Preliminary Findings for the April 2018 Agricultural Labor Base Wage Question Experiment

Benjamin M. Reist
Deputy Director
Research and Development Division

Tyler Wilson
Survey Methodologist
Survey Methodology and Technology Section
Research and Development Division

Shane T. Ball
Section Head
Survey Methodology and Technology Section
Research and Development Division

Linda J. Young
Chief Mathematical Statistician and Director
Research and Development Division

Author email: Benjamin.Reist@nass.usda.gov

1. INTRODUCTION

The U.S. Department of Agriculture's (USDA's) National Agricultural Statistics Service (NASS) has initiated a study to evaluate the impact that proposed changes in the questionnaire for the Agricultural Labor Survey will have on data quality and usability. Randomized experiments were conducted during the April and October 2018 administrations of the Agricultural Labor Survey to evaluate two survey versions of the questionnaire: 1) the original survey with a measure for gross wages and 2) a modified version with measures on gross wages, base wages, and incentive/overtime wages. The two questionnaires are to be compared using the following measures:

- Unit response rates
- Sample size to meet NASS's coefficient of variation (CV) targets
- Item nonresponse rates on gross wages
- Item nonresponse rates on base wages
- Three forms of a weighted item nonresponse rates on base wages:
    - Using the sampling weight
    - Using the nonresponse adjusted weight
    - The imputation rate

Meaningful difference/levels for each of the metrics above were determined before the fielding of the experiment. The following are the meaningful difference/levels of each of the measures:

- Unit response rates reduction of 10 percent or more
- Sample size increase of 10 percent or more
- Item nonresponse rate on gross wages of 10 percent or more
- Item nonresponse rate on base wages of 40 percent or more
- Weighted item nonresponse rates on base wages of 25 percent or more

This is a preliminary report of some of the results of the experiment in the April 2018 administration of the Agricultural Labor Survey. In Section 2, general information about the purpose, sample design and data collection methods of the Agricultural Labor Survey is provided. The design of the experiment is described in Section 3. The methodology used to analyze the data from the experiment is discussed in Section 4. The preliminary findings are presented in Section 5. The final section (6) concludes the report with overall findings, limitations and implications for the 2018 October experiment. A final report will be produced after the results from the October 2018 survey are published.

2. AGRICULTURAL LABOR SURVEY

NASS conducts the Agricultural Labor Survey twice annually. The purpose of the survey is to determine the types and number of farm workers employed and the wages they are paid.

Farm employment and wage statistics are used by federal, state and local government agencies, farm organizations, and employers for many purposes, including planning, recruitment and

placement of workers, and policymaking. The agricultural wage rate is a component of the Parity Index and is used in the establishment of minimum wage rates for domestic and foreign agricultural workers.

The survey results are published in the semi-annual Farm Labor release, issued around the 21st of May and November. Both semi-annual releases include regional and U.S. level estimates for the number of hired workers, hours worked weekly and wage rates by type of worker. The November release also includes annual average estimates at regional and U.S. levels.

2.1. Sample Design

The target population for the Agricultural Labor Survey includes all U.S. farms, where a farm is any place that produced and sold or would normally produce and sell at least $1,000 worth of agricultural products during the year.

The Labor Survey is a multiple frame survey, drawing from both the NASS List Frame and the June Area Frame. The list is an efficient sampling frame because it contains most of the farms with hired labor. The area frame provides the completeness missing from the list.

The Labor Survey has a stratified random design. Strata within states are based on peak number of hired workers, presence of intensive commodities and farm value of sales. The sample is drawn using simple random sampling within state-stratum combinations.

A new sample is drawn prior to the October data collection. The sample is split into eight replicates, each of which is representative of the population. Six of the eight replicates are randomly selected to become the sample for the October data collection. The two replicates that are not in the October data collection and four of the six replicates from the October data collection comprise the sample for the following April data collection. Approximately, 14,000 operations are in the sample for each round of data collection.

NASS targets coefficients of variation (CV) of the estimated quarterly gross wage rates of 1% at the national level and of 2% at the agricultural labor region level. The sample is allocated to meet or exceed these targets.

