072 | 0.072 | 0.072 | 0.072 |
|  | **Tot** | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |

Notice that regardless of the present state, 14% of the growers are predicted to move into code 1 at the next contact, 2.6% into code 2, 10.3% into code 3, and so forth. Thus, after many steps, the predicted distribution would be:

|  |  |  |  |
| --- | --- | --- | --- |
| **Cat** | **Description** | **Count** | **Percent** |
| **1** | Yes | 4,868.4 | 14 |
| **2** | Yes but | 904.1 | 2.6 |
| **3** | Qualified refusal | 3,581.7 | 10.3 |
| **4** | Don’t know | 9,806.3 | 28.2 |
| **5** | Incomplete | 3,373.1 | 9.7 |
| **6** | Not in sample | 9,319.4 | 26.8 |
| **7** | Skipped | 452.1 | 1.3 |
| **8** | Office codes | 2,503.7 | 7.2 |
| **Total** |  | 34,774 | 100.0 |

Finally, the initial distribution can be compared to the convergent distribution to see which categories are likely to gain or lose entries over time.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Cat** | **Description** | **InitialPercent** | **FinalPercent** | **Change** |
| **1** | Yes | 9.9 | 14 | 4.1 |
| **2** | Yes but | 2.4 | 2.6 | 0.2 |
| **3** | Qualified refusal | 8.6 | 10.3 | 1.7 |
| **4** | Don’t know | 30.5 | 28.2 | –2.3 |
| **5** | Incomplete | 7.9 | 9.7 | 1.8 |
| **6** | Not in sample | 32.8 | 26.8 | –6 |
| **7** | Skipped | 1.4 | 1.3 | –0.1 |
| **8** | Office codes | 6.4 | 7.2 | 0.8 |
| **Total** |  | 100.0 | 100.0 | 0.0 |

**CASRO Response Rate**

The calculations in this section are based upon the formulas in the following two papers.

The American Association for Public Opinion Research. *Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys*. Lenexa, Kansas: The American Association for Public Opinion Research, 2000.

Lynn P, Beerten R, Laiho J, Martin J. *Recommended Standard Final Outcome Categories and Standard Definitions of Response Rate for Social Surveys*. Colchester, Essex: The Institute for Social and Economic Research, 2001.

In addition, the results were verified using the online calculator at:

<http://www.quantitativeskills.com/sisa/calculations/casro.htm>.

The following correspondence was used between the CASRO and NAWS categories.

|  |  |  |
| --- | --- | --- |
| **CASRO letter** | **CASRO description** | **NAWS Categories** |
| A | Completed | 1 |
| B | Refused, eligible | 3 |
| C | Unknown | 2, 4, 5, 7, 8 |
| D | Ineligible | 6 |

There are various response rates defined in the papers cited above. The “overall response rate” from the Lynn paper, which agrees with RR4 in the AAPOR paper, is:

, where .

Using the initial state counts, these values would be:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Cat** | **Description** | **Count** | **Total** | **CASRO letter** |
| **1** | Yes | 3,442 | 3,442 | A |
| **3** | Qualified refusal | 850 | 850 | B |
| **2** | Yes but | 2,991 | 28,242 | C |
| **4** | Don’t know | 10,611 |
| **5** | Incomplete | 2,763 |
| **7** | Skipped | 11,396 |
| **8** | Office codes | 481 |
| **6** | Not in sample | 2,240 | 2,240 | D |
| **Total** |  | 34,774 | 34,774 |  |

*e* = 0.6571 and *RR* = 0.1506.

Using the steady state counts, the values would be:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Cat** | **Description** | **Count** | **Total** | **CASRO letter** |
| **1** | Yes | 4,868.4 | 4,868.4 | A |
| **3** | Qualified refusal | 904.1 | 904.1 | B |
| **2** | Yes but | 3,581.7 | 26,532.6 | C |
| **4** | Don’t know | 9,806.3 |
| **5** | Incomplete | 3,373.1 |
| **7** | Skipped | 9,319.4 |
| **8** | Office codes | 452.1 |
| **6** | Not in sample | 2,503.7 | 2,503.7 | D |
| **Total** |  | 34,774 | 34,774 |  |

*e* = 0.6975 and *RR* = 0.2005.

The results of the analysis showed a five-percentage point gain in the CASRO response rate from 15 to 20 percent. The overall expected response rate after a large number of cycles is 20 percent.

# Appendix E: Design Study A – Efficiency of the NAWS Sampling Design

Efficiency of the NAWS

Sampling Design

The purpose of this study was to understand the variation in crop worker responses from the National Agricultural Workers Survey (NAWS) that resulted from the survey’s complex sampling design elements, namely stratification and clustering. The study used a mixed and fixed effects ANOVA model to examine the mean squared error from each level of the NAWS’s complex sample for a significant relationship with each of the eight key variables. The research study examined two research questions: a) How optimal was the NAWS sampling design for estimating each of the key variables; and b) Were there trends in variation among clusters and strata across these variables that could inform improvements in future iterations of the survey?

Under an optimal design, the sampling frame is divided into strata such that there is maximum homogeneity among respondents within strata and heterogeneity between them. The opposite is desirable for clusters, where maximum heterogeneity is desired within clusters and homogeneity between them. In practice, conditions may be far from ideal; especially in the case of clusters, which are formed for reasons of sampling convenience or cost.

NAWS data serve a variety of stakeholders, each with different key indicators. While survey designs can be optimized for a particular key indicator or point estimate, designs generally cannot be optimized across multiple indicators. This study’s analysis examines how the survey design affects estimates of several key policy and program indicators collected by the survey.

## Method

The National Agricultural Workers Survey (NAWS) employs a multi-stage sampling design to collect information from crop workers in the continental United States. The sampling frame is divided into seasonal and regional strata to account for differences in agricultural workers over both time and space, and sampling units are chosen in clusters for cost reasons. Three cycles of data collection are conducted per year and a sample of farm workers is selected within each of twelve geographic regions. The sample of farm workers is selected by a nested sampling procedure where workers are nested within employers within Zip Code regions within counties within Farm Labor Areas (FLAs), which are the primary sampling units. The FLAs consist of single counties or groupings of counties with similar labor patterns. The employers are selected using a simple random sample of agricultural employers selected from a list of employers. Once the sample of employers is drawn, interviewers contact the selected growers or contractors to obtain access to the work site.

To test the relationship between the levels of the sampling design and the key variables, the NAWS data were analyzed using a nested mixed effects model. Stratification variables were considered fixed effects and included the fiscal year (FY), cycle (Cyc), and agricultural region (R). Clusters, which are selected from a larger population at each sampling level, are random effects. Clusters include FLAs, counties (Cou), Zip Code regions (Zip), and farm employers (F).

The NAWS data used in this study were limited to FY 2011 to 2014 data, as earlier years employed different sampling designs. Particularly, the Zip Code region was first included as a sampling variable in Federal Fiscal Year 2010. To examine the stability of the effects, the model was run on 2011–2012, 2013–2014, and 2011–2014 data. The 2011–2014 data was run with the existing weights, and with a proposed weight that included a different employer nonresponse calculation. For each set of data years and weights, the model was run eight times, once for each of the variables listed in Table 1 below.

Table 1. Response variables and their definitions.

|  |  |
| --- | --- |
| Response Variables | Definition |
| Wage | Hourly wage or hourly wage equivalent for piece rate workers. |
| FLC employer | Employed by a farm labor contractor (FLC).  |
| Indigenous | Indigenous Central or South American.  |
| Unauthorized | Lacking U.S. work authorization. |
| Number of farm employers | Number of farm employers in the past 12 months. |
| How the agricultural worker was paid  | Paid by the piece versus hourly and other forms of payment.  |
| Crop workdays  | Number of crop workdays in the past 12 months. |
| Number of children in the household | Number of children in economic household. Includes children not co-resident with respondent if supported by respondent. |

To test the significance of each level of sampling on the mean squared errors, F tests were used. The F tests account for nested effects and the mixture of random and fixed effects. When exact F tests are not available, Satterthwaite’s approximate F tests are constructed by using ratios of linear combinations of mean squares from the ANOVA table. The preferred method is to construct these linear combinations using only positive coefficients, but sometimes it is necessary to use some negative coefficients as well, as discussed in *Design and Analysis of Experiments* (Montgomery, D., 2013, p. 595).[[2]](#footnote-2) When this occurs, there is a possibility that the computed F statistic will be negative. Since the true F distribution only admits nonnegative values, a negative F statistic is interpreted as having the value 0, with a corresponding p-value of 1.

## Results

For all four analyses, Table 2 shows the sample elements that were significant. Tables 3 to 5 show the results analysis using FY 2011–2012 data, 2013–2014 data, and combined 2011–2014 data, respectively. Table 6 shows the results using the alternate weight calculations, which include a more complex grower nonresponse calculation, for fiscal years 2011–2014.

Two main patterns were found when we examined the first two sets of analyses using FY 2011–2012 data and 2013–2014 data. First, region (or the cycle/region interaction) was consistently found to be a significant effect in all except two variables. This suggests that stratification by geographic location divides farm workers into heterogeneous groups and is, therefore, an effective design strategy.

Clustering at the county level and employer level were also consistently significant effects, indicating crop workers within one cluster of employers (or county) were significantly different than farm workers in another cluster of growers (or county) for a particular combination of higher-level clustering and stratification.

No patterns emerged for other design elements, nor for the weight variables. Outside of region, county, and grower, there was little consistency in significant effects across variables. In terms of the NAWS weight and alternate weight, there were some significant effects using current NAWS weight that were non-significant when using the alternate weight, and vice versa.

