June, 2016

Non-Citrus Tree Fruit Nonresponse Bias Analysis

BACKGROUND

USDA NASS conducts two annual tree fruit surveys; the Fruit Non-citrus Tree Preliminary PDI (Production and Disposition Inquiry) and the Fruit Non-Citrus Tree Final PDI. The Preliminary PDI survey is conducted in November and the Final PDI is conducted in May. The survey of interest for this study is the May, 2016 Non-citrus Tree Fruit PDI.

The response rate for this survey was below 80 percent and the coverage was below 70 percent. A nonresponse bias analysis took place to compare our survey expansion indications to our "complete" expansion indications. This comparison was made to identify whether the survey indications were biased. Currently our sample design is Multivariate Probability Proportion to Size (MPPS) sampling, created from the different types of fruit from the sampling frame for each given record. The nonresponse adjustment is accomplished with a reweighted estimator. Each record is classified into four stratum, according to size. A nonresponse adjustment is calculated within each stratum and applied to the sampling weight.

PROXY DATA

To create a "complete" dataset, proxy data were used for the all nonresponse. The proxy data were implemented using a hierarchical structure, staring with the most recent data. First, proxy data were used from the November 2015 Tree Fruit PDI if available. Then three other recent tree fruit surveys were queried. If there were no data available at this point, sampling frame data were used from our internal database. This way, every record had usable data for each fruit crop. The primary source of the sampling frame data were from the 2012 Census of Agriculture.

Since the Census of Agriculture and the Tree Fruit PDI questions differed in structure, we adjusted some of the Tree Fruit PDI questions to reflect the census structure. For example, our annual survey separates 'Bartlett pears' and 'other pears', while the census only collected data for 'pears'. To make the comparison more on par with the census data, Bartlett and other pears were combined. Also, prunes and plums were combined as were Clingstone and Freestone peaches. The fruits of interest that are summarized are apples, apricots, sweet cherries, tart cherries, nectarines, pears, peaches, and prunes/plums. Each fruit crop has a set of states that are in the program. Apples have 29 states, apricots have 3, sweet cherries have 8, tart cherries have 7, nectarines have 2, pears have 12, peaches have 27, and prunes/plums have 5, totaling 93 indication comparisons for this study.

RESULTS

After re-running the summaries with the complete datasets, some bias was observed generally. The survey indication trend seemed to underestimate. Overall, 28 estimates were overestimated while 65 were underestimates, a 30% over and 70% under comparison. If the survey indications were unbiased, we would expect roughly the same amount of states above and below the complete indication. The proportion of indications underestimated is statistically different from 50 percent (alpha=.01, using binomial distribution).

Of the 29 apple states, ten were overestimated while 19 were underestimated. Of the three apricots and two nectarine states, all were underestimated. Of the eight sweet cherry states, three were overestimated and five were underestimated. Of the seven tart cherry states, three were overestimated while four were underestimated. Of the 12 pear states, four were overestimated and eight underestimated. Of the 27 peach states, six overestimated and 21 underestimated. For prunes/plums, two states were overestimated and three were underestimated.

Note that the “complete” dataset is not a perfect estimation of our population parameter, but it is the best obtainable comparison.

DISCUSSION

USDA NASS is currently attempting to increase response rate and coverage rates for the tree fruit surveys. Our phoning center is expanding the data collection period where possible. Our metadata team is collecting better contact information continually, including updating cell phones numbers and current home addresses. Our regional field offices are contacting our chronically inaccessible operators with new and different strategies. Our marketing department is giving more promotional information to the operators on why they should respond. Our sampling department is calibrating the maximum probability of selection (MPPS) sampling weights against sampling frame state totals. They are also creating non-response groups using the probability of selection as a measure of size. USDA NASS has also recently created a new team to look into improving response rates at the agency level.

USDA NASS research team will analyze different ways to achieve a better non-biased estimator for the Tree Fruit PDI. The research will include alternative ways to adjust for nonresponse.