2.2. Data Collection

A self-administered paper questionnaire and computer assisted telephone interviews (CATI) are the primary data collection methods for the Agricultural Labor Survey, although survey respondents may report using the NASS Agcounts website. A limited number of field interviews are also conducted. All states use the same version of the Agricultural Labor questionnaire except California. Data for two quarters are collected during each administration of the survey. Information from January and April is gathered during the April data collection; information from July and October is reported during the October data collection.

The survey reference weeks are the weeks containing the 12th day of the survey month (January, April, July, and October). Questionnaires are mailed the first day of the Reference Week in April

and October. Phone and field enumerations begin the Sunday after the reference week, and continue for approximately two and a half weeks.

2.2.1. California

NASS and the California Employment Development Department (EDD) have a joint cooperative agreement to conduct a monthly California Agricultural Labor Survey by mail. In the months that the national survey is conducted (October and April), the NASS sample is used in both the NASS and EDD preliminary estimates, while the EDD supplemental sample is used only to enhance EDD's regional estimates. In the other months, California's entire list sample is used only by EDD.

In addition to the definition of the target population described above (section 2.1), California also collects data from Agricultural Service firms. California maintains a separate list frame of Agricultural Service operations. The California State EDD prints and mails a questionnaire to a sample from this list each month. During the October and April data collections, NASS conducts phone follow-up for nonrespondents in the portion of the California sample that is used for the Agricultural Labor Survey Estimates.

3. EXPERIMANTAL DESIGN

To evaluate the effects of modifying the survey instruments to include questions about base wages and incentive/overtime wages, NASS conducted a randomized experiment evaluating two versions of the April 2018 administration of the Agricultural Labor Survey:

Version 1: The original questionnaire with questions about gross wages.

Version 2: A modified questionnaire with questions designed to collect information on gross wages, base wages and incentive/overtime wages.

Version 1 is referred to as the control; version 2 is the treatment. For the study, in addition to the six replicates that are typically used in April, one of the two remaining replicates from the October administration was included, resulting in the April 2018 sample being comprised of seven replicates. The additional replicate from the October sample and a random selection of one of the original six replicates for the April sample were assigned to version 2; the remaining five April replicates were assigned to version 1.

Data for both versions were collected simultaneously using the same data collection modes and strategies. Phone follow-up for version 2 was conducted by the Oklahoma Data Collection Center (DCC), whereas the phone follow-up for version 1 was conducted by the other five DCCs. Since version 2 of the California questionnaire was not available in time for the January administration of the Agricultural Labor Survey in California, version 2 was only collected for April in California. Additionally, the Wyoming DCC conducted phone follow-up for the California sample for both versions.

4. METHODOLOGY

The experiment was evaluated using the following criteria mentioned in the Introduction:

- Unit response rates
- Sample size to meet NASS's coefficient of variation (CV) targets
- Item nonresponse rates on gross wages
- Item nonresponse rates on base wages
- Three forms of a weighted item nonresponse rates on base wages:
  - Using the sampling weight
  - Using the nonresponse adjusted weight
  - The imputation rate

Meaningful difference/levels for each of the metrics above were determined before the fielding of the experiment as follows:

- Unit response rates reduction of 10 percent or more
- Sample size increase of 10 percent or more
- Unweighted item nonresponse rate on gross wages of 10 percent or more
- Unweighted item nonresponse rate on base wages of 40 percent or more
- Weighted item nonresponse rates on base wages of 25 percent or more

Since the differences in unit response rates and item nonresponse rates on gross wages are comparisons between the two versions of the questionnaire, logistic regression was performed to assess whether the differences in the rates between the two versions were statistically significant. The other measure only pertains to version 2, so no statistical testing can be performed.

4.1. Unit Nonresponse

For unit nonresponse rates, three logistic models were used to evaluate the statistical significance of the differences between responses at the regional and national levels.

Separate analyses were conducted for January and April. Because the base wage question was not asked in California, the January model did not include any cases from California. In April, two models were used, one that included the cases from California and one that excluded the cases from California. This second model was used since the California data collection is substantively different from the rest of the United States as discussed in section 2.2.1. All three of the models used version, region and version crossed with region as the predictors.