## Conclusions

The analysis is consistent with what we know about U.S. agriculture from USDA data and from the ethnographic record, as well as NAWS experience. Agricultural tasks and labor force vary by region and season. For example, hot southern regions have a slack season in summer when Northern regions are peaking. While some crops are grown in most regions, some crops and their tasks are concentrated in one or two regions. The significance of the stratification variables is consistent with the variations that occur in U.S. agriculture by region and cycle. The stratification variables helped optimize the design.

While not statistically desired, the significance of the employer and local geographic variables cluster variables correctly reflects the heterogeneity of agricultural tasks and workers within a region. The Census of Agriculture, ethnographic studies of the farm labor force, and NAWS experience agree that employers that grow different crops with different labor demands have different types of agricultural workers. For example, within a county, employers with H-2A versus non-H-2A agricultural workers, employers with year-round crops versus crops with short peak demands, and indigenous agricultural workers cluster in only some areas or regions within a county that grow different crops, all have different labor forces.

Given the local differences, it is not unexpected that a there is little difference between the existing weight and the alternate weight calculation. This finding is consistent with other NAWS studies that show that the NAWS results are not sensitive to minor changes in the weights.

Table 2. Significant Sample Elements.

|  |  |
| --- | --- |
|  | **Significant Effects** |
| **NAWS Variable** | **FY 2011–2012** | **FY 2013–2014** | **FY 2011–2014** | **FY 2011–2014****Alternate Weights** |
| Wage | FYCycle\*RegionCounty ZipGrower | RegionGrower | Fiscal yearClusterCounty | RegionCountyZipGrower |
| FLC | RegionCountyZipGrower | RegionZipGrower | ZipGrower | RegionCountyZipGrower |
| Indigenous | RegionCounty | RegionCycleGrower | RegionGrower | ZipGrower |
| Unauthorized | RegionCounty clusterGrower | RegionCycleCountyGrower | RegionCountyGrower | RegionCycle\*RegionCountyGrower |
| Number of farm employers | None | Grower | RegionGrower | Grower |
| How the agricultural worker was paid  | RegionCountyGrower | CycleGrower | CountyGrower | RegionCountyGrower |
| Crop workdays | Cycle\*RegionGrower | CountyGrower | RegionCycleGrower | RegionGrower |
| Number of children | County | Grower | RegionCycle\*RegionClusterCountyGrower | CycleZipGrower |

Table 3. Results from PROC MIXED analysis on NAWS data FY 2011–2012.

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  | **Weighted** |
| **NAWS Variable** | **Source** | **DF** | **SS** | **%SS** | **MS** | **F Value** | **Prob<F** |
| Wage | FY | 1 | 36.320 | 0.2% | 36.32 | 3.38 | 0.0681 |
| CYCLE(FY) | 4 | 65.808 | 0.3% | 16.45 | 1.47 | 0.2157 |
| REGION12(FY) | 22 | 690.498 | 3.1% | 31.39 | 2.57 | 0.0015 |
| CYCLE\*REGION12(FY) | 44 | 1377.756 | 6.2% | 31.31 | 1.96 | 0.0053 |
| CLUS(FY\*CYCLE\*REGIO) | 82 | 1328.833 | 6.0% | 16.21 | 0.51 | 0.961 |
| Cou(FY\*CYC\*REG\*CLUS) | 15 | 585.643 | 2.6% | 39.04 | 1.68 | 0.0596 |
| Zi(FY\*CY\*RE\*CLU\*Cou) | 170 | 3270.773 | 14.8% | 19.24 | 2.07 | <.0001 |
| F(FY\*CY\*RE\*CL\*Co\*Zi) | 307 | 2729.883 | 12.3% | 8.89 | 1.72 | <.0001 |
| Residual | 2337 | 12072.000 | 54.5% | 5.17 |  |   |
|   | Total | 2982 | 22157.515 | 100.0% |  |  |   |
| FLC | FY | 1 | 0.025 | 0.0% | 0.02 | 0.26 | 0.614 |
| CYCLE(FY) | 4 | 0.259 | 0.2% | 0.06 | 0.63 | 0.6435 |
| REGION12(FY) | 22 | 6.484 | 3.9% | 0.29 | 2.64 | 0.004 |
| CYCLE\*REGION12(FY) | 44 | 2.044 | 1.2% | 0.05 | 0.27 | 1 |
| CLUS(FY\*CYCLE\*REGIO) | 84 | 15.639 | 9.3% | 0.19 | 0.29 | 0.9995 |
| Cou(FY\*CYC\*REG\*CLUS) | 15 | 12.066 | 7.2% | 0.80 | 2.49 | 0.0024 |
| Zi(FY\*CY\*RE\*CLU\*Cou) | 171 | 42.898 | 25.5% | 0.25 | 1.26 | 0.0427 |
| F(FY\*CY\*RE\*CL\*Co\*Zi) | 311 | 55.781 | 33.1% | 0.18 | 12.84 | <.0001 |
| Residual | 2372 | 33.141 | 19.7% | 0.01 |  |   |
|   | Total | 3024 | 168.337 | 100.0% |  |  |   |
| Indigenous | FY | 1 | 0.000 | 0.0% | 0.00 | 0 | 0.9636 |
| CYCLE(FY) | 4 | 0.278 | 0.1% | 0.07 | 1.03 | 0.3927 |
| REGION12(FY) | 22 | 2.720 | 1.3% | 0.12 | 1.91 | 0.0129 |
| CYCLE\*REGION12(FY) | 44 | 2.087 | 1.0% | 0.05 | 0.79 | 0.8035 |
| CLUS(FY\*CYCLE\*REGIO) | 84 | 5.299 | 2.5% | 0.06 | 0.66 | 0.9073 |
| Cou(FY\*CYC\*REG\*CLUS) | 15 | 1.300 | 0.6% | 0.09 | 2.96 | 0.0013 |
| Zi(FY\*CY\*RE\*CLU\*Cou) | 171 | 6.541 | 3.0% | 0.04 | 0.55 | 1 |
| F(FY\*CY\*RE\*CL\*Co\*Zi) | 311 | 21.891 | 10.2% | 0.07 | 0.95 | 0.6997 |
| Residual | 2372 | 174.961 | 81.3% | 0.07 |  |   |
|   | Total | 3024 | 215.076 | 100% |  |  |   |

Yellow = 0.01 significance. Blue = 0.10 significance. Orange = Negative F value.

Table 3 (Continued). Results from PROC MIXED analysis on NAWS data FY 2011–2012.

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  | **Weighted** |
| **NAWS Variable** | **Source** | **DF** | **SS** | **%SS** | **MS** | **F Value** | **Prob<F** |
| Unauthorized:  | FY | 1 | 0.018 | 0.0% | 0.02 | 0.05 | 0.83 |
| CYCLE(FY) | 4 | 0.301 | 0.0% | 0.08 | 0.19 | 0.944 |
| REGION12(FY) | 22 | 25.810 | 3.6% | 1.17 | 2.5 | 0.0011 |
| CYCLE\*REGION12(FY) | 44 | 18.686 | 2.6% | 0.42 | 0.73 | 0.8699 |
| CLUS(FY\*CYCLE\*REGIO) | 84 | 47.556 | 6.6% | 0.57 | 2.49 | 0.0998 |
| Cou(FY\*CYC\*REG\*CLUS) | 15 | 5.274 | 0.7% | 0.35 | 0.67 | 0.8076 |
| Zi(FY\*CY\*RE\*CLU\*Cou) | 171 | 75.644 | 10.6% | 0.44 | 1.18 | 0.1157 |
| F(FY\*CY\*RE\*CL\*Co\*Zi) | 309 | 109.733 | 15.3% | 0.36 | 1.92 | <.0001 |
| Residual | 2339 | 432.350 | 60.4% | 0.18 |  |   |
|   | Total | 2989 | 715.371 | 100.0% |  |  |   |
| Number of farm employers | FY | 1 | 0.001 | 0.0% | 0.00 | 0.01 | 0.9247 |
| CYCLE(FY) | 4 | 0.032 | 0.0% | 0.01 | 0.13 | 0.9731 |
| REGION12(FY) | 22 | 1.584 | 0.9% | 0.07 | 1.11 | 0.3424 |
| CYCLE\*REGION12(FY) | 44 | 0.842 | 0.5% | 0.02 | 0.29 | 1 |
| CLUS(FY\*CYCLE\*REGIO) | 84 | 5.429 | 2.9% | 0.06 | 1.54 | 0.1582 |
| Cou(FY\*CYC\*REG\*CLUS) | 15 | 0.621 | 0.3% | 0.04 | 0.65 | 0.8263 |
| Zi(FY\*CY\*RE\*CLU\*Cou) | 171 | 10.804 | 5.8% | 0.06 | 1 | 0.5015 |
| F(FY\*CY\*RE\*CL\*Co\*Zi) | 311 | 19.633 | 10.6% | 0.06 | 1.02 | 0.3928 |
| Residual | 2372 | 146.586 | 79.0% | 0.06 |  |   |
|   | Total | 3024 | 185.532 | 100.0% |  |  |   |
| How the agricultural worker was paid  | FY | 1 | 0.003 | 0.0% | 0.00 | 0.04 | 0.8409 |
| CYCLE(FY) | 4 | 0.176 | 0.1% | 0.04 | 0.64 | 0.6322 |
| REGION12(FY) | 22 | 2.894 | 1.7% | 0.13 | 1.95 | 0.0354 |
| CYCLE\*REGION12(FY) | 44 | 5.477 | 3.3% | 0.12 | 1.32 | 0.1678 |
| CLUS(FY\*CYCLE\*REGIO) | 84 | 9.202 | 5.5% | 0.11 | 0.23 | 1 |
| Cou(FY\*CYC\*REG\*CLUS) | 15 | 8.399 | 5.0% | 0.56 | 4.01 | <.0001 |
| Zi(FY\*CY\*RE\*CLU\*Cou) | 171 | 18.395 | 10.9% | 0.11 | 0.56 | 1 |
| F(FY\*CY\*RE\*CL\*Co\*Zi) | 310 | 54.117 | 32.1% | 0.17 | 5.93 | <.0001 |
| Residual | 2368 | 69.763 | 41.4% | 0.03 |  |   |
|   | Total | 3019 | 168.425 | 100.0% |  |  |   |

Yellow = 0.01 significance. Blue = 0.10 significance. Orange = Negative F value.