4.2. Item Nonresponse Rates on Gross Wages

The models of item nonresponse rates for gross wages are similar to the models presented in section 4.1 with the following differences: (1) The response indicator was an indicator of an unusable item. (2) Each operation can report multiple worker classes. Because gross wages are collected for each worker class, an observation is recorded for each operation and reported

worker class combination. (3) Finally, since there can be multiple reports of gross wages by an operation, a random effect could be used to account for any interoperation correlation. However, almost 50 percent of operations only report one worker class, making the random effect for operation non-estimable. Not using this random effect leads to an increase in the variances, resulting in more conservative tests.

4.3.Additional Metrics

The other three metrics are calculated and compared to the levels specified above.

5.   FINDINGS

This section reports the preliminary findings of the April 2018 Labor Survey Experiment and is organized around the five metrics laid out in the introduction.

- Unit response rates
- Sample size to meet NASS's coefficient of variation (CV) targets
- Item nonresponse rates on gross wages
- Item nonresponse rates on base wages
- Weighted item nonresponse rates on base wages

5.1.Unit Response Rates

The table below shows the *p*-values for the nonresponse models described in section 4.1.

Nonresponse Models			
Predictor	January	April	
	U.S. without California	U.S. without California	All of the U.S.
Version	0.4496	0.5069	0.0463
Region	<0.0001	<0.0001	<0.0001
Version*Region	0.4409	0.4023	<0.0001

The *p*-values in red signify the effects that are statistically significant at the 5 percent level. The results of the U.S. response rate analysis without California are similar for both January and April. The lack of a significant interaction between version and region indicates that the response rate differences in versions do not vary significantly across regions. Although there is a significant difference in response rates among regions, the response rates between the two

versions is not significant (see Table A1 for more detail). In April, when California is included in the model version, region and the interaction between version and region are statistically significant at the 5 percent level. The 23.9 percent reduction in the California unit response rate for the treatment compared to the control is the only regional difference between versions that is significant at the 5 percent level ($p = <0.0001$) (see Table A2). It is also the only reduction in regional level response that is above the meaningful level of a reduction of 10 percent or more. The decrease in California led to a statistically significant decrease in the U.S. unit response rate although the observed decrease of 1 percent was well below the 10 percent reduction considered meaningful. In summary, the only region with an observed decrease in unit nonresponse exceeding 10 percent from version 1 to version 2 is California, which occurred in April.

5.2. Sample Size Increase

The sample size changes needed to meet CV targets for total gross wages (both regional and national) based on the version 2 response rates, are displayed in Tables A3 in the appendix. To meet the sample size for the national sample, the total size would need to be increased nationally by 9 percent based on January estimated CVs and 8 percent based on April estimated CVs. Most sample increases fall below the 10 percent meaningful increase.

The Southern Plains and Appalachian I regions are the only regions that need an increase in sample in either April or January. Four regions in January and two regions in April could reduce their sample sizes and still meet the CV targets.

5.3. Item Nonresponse Rates for Gross Wages

The *p*-values for the item nonresponse models described in section 4.3 are displayed in the table below.

	Gross-wages Unusable Reports Models		
Predictor	January	April	
	U.S. without California	U.S. without California	All of the U.S.
Version	0.9742	0.2751	0.9698
Region	<0.0001	0.0004	0.0006
Version*Region	0.1832	0.0001	0.0002

The *p*-values in red signify the effects that are statistically significant at the 5 percent level. For January, neither the interaction between version and region nor the main effect of version is significant. However, the Delta region had a 10.6 percent increase in item nonresponse, which

was significant at the 5 percent level ($p = 0.0085$). Since this is an increase larger than 10 percent, this increase is considered meaningful (see Table A4 in the appendix).

For April, whether considering the United States with or without California, there is a significant interaction between version and region. Three regions had a significant increase in item nonresponse for gross wages at the 5 percent confidence level, the Delta region at 8.9 percent ($p = 0.0166$), the Pacific region at 13.8 percent ($p < 0.0001$) and the Southern Plains region at 15.6 percent ($p = 0.0163$). The increases in the Pacific and Southern Plains regions are considered meaningful because they are larger than 10 percent (see Table A4 in the appendix).

5.4. Base Wage Item Nonresponse

For measures of base wage, the following item nonresponse rates were considered:

1. Unweighted item nonresponse rates
2. Weighted item nonresponse rates using sampling weights
3. Weighted item nonresponse rates using nonresponse adjusted weights
4. Imputation rate (i.e. the proportion of the estimate of total base wages imputed)

The definition of each of these rates is given in Table A8 in the appendix. One issue with using traditional weighted item nonresponse rates (*i.e.*, 2 and 3) for populations with a highly skewed outcome variables is that the sampling weight is usually inversely related to the contribution record of the estimate. This is because to reduce sampling error, units that have a higher likelihood of having large values, in this case wages, are sampled at a higher rate than units that do not. A traditional weighted item nonresponse rate gives more weight to missing values with higher weights that in general have lower base wages. For more information about this issue in business and establishment surveys and some alternative measures, see Thompson and Oliver (2012).

The U.S. Census Bureau uses an item nonresponse rate for business and establishment surveys as an alternative to a weighted item nonresponse rate (see appendix D3-B of the U.S. Census Bureau Statistical Quality Standards, 2013). This rate is the proportion of the estimate that is based on imputed values. The imputation rate takes into consideration both weight and the amount of base wages either reported or imputed for the operation.

At the national level, none of the four measures show a meaningful level of item nonresponse for base wages. In April, only the imputation rate rises to a meaningful level for both the United States with or without California at 28.4 percent and 26.6 percent respectively.

The following table shows the percentage of the regions that show a meaningful level of item nonresponse in base wages (base wages pertains only to version 2 so no statistical testing was performed).

	Item Nonresponse Rate			
Month	Unweighted	Sampling Weight	Nonresponse Weight	Imputation Rate
January	3 (17.6%)	4 (23.5%)	4 (23.5%)	2 (11.8%)
April	3 (16.7%)	8 (44.4%)	7 (38.9%)	4 (22.2%)

The four measures show that item nonresponse to base wages was a bigger issue in April compared to January (see Tables A6 and A7). Additionally, in both January and April, fewer regions were identified by the imputation rate when compared to either of the weighted measures. This implies that item nonresponse is more of an issue for operations with the larger weights and lower levels of base wages. Finally, there are only two instances where all four measures rise to a meaningful level, the California and Pacific regions in April. These regions had April imputation rates of 59.2 percent and 58.2 percent, respectively.

6. CONCLUSION

The current Agricultural Labor Survey collects gross wages. It has been proposed that the base wage, *e.g.* base pay, as well as incentives (including overtime pay), also be collected. To assess the potential impact of such a change, a new version of the questionnaire with questions to acquire the additional information was compared to the current version. In this report, statistical findings indicate very few differences in response rates between versions at the national level. In addition, some results in the amount of usable data collected for base wages indicate possible issues that will require further research. As is true with any experiment performed in an operational environment, there were limitations that need to be addressed.

6.1. Limitations

The treatment survey sample was administered in one call center, whereas the control survey sample was evenly distributed between the remaining four call centers. Thus, interviewer effects are confounded with version, which could explain some of the differences or lack of differences found in response rates and usable data.

Evidence from recent behavior coding research on interviewers during the January and April data collection processes indicates a consistent lack of adherence to reading the base wage questions. Some interviewers were found reverting to the original base wage wording and not reading the new questions. A final report on this research is pending.

California conducts the Agricultural Labor Survey differently than other regions or states. In California, the survey is administered monthly by the California Employment Development Division. Response is encouraged by mail or fax although some interviews are conducted by telephone from Wyoming and questions are read from a paper version. This led to two analyses

for the April survey: (1) analyzing all of the United States and (2) considering separate analyses for California and the rest of the United States.

In the upcoming October 2018 Agricultural Labor Survey Experiment, the first two limitations have been addressed. To allow for an evaluation of potential call-center effects, both versions of the questionnaire will be administered in each of the data collection centers. Additionally, based on the behavior coding and prior cognitive and usability testing, the CATI instrument has been redesigned to address some of the issues interviewers were having administering this survey over the phone and separating overtime wages and incentives. Finally, interviewer training was updated to provide clearer guidance on how to administer the questionnaire.

6.2. Final Summary

The results of the April 2018 Agricultural Labor Survey Experiment on the inclusion of the new wage question for the United States, excluding California are:

1. Response Rates
   - No significant impact in response rates between survey versions
   - No significant impact in item nonresponse for gross wages at the national level although some differences are observed at the regional level
   - Significant and meaningful impacts in item nonresponse rates for gross wages for a few regions
   - At the national level, only one of the four measures of item nonresponse for base wages rose to a meaningful level in April and none for January.
   - The national samples would need to be increased to meet NASS CV targets, but this increase is less than 10 percent
   - Measures of item nonresponse show a consistent meaningful level of item nonresponse for base wage in the Pacific region.