Table 3 (Continued). Results from PROC MIXED analysis on NAWS data FY 2011–2012.

|  |  |  |  |
| --- | --- | --- | --- |
|   |  |  | **Weighted** |
| **NAWS Variable** | **Source** | **DF** | **SS** | **%SS** | **MS** | **F Value** | **Prob<F** |
| Crop workdays | FY | 1 | 419.6 | 0.0% | 419.55 | 0.04 | 0.8365 |
| CYCLE(FY) | 4 | 70307.0 | 0.2% | 17577 | 1.76 | 0.1394 |
| REGION12(FY) | 22 | 808811.0 | 2.9% | 36764 | 3.4 | <.0001 |
| CYCLE\*REGION12(FY) | 44 | 1258992.0 | 4.4% | 28613 | 2.29 | 0.0005 |
| CLUS(FY\*CYCLE\*REGIO) | 84 | 1014625.0 | 3.6% | 12079 | 1.59 | 0.2816 |
| Cou(FY\*CYC\*REG\*CLUS) | 15 | 183258.0 | 0.6% | 12217 | 0.63 | 0.8476 |
| Zi(FY\*CY\*RE\*CLU\*Cou) | 171 | 2830477.0 | 10.0% | 16552 | 1.07 | 0.309 |
| F(FY\*CY\*RE\*CL\*Co\*Zi) | 311 | 4537626.0 | 16.0% | 14590 | 1.96 | <.0001 |
| Residual | 2370 | 17608544.0 | 62.2% | 7430 |  |   |
|   | Total | 3022 | 28313059.6 | 100.0% |  |  |   |
| Number of children | FY | 1 | 1.525 | 0.0% | 1.52 | 0.61 | 0.4355 |
| CYCLE(FY) | 4 | 3.851 | 0.1% | 0.96 | 0.38 | 0.8247 |
| REGION12(FY) | 22 | 82.753 | 1.3% | 3.76 | 1.45 | 0.1088 |
| CYCLE\*REGION12(FY) | 44 | 74.542 | 1.2% | 1.69 | 0.61 | 0.958 |
| CLUS(FY\*CYCLE\*REGIO) | 84 | 244.887 | 3.8% | 2.92 | 0.58 | 0.9504 |
| Cou(FY\*CYC\*REG\*CLUS) | 15 | 74.798 | 1.2% | 4.99 | 4.4 | <.0001 |
| Zi(FY\*CY\*RE\*CLU\*Cou) | 171 | 229.738 | 3.6% | 1.34 | 0.69 | 0.9948 |
| F(FY\*CY\*RE\*CL\*Co\*Zi) | 311 | 610.463 | 9.6% | 1.96 | 0.92 | 0.8306 |
| Residual | 2372 | 5065.963 | 79.3% | 2.14 |  |   |
|   | Total | 3024 | 6388.520 | 100.0% |   |   |   |

Yellow = 0.01 significance. Blue = 0.10 significance. Orange = Negative F value.

Table 4. Results from PROC MIXED analysis on NAWS data FY 2013–2014.

|  |  |  |  |
| --- | --- | --- | --- |
|   |  |  | **Weighted** |
| **NAWS Variable** | **Source** | **DF** | **SS** | **%SS** | **MS** | **F Value** | **Prob<F** |
| Wage  | FY | 1 | 0 | 0.00% | 0 | 0 | 1 |
| CYCLE(FY) | 4 | 27.24 | 0.09% | 6.81 | 0.56 | 0.6889 |
| REGION12(FY) | 22 | 488.2 | 1.60% | 22.191 | 1.53 | 0.0694 |
| CYCLE\*REGION12(FY) | 44 | 776.702 | 2.55% | 17.652 | 0.99 | 0.4973 |
| CLUS(FY\*CYCLE\*REGIO) | 125 | 2775.667 | 9.11% | 22.205 | 5.45 | 0.3094 |
| Cou(FY\*CYC\*REG\*CLUS) | 11 | 100.067 | 0.33% | 9.097 | 0.59 | 0.8389 |
| Zi(FY\*CY\*RE\*CLU\*Cou) | 141 | 2369.762 | 7.78% | 16.806 | 1.09 | 0.2457 |
| F(FY\*CY\*RE\*CL\*Co\*Zi) | 549 | 7433.598 | 24.39% | 13.54 | 2.67 | <.0001 |
| Residual | 3256 | 16505 | 54.16% | 5.069 |  |   |
|   | Total | 4153 | 30476.24 | 100.00% |  |  |   |
| FLC | FY | 1 | 0.027 | 0.01% | 0.027 | 0.55 | 0.4598 |
| CYCLE(FY) | 4 | 0.0542 | 0.02% | 0.014 | 0.23 | 0.9184 |
| REGION12(FY) | 22 | 4.18 | 1.64% | 0.19 | 2.67 | 0.0002 |
| CYCLE\*REGION12(FY) | 44 | 1.442 | 0.57% | 0.033 | 0.37 | 0.9999 |
| CLUS(FY\*CYCLE\*REGIO) | 126 | 14.537 | 5.72% | 0.115 | -0.59 | 1 |
| Cou(FY\*CYC\*REG\*CLUS) | 11 | 0.016 | 0.01% | 0.001 | 0 | 1 |
| Zi(FY\*CY\*RE\*CLU\*Cou) | 140 | 66.46 | 26.14% | 0.475 | 1.87 | <.0001 |
| F(FY\*CY\*RE\*CL\*Co\*Zi) | 549 | 116.21 | 45.72% | 0.212 | 13.78 | <.0001 |
| Residual | 3337 | 51.275 | 20.17% | 0.015 |  |   |
|   | Total | 4234 | 254.2012 | 100.00% |  |  |   |
| Indigenous | FY | 1 | 0.035 | 0.02% | 0.035 | 0.72 | 0.3978 |
| CYCLE(FY) | 4 | 0.19 | 0.10% | 0.048 | 0.92 | 0.4493 |
| REGION12(FY) | 22 | 2.017 | 1.02% | 0.092 | 1.72 | 0.0239 |
| CYCLE\*REGION12(FY) | 44 | 3.572 | 1.81% | 0.081 | 1.4 | 0.0601 |
| CLUS(FY\*CYCLE\*REGIO) | 126 | 7.742 | 3.93% | 0.061 | -1.78 | 1 |
| Cou(FY\*CYC\*REG\*CLUS) | 11 | 0.039 | 0.02% | 0.004 | 0.06 | 1 |
| Zi(FY\*CY\*RE\*CLU\*Cou) | 140 | 9.252 | 4.70% | 0.066 | 1.08 | 0.2794 |
| F(FY\*CY\*RE\*CL\*Co\*Zi) | 549 | 31.762 | 16.13% | 0.058 | 1.36 | <.0001 |
| Residual | 3337 | 142.25 | 72.26% | 0.043 |  |   |
|   | Total | 4234 | 196.859 | 100% |   |   |   |

Yellow = 0.01 significance. Blue = 0.10 significance. Orange = Negative F value.

Table 4 (Continued). Results from PROC MIXED analysis on NAWS data FY 2013–2014.