Even though the new version of the questionnaire satisfied the criteria for implementation for all but one of the national measures, there is evidence that there could be some issues between versions at the regional level when reporting base wages. This is particularly true in California and the Pacific region. Regional level analysis will need to be a focus of the 2018 October Agricultural Labor Base Wage Experiment. Additionally, more research needs to be done to better understand why the new version of the questionnaire performed so poorly in California.

References

Thompson, K. J. and Oliver, B. E. (2012). Response Rates in Business Surveys: Going Beyond the Usual Performance Measures. *Journal of Official Statistics*, 28, 221-237.

U.S. Census Bureau (2013) *U.S. Census Bureau Statistical Quality Standards*. Washington, DC. available at https://www.census.gov/content/dam/Census/about/about-the-bureau/policies_and_notices/quality/statistical-quality-standards/Quality_Standards.pdf (accessed 9/27/2018)

Appendix

Table A1: January Overall Unit Response Rate Treatment versus Control by Region

	Sample Size		Response Rate			
Region	**Control**	**Treatment**	**Control**	**Treatment**	**Difference**	**P-value**
Florida	419	152	53.0%	55.9%	2.9%	0.9220
Hawaii	351	124	47.3%	46.0%	-1.3%	0.9450
Northeast I	591	209	46.7%	50.2%	3.5%	0.7808
Northeast II	445	152	52.8%	52.0%	-0.8%	0.8991
Appalachian I	546	184	53.8%	54.3%	0.5%	0.9439
Appalachian II	633	220	58.9%	61.8%	2.9%	0.8413
Southeast	553	175	58.2%	65.7%	7.5%	0.1398
Lake	696	233	56.6%	51.9%	-4.7%	0.0845
Cornbelt I	753	260	52.6%	58.1%	5.5%	0.1027
Cornbelt II	639	231	52.6%	54.5%	2.0%	0.8762
Delta	669	198	66.8%	63.1%	-3.7%	0.2883
Northern Plains	895	288	49.1%	50.3%	1.3%	0.6216
Southern Plains	913	269	50.7%	51.3%	0.6%	0.9900
Mountain I	347	122	48.1%	55.7%	7.6%	0.4334
Mountain II	296	110	57.4%	49.1%	-8.3%	0.0680
Mountain III	285	102	53.0%	45.1%	-7.9%	0.1293
Pacific	536	184	45.9%	44.6%	-1.3%	0.4097
U.S. without CA	9,567	3,213	53.3%	53.9%	0.7%	0.4496

Table A2: April Overall Unit Response Rate Treatment versus Control by Region

Region	Sample Size		Response Rate			
	Control	Treatment	Control	Treatment	Difference	P-value
California	706	259	66.0%	42.1%	-23.9%	<.0001
Florida	419	152	53.7%	56.6%	2.9%	0.9434
Hawaii	351	124	47.6%	46.0%	-1.6%	0.7405
Northeast I	591	209	46.7%	50.2%	3.5%	0.9220
Northeast II	445	152	52.8%	53.3%	0.5%	0.6981
Appalachian I	546	184	53.7%	54.9%	1.2%	0.8134
Appalachian II	633	220	59.9%	61.4%	1.5%	0.8816
Southeast	553	175	58.0%	65.7%	7.7%	0.1295
Lake	696	233	56.6%	51.9%	-4.7%	0.0845
Cornbelt I	753	260	52.9%	58.8%	6.0%	0.0769
Cornbelt II	639	231	52.9%	54.1%	1.2%	0.9746
Delta	669	198	67.1%	63.1%	-4.0%	0.2545
Northern Plains	895	288	49.1%	50.7%	1.6%	0.5531
Southern Plains	913	269	51.3%	51.3%	0.0%	0.8914
Mountain I	347	122	47.8%	55.7%	7.9%	0.4044
Mountain II	296	110	57.4%	49.1%	-8.3%	0.0680
Mountain III	285	102	52.6%	45.1%	-7.5%	0.1446
Pacific	536	184	45.0%	45.1%	0.1%	0.6198
U.S. without CA	9,567	3,213	53.4%	54.1%	0.7%	0.5069
United States	10,273	3,472	54.3%	53.2%	-1.0%	0.0463

Table A3: Percentage Change in Sample Sizes Needed to Meet CV Targets

Region	January	April
	% Difference	
California	N/A	0
Florida	0	0
Hawaii	-56	-20
Northeast I	0	0
Northeast II	0	0
Appalachian I	87	0
Appalachian II	0	0
Southeast	0	0
Lake	0	0
Cornbelt I	-35	0
Cornbelt II	0	0
Delta	0	0
Northern Plains	0	0
Southern Plains	230	166
Mountain I	0	0
Mountain II	-35	0
Mountain III	-57	-60
Pacific	0	0
United States	9	8

Table A4: Item Nonresponse for Gross Wages January 2018

Region	Total Rows		Percent of Rows with Missing Gross Wages			
	Control	Treatment	Control	Treatment	Difference	P-Value
Florida	224	68	17.0%	13.2%	-3.7%	0.4649
Hawaii	222	29	15.3%	0.0%	-15.3%	0.9752
Northeast I	266	52	6.0%	13.5%	7.4%	0.0650
Northeast II	185	42	7.0%	7.1%	0.1%	0.9789
Appalachian I	249	44	8.0%	6.8%	-1.2%	0.7828
Appalachian II	163	36	5.5%	5.6%	0.1%	0.9935
Southeast	206	62	16.0%	16.1%	0.1%	0.9836
Lake	231	79	3.9%	2.5%	-1.4%	0.5745
Cornbelt I	193	67	6.2%	1.5%	-4.7%	0.1602
Cornbelt II	121	35	7.4%	8.6%	1.1%	0.8248
Delta	251	73	7.2%	17.8%	10.6%	0.0085
Northern Plains	234	39	16.7%	12.8%	-3.8%	0.5467
Southern Plains	166	26	10.8%	19.2%	8.4%	0.2276
Mountain I	152	39	5.3%	2.6%	-2.7%	0.4875
Mountain II	120	18	10.0%	5.6%	-4.4%	0.7329
Mountain III	172	56	8.1%	5.4%	-2.8%	0.3219
Pacific	255	67	5.1%	14.9%	9.8%	0.1233
U.S. without CA	3,410	754	9.2%	9.4%	0.2%	0.9742

Table A5: Item Nonresponse for Gross Wages April 2018

Region	Total Rows		Percent of Rows with Missing Gross Wages			
	Control	Treatment	Control	Treatment	Difference	P-Value
California	1,048	266	0.4%	0.0%	-0.4%	0.9761
Florida	221	69	16.3%	7.2%	-9.1%	0.0674
Hawaii	223	30	12.1%	3.3%	-8.8%	0.1820
Northeast I	299	77	5.4%	11.7%	6.3%	0.0522
Northeast II	224	54	9.4%	3.7%	-5.7%	0.1908
Appalachian I	287	67	3.8%	7.5%	3.6%	0.2061
Appalachian II	180	55	5.6%	5.5%	-0.1%	0.9771
Southeast	236	74	12.7%	12.2%	-0.5%	0.9010
Lake	269	94	3.7%	1.1%	-2.7%	0.2260
Cornbelt I	227	76	8.8%	2.6%	-6.2%	0.0911
Cornbelt II	147	41	5.4%	7.3%	1.9%	0.6523
Delta	363	86	8.5%	17.4%	8.9%	0.0166
Northern Plains	254	42	15.7%	11.9%	-3.8%	0.5220
Southern Plains	176	29	8.5%	24.1%	15.6%	0.0163
Mountain I	174	45	5.7%	4.4%	-1.3%	0.7329
Mountain II	136	24	11.0%	4.2%	-6.9%	0.3219
Mountain III	171	57	10.5%	3.5%	-7.0%	0.1233
Pacific	330	90	3.9%	17.8%	13.8%	<.0001
U.S. without CA	3,917	1,010	8.5%	8.7%	0.2%	0.7897
United States	4,965	1,276	6.7%	6.9%	0.1%	0.8499