|  |  |  |  |
| --- | --- | --- | --- |
|   |  |  | **Weighted** |
| **NAWS Variable** | **Source** | **DF** | **SS** | **%SS** | **MS** | **F Value** | **Prob<F** |
| Unauthorized  | FY | 1 | 0.222 | 0.03% | 0.222 | 1.07 | 0.3012 |
| CYCLE(FY) | 4 | 0.57 | 0.07% | 0.142 | 0.64 | 0.6368 |
| REGION12(FY) | 22 | 12.214 | 1.45% | 0.555 | 2.26 | 0.002 |
| CYCLE\*REGION12(FY) | 44 | 16.517 | 1.95% | 0.375 | 1.46 | 0.0763 |
| CLUS(FY\*CYCLE\*REGIO) | 126 | 48.601 | 5.75% | 0.386 | 0.32 | 0.9969 |
| Cou(FY\*CYC\*REG\*CLUS) | 11 | 9.888 | 1.17% | 0.899 | 3.04 | 0.0007 |
| Zi(FY\*CY\*RE\*CLU\*Cou) | 140 | 40.25 | 4.76% | 0.287 | 0.79 | 0.9525 |
| F(FY\*CY\*RE\*CL\*Co\*Zi) | 549 | 180.141 | 21.31% | 0.328 | 2.02 | <.0001 |
| Residual | 3304 | 536.768 | 63.51% | 0.162 |  |   |
|   | Total | 4201 | 845.171 | 100.00% |  |  |   |
| Number of farm employers | FY | 1 | 0.0001 | 0.00% | 0.001 | 0 | 0.9602 |
| CYCLE(FY) | 4 | 0.041 | 0.02% | 0.01 | 0.2 | 0.9403 |
| REGION12(FY) | 22 | 1.499 | 0.58% | 0.068 | 1.34 | 0.134 |
| CYCLE\*REGION12(FY) | 44 | 1.747 | 0.67% | 0.04 | 0.8 | 0.815 |
| CLUS(FY\*CYCLE\*REGIO) | 126 | 5.223 | 2.01% | 0.041 | -0.64 | 1 |
| Cou(FY\*CYC\*REG\*CLUS) | 11 | 0.02 | 0.01% | 0.002 | 0.02 | 1 |
| Zi(FY\*CY\*RE\*CLU\*Cou) | 140 | 11.996 | 4.61% | 0.086 | 0.82 | 0.9215 |
| F(FY\*CY\*RE\*CL\*Co\*Zi) | 549 | 52.747 | 20.26% | 0.096 | 1.71 | <.0001 |
| Residual | 3337 | 187.07 | 71.86% | 0.056 |  |   |
|   | Total | 4234 | 260.3431 | 100.00% |  |  |   |
| How the agricultural worker was paid  | FY | 1 | 0.01 | 0.01% | 0.01 | 0.27 | 0.6063 |
| CYCLE(FY) | 4 | 0.274 | 0.15% | 0.068 | 1.57 | 0.1839 |
| REGION12(FY) | 22 | 1.413 | 0.77% | 0.064 | 1.37 | 0.1324 |
| CYCLE\*REGION12(FY) | 44 | 4.608 | 2.52% | 0.105 | 1.85 | 0.0039 |
| CLUS(FY\*CYCLE\*REGIO) | 126 | 10.208 | 5.58% | 0.081 | -0.76 | 1 |
| Cou(FY\*CYC\*REG\*CLUS) | 11 | 0.206 | 0.11% | 0.019 | 0.2 | 0.9973 |
| Zi(FY\*CY\*RE\*CLU\*Cou) | 140 | 8.781 | 4.80% | 0.063 | 0.3 | 1 |
| F(FY\*CY\*RE\*CL\*Co\*Zi) | 549 | 96.164 | 52.61% | 0.175 | 9.54 | <.0001 |
| Residual | 3330 | 61.132 | 33.44% | 0.018 |  |   |
|   | Total | 4227 | 182.796 | 100.00% |   |   |   |

Yellow = 0.01 significance. Blue = 0.10 significance. Orange = Negative F value.

Table 4 (Continued). Results from PROC MIXED analysis on NAWS data FY 2013–2014.

|  |  |  |  |
| --- | --- | --- | --- |
|   |  |  | **Weighted** |
| **NAWS Variable** | **Source** | **DF** | **SS** | **%SS** | **MS** | **F Value** | **Prob<F** |
| Crop workdays | FY | 1 | 10705 | 0.03% | 10705 | 0.84 | 0.359 |
| CYCLE(FY) | 4 | 37086 | 0.10% | 9271.388 | 0.62 | 0.6505 |
| REGION12(FY) | 22 | 343555 | 0.96% | 15616 | 0.9 | 0.594 |
| CYCLE\*REGION12(FY) | 44 | 466319 | 1.31% | 10598 | 0.51 | 0.9944 |
| CLUS(FY\*CYCLE\*REGIO) | 126 | 3748514 | 10.51% | 29750 | 1.04 | 0.5413 |
| Cou(FY\*CYC\*REG\*CLUS) | 11 | 255934 | 0.72% | 23267 | 1.89 | 0.0401 |
| Zi(FY\*CY\*RE\*CLU\*Cou) | 140 | 1726483 | 4.84% | 12332 | 0.88 | 0.8197 |
| F(FY\*CY\*RE\*CL\*Co\*Zi) | 549 | 6980445 | 19.58% | 12715 | 1.92 | <.0001 |
| Residual | 3336 | 22089112 | 61.95% | 6621.437 |  |   |
|   | Total | 4233 | 35658153 | 100.00% |  |  |   |
| Number of children | FY | 1 | 0.441 | 0.01% | 0.441 | 0.24 | 0.6262 |
| CYCLE(FY) | 4 | 7.698 | 0.11% | 1.924 | 0.98 | 0.4163 |
| REGION12(FY) | 22 | 48.517 | 0.68% | 2.205 | 1.08 | 0.3703 |
| CYCLE\*REGION12(FY) | 44 | 75.933 | 1.06% | 1.726 | 0.78 | 0.8379 |
| CLUS(FY\*CYCLE\*REGIO) | 126 | 317.724 | 4.43% | 2.522 | 2.59 | 0.3363 |
| Cou(FY\*CYC\*REG\*CLUS) | 11 | 16.015 | 0.22% | 1.456 | 0.73 | 0.7139 |
| Zi(FY\*CY\*RE\*CLU\*Cou) | 140 | 280.235 | 3.90% | 2.002 | 0.92 | 0.709 |
| F(FY\*CY\*RE\*CL\*Co\*Zi) | 549 | 1133.444 | 15.79% | 2.065 | 1.3 | <.0001 |
| Residual | 3337 | 5296.403 | 73.80% | 1.587 |  |   |
|   | Total | 4234 | 7176.41 | 100.00% |   |   |   |

Yellow = 0.01 significance. Blue = 0.10 significance. Orange = Negative F value.

Table 5. Results from PROC MIXED analysis on NAWS data FY 2011–2014.

|  |  |  |  |
| --- | --- | --- | --- |
|   |  |  | **Weighted** |
| **NAWS Variable** | **Source** | **DF** | **SS** | **%SS** | **MS** | **F Value** | **Prob<F** |
| Wage  | FY | 3 | 28.5001 | 0.00% | 9.5 | 4.32 | 0.0059 |
| CYCLE(FY) | 8 | 68.6817 | 0.00% | 8.5852 | -5.37 | 1 |
| REGION12(FY) | 44 | 1113.3968 | 0.00% | 25.3045 | -5.63 | 1 |
| CYCLE\*REGION12(FY) | 88 | 1790.8131 | 0.00% | 20.3501 | -2.71 | 1 |
| CLUS(FY\*CYCLE\*REGIO) | 207 | 3.71E+13 | 100.00% | 1.782E+11 | 43.19 | <.0001 |
| Cou(FY\*CYC\*REG\*CLUS) | 26 | 700.8112 | 0.00% | 26.9543 | 1.39 | 0.0999 |
| Zi(FY\*CY\*RE\*CLU\*Cou) | 310 | 5601.0309 | 0.00% | 18.0097 | 1.38 | 0.0003 |
| F(FY\*CY\*RE\*CL\*Co\*Zi) | 856 | 10131 | 0.00% | 11.8351 | 2.33 | <.0001 |
| Residual | 5594 | 28437 | 0.00% | 5.0853 |  |   |
|   | Total | 7136 | 3.706E+13 | 100.00% |   |   |   |
| FLC | FY | 3 | 0.0144 | 0.00% | 0.0048 | -0.09 | 1.0000 |
| CYCLE(FY) | 8 | 0.3107 | 0.00% | 0.0388 | -0.25 | 1 |
| REGION12(FY) | 44 | 9.1529 | 0.00% | 0.208 | -1.04 | 1 |
| CYCLE\*REGION12(FY) | 88 | 35.8444 | 0.00% | 0.4073 | -0.49 | 1 |
| CLUS(FY\*CYCLE\*REGIO) | 210 | 9217516990 | 100.00% | 43684915 | 0.17 | 1 |
| Cou(FY\*CYC\*REG\*CLUS) | 26 | 11.9887 | 0.00% | 0.4611 | 1.2 | 0.2302 |
| Zi(FY\*CY\*RE\*CLU\*Cou) | 311 | 109.0457 | 0.00% | 0.3495 | 1.5 | <.0001 |
| F(FY\*CY\*RE\*CL\*Co\*Zi) | 860 | 171.5981 | 0.00% | 0.1995 | 13.53 | <.0001 |
| Residual | 5709 | 84.1752 | 0.00% | 0.0148 |  |   |
|   | Total | 7259 | 9217517412 | 100.00% |  |  |   |
| Indigenous | FY | 3 | 0.002 | 0.00% | 0.0007 | 0.01 | 0.9982 |
| CYCLE(FY) | 8 | 0.3915 | 0.00% | 0.0489 | 0.91 | 0.5076 |
| REGION12(FY) | 44 | 4.3526 | 0.00% | 0.0989 | 1.85 | 0.0009 |
| CYCLE\*REGION12(FY) | 88 | 5.4957 | 0.00% | 0.083 | 1.39 | 0.1193 |
| CLUS(FY\*CYCLE\*REGIO) | 210 | 308653938 | 100.00% | 0.0834 | -1.13 | 1 |
| Cou(FY\*CYC\*REG\*CLUS) | 26 | 1.3191 | 0.00% | 0.019 | 0.98 | 0.4933 |
| Zi(FY\*CY\*RE\*CLU\*Cou) | 311 | 15.7417 | 0.00% | 0.109 | 0.8 | 0.9892 |
| F(FY\*CY\*RE\*CL\*Co\*Zi) | 860 | 53.4709 | 0.00% | 0.109 | 1.12 | 0.0111 |
| Residual | 5,709 | 315.949 | 0.00% | 0.052 |  |   |
|   | Total | 7,259 | 308654334.7 | 100% |   |   |   |

Yellow = 0.01 significance. Blue = 0.10 significance. Orange = Negative F value.