Table A6: Item Nonresponse for Base Wages January 2018

Region	Total Rows	Item Nonresponse Rates			
		Unweighted	Sampling Weight	Nonresponese Weight	Imputation Rate
Florida	69	20.30%	10.1%	12.1%	0.0%
Hawaii	29	20.70%	15.9%	15.1%	7.80%
Northeast I	52	19.20%	14.4%	11.8%	10.50%
Northeast II	41	41.50%	35.3%	37.3%	4.70%
Appalachian I	44	15.90%	5.2%	5.7%	5.10%
Appalachian II	36	11.10%	27.3%	24.3%	14.30%
Southeast	62	51.60%	32.3%	34.2%	0.0%
Lake	80	26.30%	22.0%	22.1%	2.60%
Cornbelt I	65	29.20%	14.5%	16.5%	18.90%
Cornbelt II	35	34.30%	31.7%	30.7%	12.70%
Delta	72	27.80%	21.7%	23.1%	0.0%
Northern Plains	39	33.30%	13.8%	13.7%	47.30%
Southern Plains	26	15.40%	23.0%	22.3%	9.00%
Mountain	39	30.80%	34.1%	40.9%	21.50%
Mountain II	18	5.60%	16.0%	11.5%	8.00%
Mountain III	55	3.60%	2.9%	2.5%	8.90%
Pacific	67	46.30%	16.2%	23.0%	44.80%
National (no CA)	829	27.10%	20.4%	20.7%	13.40%

Table A7: Item Nonresponse for Base Wages April 2018

Region	Total Rows	Item Nonresponse Rates			
		Unweighted	Sampling Weight	Nonresponese Weight	Imputation Rate
California	265	60.0%	60.4%	60.8%	59.2%
Florida	70	18.6%	8.9%	10.4%	0.0%
Hawaii	30	30.0%	40.7%	42.1%	13.7%
Northeast I	77	22.1%	13.1%	11.8%	13.5%
Northeast II	53	34.0%	31.9%	32.1%	4.3%
Appalachian I	67	22.4%	14.7%	14.9%	5.2%
Appalachian II	55	9.1%	25.8%	22.5%	14.2%
Southeast	74	48.6%	25.7%	27.1%	0.0%
Lake	95	24.2%	22.3%	21.9%	6.6%
Cornbelt I	74	21.6%	13.3%	15.0%	11.8%
Cornbelt II	41	29.3%	29.5%	30.2%	14.3%
Delta	86	25.6%	23.4%	24.5%	0.6%
Northern Plains	42	28.6%	11.9%	11.7%	45.1%
Southern Plains	29	17.2%	18.2%	18.0%	8.7%
Mountain	45	28.9%	44.4%	50.3%	29.8%
Mountain II	24	0.0%	0.0%	0.0%	7.3%
Mountain III	56	1.8%	0.6%	0.7%	3.8%
Pacific	90	56.7%	27.0%	33.3%	58.2%
U.S. without CA	1,008	26.6%	19.8%	20.3%	26.6%
United States	1,273	33.5%	23.1%	24.5%	28.4%

Table A8: Item Nonresponse Formulas

Type	Definition
Unweighted	$$\frac{\sum_{i \epsilon F} \sum_{j \epsilon R_i} u_{ij}}{r}$$
Sampling Weight	$$\frac{\sum_{i \epsilon F} \sum_{j \epsilon R_i} w_i^S u_{ij}}{\sum_{i \epsilon F} \sum_{j \epsilon R_i} w_i^S}$$
Nonresponse Adjusted Weight	$$\frac{\sum_{i \epsilon F} \sum_{j \epsilon R_i} w_i^A u_{ij}}{\sum_{i \epsilon F} \sum_{j \epsilon R_i} w_i^A}$$
Imputation Rate	$$\frac{\sum_{i \epsilon F} \sum_{j \epsilon R_i} w_i^A u_{ij} t_{ij}}{\sum_{i \epsilon F} \sum_{j \epsilon R_i} w_i^A t_{ij}}$$

F = the set of farms who responded

R_i = the set of rows for farm i

u_{ij} = the indicator that value of the given item is unusable in row j for farm i

r = the total number of rows across all responding farms

w_i^S = the sampling weight for farm i

w_i^A = the sampling weight for farm i

t_{ij} = the value of the given item given item in row j for farm i (Note: this could be imputed)