Table 5 (Continued). Results from PROC MIXED analysis on NAWS data from FY 2011–2014.

|  |  |  |  |
| --- | --- | --- | --- |
|   |  |  | **Weighted** |
| **NAWS Variable** | **Source** | **DF** | **SS** | **%SS** | **MS** | **F Value** | **Prob<F** |
| Unauthorized  | FY | 3 | 0.0279 | 0.00% | 0.0093 | 0.03 | 0.9913 |
| CYCLE(FY) | 8 | 0.6978 | 0.04% | 0.0872 | 0.3 | 0.9655 |
| REGION12(FY) | 44 | 37.7851 | 2.43% | 0.8588 | 2.61 | <.0001 |
| CYCLE\*REGION12(FY) | 88 | 34.6133 | 2.23% | 0.3933 | 1.01 | 0.4652 |
| CLUS(FY\*CYCLE\*REGIO) | 210 | 95.8132 | 6.16% | 0.4541 | 0.85 | 0.7117 |
| Cou(FY\*CYC\*REG\*CLUS) | 26 | 15.1243 | 0.97% | 0.5817 | 1.44 | 0.0774 |
| Zi(FY\*CY\*RE\*CLU\*Cou) | 311 | 115.3372 | 7.42% | 0.3709 | 1.01 | 0.4625 |
| F(FY\*CY\*RE\*CL\*Co\*Zi) | 858 | 289.0733 | 18.60% | 0.3369 | 1.97 | <.0001 |
| Residual | 5643 | 966.0589 | 62.14% | 0.1712 |  |   |
|   | Total | 7191 | 1554.531 | 100.00% |  |  |   |
| Number of farm employers | FY | 3 | 0.004 | 0.00% | 0.0013 | 0.03 | 0.9946 |
| CYCLE(FY) | 8 | 0.0624 | 0.00% | 0.0078 | 0.17 | 0.9947 |
| REGION12(FY) | 44 | 2.5698 | 0.00% | 0.0584 | 1.34 | 0.0796 |
| CYCLE\*REGION12(FY) | 88 | 7.2795 | 0.00% | 0.0827 | -6.41 | 1 |
| CLUS(FY\*CYCLE\*REGIO) | 210 | 32689216275 | 100.00% | 154925196 | -19.3 | 1 |
| Cou(FY\*CYC\*REG\*CLUS) | 26 | 0.6309 | 0.00% | 0.0243 | 0.31 | 0.9996 |
| Zi(FY\*CY\*RE\*CLU\*Cou) | 311 | 22.7212 | 0.00% | 0.0728 | 0.82 | 0.9773 |
| F(FY\*CY\*RE\*CL\*Co\*Zi) | 860 | 72.2348 | 0.00% | 0.084 | 1.44 | <.0001 |
| Residual | 5709 | 332.5911 | 0.00% | 0.0583 |  |   |
|   | Total | 7259 | 32689216713 | 100.00% |  |  |   |
| How the agricultural worker was paid  | FY | 3 | 0.0048 | 0.00% | 0.0016 | -0.13 | 1 |
| CYCLE(FY) | 8 | 0.4393 | 0.00% | 0.0549 | -0.81 | 1 |
| REGION12(FY) | 44 | 4.1709 | 0.00% | 0.0948 | -0.9 | 1 |
| CYCLE\*REGION12(FY) | 88 | 26.1659 | 0.00% | 0.2973 | -0.74 | 1 |
| CLUS(FY\*CYCLE\*REGIO) | 210 | 1326627221 | 100.00% | 6287333 | 0.06 | 1 |
| Cou(FY\*CYC\*REG\*CLUS) | 26 | 8.5387 | 0.00% | 0.3284 | 2.96 | <.0001 |
| Zi(FY\*CY\*RE\*CLU\*Cou) | 311 | 27.0397 | 0.00% | 0.0867 | 0.43 | 1 |
| F(FY\*CY\*RE\*CL\*Co\*Zi) | 859 | 149.8953 | 0.00% | 0.1745 | 7.62 | <.0001 |
| Residual | 5,698 | 130.4381 | 0.00% | 0.0229 |  |   |
|   | Total | 7,247 | 1326627568 | 100.00% |   |   |   |

Yellow = 0.01 significance. Blue = 0.10 significance. Orange = Negative F value.

Table 5 (Continued). Results from PROC MIXED analysis on NAWS data from FY 2011–2014.

|  |  |  |  |
| --- | --- | --- | --- |
|   |  |  | **Weighted** |
| **NAWS Variable** | **Source** | **DF** | **SS** | **%SS** | **MS** | **F Value** | **Prob<F** |
| Crop workdays | FY | 3 | 9104.1261 | 0.01% | 3034.7087 | 0.25 | 0.8621 |
| CYCLE(FY) | 8 | 89735 | 0.14% | 11217 | 0.82 | 0.5821 |
| REGION12(FY) | 44 | 1047362 | 1.61% | 23804 | 1.58 | 0.0143 |
| CYCLE\*REGION12(FY) | 88 | 3376998 | 5.18% | 38375 | 2.02 | <.0001 |
| CLUS(FY\*CYCLE\*REGIO) | 210 | 4653996 | 7.14% | 22057 | 1.58 | 0.1767 |
| Cou(FY\*CYC\*REG\*CLUS) | 26 | 437863 | 0.67% | 16841 | 1.07 | 0.3791 |
| Zi(FY\*CY\*RE\*CLU\*Cou) | 311 | 4534980 | 6.95% | 14582 | 1 | 0.4888 |
| F(FY\*CY\*RE\*CL\*Co\*Zi) | 860 | 11485999 | 17.61% | 13356 | 1.93 | <.0001 |
| Residual | 5706 | 39570101 | 60.68% | 6936.0388 |  |   |
|   | Total | 7256 | 65206138.13 | 100.00% |  |  |   |
| HHKID: Number of kids in household | FY | 3 | 1.1666 | 0.00% | 0.3889 | 0.2 | 0.8821 |
| CYCLE(FY) | 8 | 11.1769 | 0.00% | 1.3971 | 0.83 | 0.58 |
| REGION12(FY) | 44 | 140.5282 | 0.00% | 3.1938 | 1.95 | 0.0004 |
| CYCLE\*REGION12(FY) | 88 | 1772.949 | 0.00% | 20.1471 | 15.92 | <.0001 |
| CLUS(FY\*CYCLE\*REGIO) | 210 | 3.92E+12 | 100.00% | 1.859E+10 | 27.57 | 0.0002 |
| Cou(FY\*CYC\*REG\*CLUS) | 26 | 88.5032 | 0.00% | 3.404 | 2.04 | 0.0025 |
| Zi(FY\*CY\*RE\*CLU\*Cou) | 311 | 508.3302 | 0.00% | 1.6293 | 0.79 | 0.9909 |
| F(FY\*CY\*RE\*CL\*Co\*Zi) | 860 | 1739.4251 | 0.00% | 2.0226 | 1.12 | 0.014 |
| Residual | 5709 | 10326 | 0.00% | 1.8093 |  |   |
|   | Total | 7259 | 3.9216E+12 | 100.00% |   |   |   |

Yellow = 0.01 significance. Blue = 0.10 significance. Orange = Negative F value.

Table 6. Results from PROC MIXED analysis using alternate weights on NAWS data FY 2011–2014.

|  |  |  |  |
| --- | --- | --- | --- |
|   |  |  | **Alternate Weighted** |
| **NAWS Variable** | **Source** | **DF** | **SS** | **%SS** | **MS** | **F Value** | **Prob<F** |
| Wage | FY | 3 | 12.5785 | 0.00% | 4.1928 | 0.7 | 0.5502 |
| CYCLE(FY) | 8 | 56.3181 | 0.00% | 7.0398 | 0.72 | 0.671 |
| REGION12(FY) | 44 | 387.9777 | 0.00% | 8.8177 | 1.49 | 0.0204 |
| CYCLE\*REGION12(FY) | 88 | 2789024 | 0.00% | 31693 | -11 | 1 |
| CLUS(FY\*CYCLE\*REGIO) | 207 | 4.01E+03 | 0.00% | 19.384 | 0.92 | 0.6319 |
| Cou(FY\*CYC\*REG\*CLUS) | 26 | 562.1268 | 0.00% | 21.6203 | 1.8 | 0.0083 |
| Zi(FY\*CY\*RE\*CLU\*Cou) | 310 | 5.35E+13 | 100.00% | 1.73E+11 | 342.13 | <.0001 |
| F(FY\*CY\*RE\*CL\*Co\*Zi) | 858 | 10980 | 0.00% | 12.7972 | 2.52 | <.0001 |
| Residual | 5592 | 28413 | 0.00% | 5.081 |  |   |
|   | Total | 7136 | 5.35E+13 | 100.00% |   |   |   |
| FLC | FY | 3 | 0.0225 | 0.00% | 0.0075 | 0.22 | 0.8832 |
| CYCLE(FY) | 8 | 0.1272 | 0.00% | 0.0159 | 0.32 | 0.9593 |
| REGION12(FY) | 44 | 1.9948 | 0.01% | 0.0453 | 2.13 | <.0001 |
| CYCLE\*REGION12(FY) | 88 | 17.3945 | 0.06% | 0.1977 | 0.02 | 1 |
| CLUS(FY\*CYCLE\*REGIO) | 210 | 29.2476 | 0.10% | 0.1393 | 0.29 | 1 |
| Cou(FY\*CYC\*REG\*CLUS) | 26 | 11.7641 | 0.04% | 0.4525 | 2.54 | <.0001 |
| Zi(FY\*CY\*RE\*CLU\*Cou) | 311 | 29236 | 98.89% | 94.0052 | 0 | 1 |
| F(FY\*CY\*RE\*CL\*Co\*Zi) | 862 | 176.2472 | 0.60% | 0.2045 | 12.84 | <.0001 |
| Residual | 5707 | 90.9129 | 0.31% | 0.0159 |  |   |
|   | Total | 7259 | 29563.71 | 100.00% |  |  |   |
| Indigenous | FY | 3 | 0.0026 | 0.00% | 0.0009 | 0.02 | 0.9974 |
| CYCLE(FY) | 8 | 0.3343 | 0.00% | 0.0418 | 0.74 | 0.6561 |
| REGION12(FY) | 44 | 0.3836 | 0.00% | 0.0087 | 0.16 | 1 |
| CYCLE\*REGION12(FY) | 88 | 1.5175 | 0.00% | 0.0172 | 0.02 | 1 |
| CLUS(FY\*CYCLE\*REGIO) | 210 | 12.3398 | 0.00% | 0.0588 | 4.96 | 0.1312 |
| Cou(FY\*CYC\*REG\*CLUS) | 26 | 0.7232 | 0.00% | 0.0278 | 0.43 | 0.9947 |
| Zi(FY\*CY\*RE\*CLU\*Cou) | 311 | 6.34E+09 | 100.00% | 20393874 | 27.91 | <.0001 |
| F(FY\*CY\*RE\*CL\*Co\*Zi) | 862 | 57.1676 | 0.00% | 0.0663 | 1.22 | <.0001 |
| Residual | 5707 | 311.2675 | 0.00% | 0.0545 |  |   |
|   | Total | 7259 | 6.34E+09 | 100% |   |   |   |

Yellow = 0.01 significance. Blue = 0.10 significance. Orange = Negative F value.

Table 6 (Continued). Results from PROC MIXED analysis using alternate weights on NAWS data FY 2011–2014.

|  |  |  |  |
| --- | --- | --- | --- |
|   |  |  | **Alternate Weighted** |
| **NAWS Variable** | **Source** | **DF** | **SS** | **%SS** | **MS** | **F Value** | **Prob<F** |
| Unauthorized | FY | 3 | 5.01E-11 | 0.00% | 1.67E-11 | 0 | 1 |
| CYCLE(FY) | 8 | 0.4396 | 0.00% | 0.0549 | 0.23 | 0.9846 |
| REGION12(FY) | 44 | 35.0851 | 0.00% | 0.7974 | 2.8 | <.0001 |
| CYCLE\*REGION12(FY) | 88 | 7.8E+10 | 34.42% | 8.86E+08 | 3.07 | <.0001 |
| CLUS(FY\*CYCLE\*REGIO) | 210 | 81.1948 | 0.00% | 0.3866 | 0.86 | 0.6908 |
| Cou(FY\*CYC\*REG\*CLUS) | 26 | 12.4558 | 0.00% | 0.4791 | 1.51 | 0.0486 |
| Zi(FY\*CY\*RE\*CLU\*Cou) | 311 | 1.49E+11 | 65.58% | 4.78E+08 | 0.94 | 0.6981 |
| F(FY\*CY\*RE\*CL\*Co\*Zi) | 861 | 330.3556 | 0.00% | 0.383688 | 2.25 | <.0001 |
| Residual | 5640 | 959.9278 | 0.00% | 0.1702 |  |   |
|   | Total | 7191 | 2.27E+11 | 100.00% |  |  |   |
| Number of farm employers | FY | 3 | 0.0136 | 0.00% | 0.0045 | 0.08 | 0.9722 |
| CYCLE(FY) | 8 | 0.0517 | 0.00% | 0.0646 | 0.11 | 0.9987 |
| REGION12(FY) | 44 | 0.4276 | 0.00% | 0.0097 | 0.16 | 1 |
| CYCLE\*REGION12(FY) | 88 | 1.5355 | 0.00% | 0.0174 | -0.14 | 1 |
| CLUS(FY\*CYCLE\*REGIO) | 210 | 10.0399 | 0.00% | 0.0478 | 7.38 | 0.4297 |
| Cou(FY\*CYC\*REG\*CLUS) | 26 | 0.8283 | 0.00% | 0.0319 | 0.39 | 0.9978 |
| Zi(FY\*CY\*RE\*CLU\*Cou) | 311 | 1.94E+08 | 100.00% | 622190 | 0.38 | 1 |
| F(FY\*CY\*RE\*CL\*Co\*Zi) | 862 | 74.4382 | 0.00% | 0.0864 | 1.44 | <.0001 |
| Residual | 5707 | 341.8114 | 0.00% | 0.0599 |  |   |
|   | Total | 7259 | 1.94E+08 | 100.00% |  |  |   |
| How the agricultural worker was paid  | FY | 3 | 0.0118 | 0.00% | 0.0393 | 0.16 | 0.9234 |
| CYCLE(FY) | 8 | 0.2651 | 0.00% | 0.0331 | 0.87 | 0.5431 |
| REGION12(FY) | 44 | 3.6048 | 0.00% | 0.0819 | 1.77 | 0.0078 |
| CYCLE\*REGION12(FY) | 88 | 1075.284 | 0.00% | 12.2191 | 28.44 | 0.7352 |
| CLUS(FY\*CYCLE\*REGIO) | 210 | 17.7519 | 0.00% | 0.0845 | 0.25 | 1 |
| Cou(FY\*CYC\*REG\*CLUS) | 26 | 8.7125 | 0.00% | 0.3351 | 2.01 | 0.0021 |
| Zi(FY\*CY\*RE\*CLU\*Cou) | 311 | 27016562 | 99.99% | 86870 | 0.01 | 1 |
| F(FY\*CY\*RE\*CL\*Co\*Zi) | 861 | 163.9983 | 0.00% | 0.1905 | 8.81 | <.0001 |
| Residual | 5696 | 123.0968 | 0.00% | 0.0216 |  |   |
|   | Total | 7247 | 27017955 | 100.00% |   |   |   |

Yellow = 0.01 significance. Blue = 0.10 significance. Orange = Negative F value.

Table 6 (Continued). Results from PROC MIXED analysis using alternate weights on NAWS data FY 2011–2014.

|  |  |  |  |
| --- | --- | --- | --- |
|   |  |  | **Alternate Weighted** |
| **NAWS Variable** | **Source** | **DF** | **SS** | **%SS** | **MS** | **F Value** | **Prob<F** |
| Crop workdays | FY | 3 | 9768.902 | 0.02% | 3256.301 | 0.34 | 0.7951 |
| CYCLE(FY) | 8 | 95253 | 0.15% | 11907 | 1.01 | 0.4258 |
| REGION12(FY) | 44 | 918173 | 1.41% | 20868 | 1.4 | 0.0566 |
| CYCLE\*REGION12(FY) | 88 | 1649773 | 2.54% | 18747 | 1.01 | 0.4579 |
| CLUS(FY\*CYCLE\*REGIO) | 210 | 4501598 | 6.92% | 21436 | 1.59 | 0.1961 |
| Cou(FY\*CYC\*REG\*CLUS) | 26 | 401363 | 0.62% | 15437 | 1.1 | 0.3321 |
| Zi(FY\*CY\*RE\*CLU\*Cou) | 311 | 4093862 | 6.29% | 13164 | 0.91 | 0.8344 |
| F(FY\*CY\*RE\*CL\*Co\*Zi) | 862 | 12520678 | 19.24% | 14525 | 2.03 | <.0001 |
| Residual | 5704 | 40885893 | 62.83% | 7167.934 |  |   |
|   | Total | 7256 | 65076362 | 100.00% |  |  |   |
| HHKID: Number of kids in household | FY | 3 | 1.0023 | 0.00% | 0.3341 | 0.18 | 0.9076 |
| CYCLE(FY) | 8 | 8.4057 | 0.00% | 1.0507 | 0.56 | 0.81 |
| REGION12(FY) | 44 | 22.8867 | 0.00% | 0.5202 | 0.29 | 1 |
| CYCLE\*REGION12(FY) | 88 | 14575 | 0.00% | 165.6236 | 5.43 | 0.0674 |
| CLUS(FY\*CYCLE\*REGIO) | 210 | 4.48E+02 | 0.00% | 2.1336 | 0.79 | 0.7722 |
| Cou(FY\*CYC\*REG\*CLUS) | 26 | 69.6914 | 0.00% | 2.6804 | 1.24 | 0.1886 |
| Zi(FY\*CY\*RE\*CLU\*Cou) | 311 | 1.43E+11 | 100.00% | 4.6E+08 | 15.99 | <.0001 |
| F(FY\*CY\*RE\*CL\*Co\*Zi) | 862 | 1918.4 | 0.00% | 2.2255 | 1.26 | <.0001 |
| Residual | 5707 | 10056 | 0.00% | 1.7621 |  |   |
|   | Total | 7259 | 1.43E+11 | 100.00% |   |   |   |

Yellow = 0.01 significance. Blue = 0.10 significance. Orange = Negative F value.

# Appendix F: Design Study B – Optimal Interview Allocations for NAWS Sampling

**Optimal Interview Allocations for NAWS Sampling**

The purpose of this study was to determine optimal interview allocations (i.e., the number of employees who should be interviewed) for the National Agricultural Workers Survey. The NAWS statisticians calculated optimal allocations for cost reduction and for statistical efficiency (minimizing standard errors) for each of the 3 cycles and 12 sampling regions used to stratify the NAWS sample. The goal was to gain more information about how to reduce interviewing costs and improve the precision of point estimates.

## Method

The NAWS data for fiscal years 2016–2017 were used for this analysis. The data include interview allocations, population sizes, and marginal travel costs per interview for the 36 sampling strata (3 cycles times 12 regions).

Optimal allocations were calculated for nine variables that are considered key findings from the NAWS. These findings have been used for policy or program planning. The selected variables included four continuous variables and five binary variables. The continuous variables were:

* The worker’s hourly wage or hourly equivalent wage if a piece rate worker;
* Number of farm employers in the past 12 months;
* Number of farm work days in the past 12 months; and
* Number of children in the household.

The binary variables were coded as one if:

* The employer was an agricultural producer and not a labor contractor;
* The worker lacked work authorization;
* The worker had only one farm employer;
* The worker was paid an hourly wage as opposed to a piece rate or salary; and
* The number of children in household was three or fewer.

These five binary variables were coded as zero otherwise.

For each of the 36 NAWS strata, a set of 10 allocations were calculated, one for each of the nine variables and an overall allocation which was the result of averaging the allocations for the nine variables. For each of the 36 strata, 2 sets of 10 allocations were calculated. The optimal allocation achieves both statistical and cost efficiency. The Neyman allocation is a special case of optimal allocation that assumes the cost of each stratum is approximately equal and thus calculates statistical efficiency only.

Each of the optimal allocations was calculated using the following equation for stratum *h*:

$$n\_{h =}\frac{\left(\frac{N\_{h}×S\_{h}}{\sqrt{C\_{h}}}\right)}{\sum\_{}^{}\left(\frac{N\_{h} ×S\_{h}}{\sqrt{C\_{h}}}\right)}×n$$

where

 nh = allocation size

Nh = population estimate (average 2016-2017)

Sh = standard deviation

Ch = unit cost (cost per interview)

Ʃ = total for all three cycles and 12 regions

n = total sample size

Each of the Neyman allocations was calculated using the following equation for stratum *h*:

$$n\_{h =}\frac{N\_{h}×S\_{h}}{\sum\_{}^{}\left(N\_{h}×S\_{h}\right)}×n$$

where

 nh = allocation size

Nh = population estimate (average)

Sh = standard deviation

Ʃ = total for all three cycles and 12 regions

n = total sample size

## Results

In general, the results showed that optimal allocations would increase interview allocations in all three cycles for the largest farm labor regions: California (CA) and the Pacific Northwest (PC). Florida (FL) showed modest increases in the fall and spring cycles. The western mountain states regions – Mountain I, II (MT12), and Mountain III (MT3) – had increased allocations in the summer cycles. Allocations for the remaining five regions – Northeast I (NE1), Northeast II (NE2), Appalachia (AP), Corn Belt/Northern Plains (CBNP), and Southern Plains (SP) – were reduced in all cycles.

As would be expected from the formula, optimal allocations differed by variable with variables having higher standard errors requiring larger allocations than those with lower standard errors. This led to large increases in allocations for certain variables for specific regions. In California, the optimal allocation for the type of employer increased the sample size by more than 200 interviews in the fall and winter cycles, which was more than a 50 percent increase. In the summer cycle, the allocation increased by 185 interviews, which was a 44 percent increase. This increase likely resulted from the finding that California has a higher share of farm labor contractors than other regions. The closer a binary variable is to 50 percent, the larger is its variance. Another large increase occurred with the optimal allocation for the number of farm employers in the Pacific Northwest. The optimal allocation called for increasing the interview allocation in the fall cycle by 95 percent (an additional 84 workers) and by 45 percent for the spring and summer cycles.

Tables 1–7 below present the results of the analysis. Table 1 summarizes the analysis. For each cycle/region stratum, it shows the average optimal allocation, the Neyman allocation, the current allocation, and the difference in the number of farm workers to be interviewed if the optimal or Neyman allocations were used. The current allocation is based on a planned sample size of 2,458. The optimal allocations sum to a slightly smaller number due to rounding the allocations to whole workers. Tables 2–4 show the 10 optimal allocations for each of the 12 regions in the fall, spring, and summer cycles respectively. Tables 5–7 show the Neyman allocations for each of the 12 regions in the fall, spring, and summer cycles respectively.

Table 1. Difference Between Current and Optimal or Neyman Allocation.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Cycle** | **Region** | **Optimal allocation** | **Neyman allocation** | **Current allocation** | **Difference if optimal allocation is used** | **Difference if Neyman allocation is used** |
| Fall | AP | 37 | 46 | 49 | -12 | -3 |
|  | CA | 394 | 368 | 350 | 44 | 18 |
|  | CBNP | 24 | 60 | 75 | -51 | -15 |
|  | DLSE | 36 | 45 | 47 | -11 | -2 |
|  | FL | 56 | 56 | 53 | 3 | 3 |
|  | LK | 25 | 33 | 36 | -11 | -3 |
|  | MT12 | 21 | 25 | 27 | -6 | -2 |
|  | MT3 | 29 | 28 | 29 | 0 | -1 |
|  | NE1 | 23 | 23 | 29 | -6 | -6 |
|  | NE2 | 12 | 20 | 24 | -12 | -4 |
|  | PC | 118 | 99 | 88 | 30 | 11 |
|  | SP | 40 | 34 | 42 | -2 | -8 |
|  |  |  |  |  |  |  |
| Spring | AP | 22 | 28 | 29 | -7 | -1 |
|  | CA | 399 | 335 | 317 | 82 | 18 |
|  | CBNP | 22 | 37 | 45 | -23 | -8 |
|  | DLSE | 35 | 37 | 40 | -5 | -3 |
|  | FL | 68 | 68 | 64 | 4 | 4 |
|  | LK | 15 | 18 | 19 | -4 | -1 |
|  | MT12 | 15 | 18 | 18 | -3 | 0 |
|  | MT3 | 24 | 25 | 25 | -1 | 0 |
|  | NE1 | 10 | 12 | 15 | -5 | -3 |
|  | NE2 | 17 | 18 | 22 | -5 | -4 |
|  | PC | 75 | 67 | 60 | 15 | 7 |
|  | SP | 19 | 24 | 29 | -10 | -5 |
|  |  |  |  |  |  |  |
| Summer | AP | 42 | 50 | 53 | -11 | -3 |
|  | CA | 420 | 406 | 384 | 36 | 22 |
|  | CBNP | 50 | 67 | 83 | -33 | -16 |
|  | DLSE | 44 | 47 | 49 | -5 | -2 |
|  | FL | 35 | 47 | 44 | -9 | 3 |
|  | LK | 27 | 34 | 36 | -9 | -2 |
|  | MT12 | 35 | 29 | 30 | 5 | -1 |
|  | MT3 | 24 | 20 | 21 | 3 | -1 |
|  | NE1 | 23 | 21 | 27 | -4 | -6 |
|  | NE2 | 28 | 27 | 32 | -4 | -5 |
|  | PC | 168 | 149 | 131 | 37 | 18 |
|  | SP | 20 | 30 | 36 | -16 | -6 |

Table 2. Optimal Allocations for the Fall Cycle.

|  |  |  |
| --- | --- | --- |
| **Region** | **Optimal allocation** | **Current Allocation** |
|  | **Wage** | **Employer** | **Unauthorized** | **Number of Farm Employers (cont.)** | **Number of farm employers (binary)** | **Paid by hour** | **Number of farm work days** | **Number of children (cont.)** | **Number of children (binary)** | **Average** |  |
| AP | 35 | 23 | 41 | 35 | 36 | 40 | 43 | 39 | 38 | 37 | 49 |
| CA | 354 | 567 | 392 | 369 | 398 | 348 | 356 | 377 | 381 | 394 | 350 |
| CBNP | 38 | 9 | 24 | 9 | 16 | 30 | 32 | 29 | 27 | 24 | 75 |
| DLSE | 36 | 29 | 39 | 28 | 35 | 40 | 40 | 40 | 40 | 36 | 47 |
| FL | 42 | 32 | 53 | 79 | 63 | 60 | 51 | 59 | 65 | 56 | 53 |
| LK | 31 | 8 | 24 | 30 | 26 | 29 | 27 | 24 | 22 | 25 | 36 |
| MT12 | 27 | 6 | 22 | 18 | 18 | 30 | 25 | 22 | 21 | 21 | 27 |
| MT3 | 24 | 45 | 24 | 26 | 31 | 29 | 31 | 26 | 22 | 29 | 29 |
| NE1 | 39 | 9 | 25 | 14 | 22 | 33 | 33 | 16 | 14 | 23 | 29 |
| NE2 | 16 | 4 | 15 | 11 | 12 | 12 | 15 | 13 | 11 | 12 | 24 |
| PC | 122 | 61 | 110 | 172 | 120 | 114 | 118 | 119 | 124 | 118 | 88 |
| SP | 48 | 13 | 45 | 23 | 39 | 54 | 44 | 47 | 46 | 40 | 42 |

Table 3. Optimal Allocations for the Spring Cycle.

|  |  |  |
| --- | --- | --- |
| **Region** | **Optimal allocation** | **Current Allocation** |
|  | **Wage** | **Employer** | **Unauthorized** | **Number of Farm Employers (cont.)** | **Number of farm employers (binary)** | **Paid by hour** | **Number of farm work days** | **Number of children (cont.)** | **Number of children (binary)** | **Average** |  |
| AP | 21 | 14 | 24 | 21 | 21 | 24 | 25 | 23 | 23 | 22 | 29 |
| CA | 359 | 574 | 397 | 374 | 404 | 352 | 361 | 382 | 386 | 399 | 317 |
| CBNP | 34 | 8 | 22 | 8 | 15 | 27 | 29 | 26 | 25 | 22 | 45 |
| DLSE | 34 | 28 | 38 | 27 | 33 | 38 | 38 | 39 | 39 | 35 | 40 |
| FL | 51 | 39 | 65 | 97 | 76 | 73 | 63 | 71 | 79 | 68 | 64 |
| LK | 19 | 5 | 15 | 19 | 16 | 18 | 17 | 15 | 14 | 15 | 19 |
| MT12 | 20 | 4 | 16 | 13 | 13 | 21 | 17 | 16 | 15 | 15 | 18 |
| MT3 | 20 | 38 | 20 | 22 | 26 | 25 | 26 | 22 | 19 | 24 | 25 |
| NE1 | 17 | 4 | 11 | 6 | 10 | 14 | 14 | 7 | 6 | 10 | 15 |
| NE2 | 22 | 6 | 21 | 15 | 17 | 17 | 21 | 18 | 15 | 17 | 22 |
| PC | 77 | 39 | 70 | 109 | 76 | 72 | 75 | 76 | 79 | 75 | 60 |
| SP | 23 | 6 | 21 | 11 | 18 | 25 | 21 | 22 | 22 | 19 | 29 |

Table 4. Optimal Allocations for the Summer Cycle.

|  |  |  |
| --- | --- | --- |
| **Region** | **Optimal allocation** | **Current Allocation** |
|  | **Wage** | **Employer** | **Unauthorized** | **Number of Farm Employers (cont.)** | **Number of farm employers (binary)** | **Paid by hour** | **Number of farm work days** | **Number of children (cont.)** | **Number of children (binary)** | **Average** |  |
| AP | 41 | 27 | 47 | 40 | 41 | 47 | 49 | 45 | 44 | 42 | 53 |
| CA | 378 | 605 | 418 | 394 | 425 | 371 | 380 | 403 | 407 | 420 | 384 |
| CBNP | 79 | 18 | 50 | 18 | 34 | 64 | 67 | 61 | 57 | 50 | 83 |
| DLSE | 44 | 36 | 48 | 34 | 42 | 49 | 49 | 49 | 49 | 44 | 49 |
| FL | 26 | 20 | 33 | 49 | 39 | 37 | 32 | 36 | 40 | 35 | 44 |
| LK | 34 | 9 | 27 | 33 | 29 | 32 | 29 | 27 | 25 | 27 | 36 |
| MT12 | 45 | 10 | 36 | 30 | 30 | 49 | 40 | 37 | 35 | 35 | 30 |
| MT3 | 20 | 38 | 20 | 22 | 26 | 25 | 26 | 22 | 19 | 24 | 21 |
| NE1 | 39 | 9 | 24 | 14 | 22 | 33 | 32 | 16 | 14 | 23 | 27 |
| NE2 | 36 | 10 | 35 | 24 | 27 | 29 | 34 | 29 | 25 | 28 | 32 |
| PC | 174 | 87 | 156 | 244 | 171 | 162 | 169 | 170 | 177 | 168 | 131 |
| SP | 24 | 6 | 22 | 11 | 19 | 26 | 22 | 23 | 23 | 20 | 36 |

Table 5. Neyman Allocations for the Fall Cycle.

|  |  |  |
| --- | --- | --- |
| **Region** | **Neyman allocation** | **Current Allocation** |
|  | **Wage** | **Employer** | **Unauthorized** | **Number of Farm Employers (cont.)** | **Number of farm employers (binary)** | **Paid by hour** | **Number of farm work days** | **Number of children (cont.)** | **Number of children (binary)** | **Average** |  |
| AP | 44 | 31 | 51 | 45 | 46 | 50 | 53 | 49 | 48 | 46 | 49 |
| CA | 323 | 554 | 363 | 352 | 375 | 318 | 326 | 348 | 353 | 368 | 350 |
| CBNP | 93 | 23 | 60 | 22 | 42 | 75 | 80 | 73 | 69 | 60 | 75 |
| DLSE | 43 | 38 | 49 | 36 | 43 | 49 | 49 | 50 | 50 | 45 | 47 |
| FL | 41 | 34 | 53 | 81 | 63 | 59 | 50 | 58 | 64 | 56 | 53 |
| LK | 41 | 12 | 33 | 42 | 36 | 38 | 35 | 32 | 30 | 33 | 36 |
| MT12 | 32 | 8 | 26 | 22 | 22 | 35 | 29 | 27 | 26 | 25 | 27 |
| MT3 | 23 | 46 | 23 | 26 | 31 | 28 | 30 | 26 | 22 | 28 | 29 |
| NE1 | 38 | 10 | 25 | 15 | 23 | 32 | 32 | 16 | 14 | 23 | 29 |
| NE2 | 25 | 8 | 25 | 18 | 20 | 20 | 24 | 21 | 18 | 20 | 24 |
| PC | 100 | 54 | 92 | 148 | 102 | 94 | 98 | 99 | 103 | 99 | 88 |
| SP | 40 | 11 | 38 | 20 | 33 | 45 | 37 | 40 | 39 | 34 | 42 |

Table 6. Neyman Allocations for the Spring Cycle.

|  |  |  |
| --- | --- | --- |
| **Region** | **Neyman allocation** | **Current Allocation** |
|  | **Wage** | **Employer** | **Unauthorized** | **Number of Farm Employers (cont.)** | **Number of farm employers (binary)** | **Paid by hour** | **Number of farm work days** | **Number of children (cont.)** | **Number of children (binary)** | **Average** |  |
| AP | 26 | 19 | 31 | 27 | 28 | 30 | 32 | 29 | 29 | 28 | 29 |
| CA | 294 | 504 | 331 | 321 | 341 | 290 | 297 | 317 | 321 | 335 | 317 |
| CBNP | 57 | 14 | 37 | 14 | 26 | 46 | 49 | 45 | 42 | 37 | 45 |
| DLSE | 36 | 31 | 40 | 30 | 36 | 40 | 40 | 41 | 41 | 37 | 40 |
| FL | 50 | 41 | 64 | 98 | 76 | 71 | 61 | 70 | 78 | 68 | 64 |
| LK | 22 | 6 | 18 | 23 | 19 | 21 | 19 | 17 | 16 | 18 | 19 |
| MT12 | 23 | 6 | 19 | 16 | 16 | 25 | 21 | 19 | 18 | 18 | 18 |
| MT3 | 20 | 41 | 20 | 23 | 27 | 25 | 26 | 23 | 19 | 25 | 25 |
| NE1 | 20 | 5 | 13 | 8 | 12 | 17 | 17 | 8 | 7 | 12 | 15 |
| NE2 | 23 | 7 | 22 | 16 | 18 | 18 | 22 | 19 | 16 | 18 | 22 |
| PC | 68 | 37 | 63 | 101 | 69 | 64 | 67 | 68 | 71 | 67 | 60 |
| SP | 29 | 8 | 27 | 15 | 24 | 32 | 26 | 28 | 28 | 24 | 29 |

Table 7. Neyman Allocations for the Summer Cycle.

|  |  |  |
| --- | --- | --- |
| **Region** | **Neyman allocation** | **Current Allocation** |
|  | **Wage** | **Employer** | **Unauthorized** | **Number of Farm Employers (cont.)** | **Number of farm employers (binary)** | **Paid by hour** | **Number of farm work days** | **Number of children (cont.)** | **Number of children (binary)** | **Average** |  |
| AP | 47 | 33 | 55 | 48 | 49 | 54 | 57 | 53 | 52 | 50 | 53 |
| CA | 356 | 611 | 401 | 388 | 413 | 351 | 359 | 384 | 389 | 406 | 384 |
| CBNP | 104 | 26 | 67 | 25 | 47 | 84 | 89 | 82 | 77 | 67 | 83 |
| DLSE | 45 | 40 | 51 | 37 | 45 | 50 | 50 | 52 | 52 | 47 | 49 |
| FL | 34 | 28 | 44 | 67 | 53 | 49 | 42 | 48 | 53 | 47 | 44 |
| LK | 42 | 12 | 34 | 43 | 37 | 39 | 36 | 33 | 31 | 34 | 36 |
| MT12 | 38 | 9 | 31 | 26 | 26 | 41 | 34 | 31 | 30 | 29 | 30 |
| MT3 | 16 | 33 | 17 | 19 | 22 | 20 | 21 | 18 | 16 | 20 | 21 |
| NE1 | 36 | 9 | 23 | 14 | 21 | 30 | 30 | 15 | 13 | 21 | 27 |
| NE2 | 34 | 10 | 34 | 24 | 27 | 27 | 33 | 28 | 24 | 27 | 32 |
| PC | 150 | 81 | 138 | 222 | 153 | 141 | 147 | 149 | 156 | 149 | 131 |
| SP | 36 | 10 | 34 | 18 | 30 | 40 | 33 | 35 | 35 | 30 | 36 |

1. Office of Management and Budget (2016). Standards and Guidelines for Statistical Surveys. Retrieved from <https://unstats.un.org/unsd/dnss/docs-nqaf/USA_standards_stat_surveys.pdf> [↑](#footnote-ref-1)
2. Montgomery, D. (Eighth ed., 2013). *Design and Analysis of Experiments.* New York, NY: Wiley. [↑](#footnote-ref-2)