Questions and Answers Concerning Coverage Adjustment Cells and CV's

(OMB had several questions asking for additional information on the methodology NASS would be using to calculate expansions, weights, CV's, or the handling of non-response questionnaires. The following text provides the questions and answers provided to OMB.)

Q. How are coverage-adjustment cells created and defined for the 2013 Census of Aquaculture?

A. The weighting cells will be most likely based on reported aquaculture sales obtained from the Census of Agriculture, but general sales data obtained from Census of Agriculture could be used as well. Size groups based on sales groups will be used- probably not to exceed 4 or 5 groups- i.e., group_1: sales below $50,000, group_2: sales between $50,000 and $200,000, group_3: sales between $200,000 and $500,000, group_4: sales above $500,000. Census coverage rates are in general strongly related to the economic size of the farm. Smaller farms have relatively low coverage rates and large farms have relatively high coverage rates.

Q. Can you show the formula for the target ratio used for the coverage adjustment?

A. $$w_g = \frac{\sum_i C_{ig}}{r_g}$$ where; w_g is the ratio for all respondents in adjustment cell, g. The numerator is the sum of the Census of Agriculture weights, C_i for each record (i) that falls into the cell (respondents and non-respondents alike). This number is referred to in the write-up as the "target". The denominator (r_g) is the number of responding records in weighting cell g.

Q. How is this ratio applied in order to "attain the coverage target for the cell"?

A. In the simplest case as shown above, the non-response and coverage adjustment is carried out simultaneously, and the weight for each responding in-scope aquaculture farm in weighting group g is simply w_g .

Q. Can some "Must" case records ever have a weight greater than 1?

A. Yes.

Q. What is the threshold definition of a “Must” record for this aquaculture survey?

A. All records for which Aquaculture sales data on the 2012 Census of Agriculture was greater than or equal to $150,000. This definition will be consistent across all states. [(K1345 + K1365 (Census of Agriculture, Section 17 aquaculture sales data)) >= $150,000]

Q. What were the CVs for the “major items of interest at the state level” from the 2005 survey cycle? Were they about 5%?

A. None were ever calculated.

Q. If so, is that the basis for your projection for this survey cycle? If not, can you explain the basis for projecting that CVs will be in the range of 5%?

A. CVs are projected based on results of the 2009 Horticulture census, which used a similar methodology. We expect a high response rate and a census is being taken. It does, of course, depend on the coverage and response rates on the Census of Agriculture. To insure coverage on the Census of Agriculture, possible aquaculture operations were flagged to not allow non-response. For some states CVs could be considerably higher, or lower, than 5%.

Q. If the reporting population is a census, are sample means and sample standard deviations calculated from this Census?

A. Yes, for some items, as yet to be determined.

Q. Can you show the formula used to calculate CVs?

A. Yes, see the short technical write-up that follows.

The estimator of an arbitrary total, t_y :

$$\hat{t}_y = \sum_g \left(\hat{N}_g / \sum_i a_{gi} \right) a_g \sum_i y_{gi} = \sum_g w_g \sum_i y_{ig}$$

Where $a_g = N_g / n_g$ which represents the nonresponse adjusted weight for each responding record *i* in adjustment group *g*.

$\hat{N}_g$ is the sum of the fully adjusted census weights for the records in group *g*. This includes both respondents and nonrespondents.

The variance estimator is:

$$\hat{V}\left(\hat{t}_y\right)=\sum_g a_g\left(a_g-1\right)\sum_i\left(q_{gi}e_{gi}\right)^2+\sum_g \hat{V}\left(\hat{N}_g\right)*\left[\left(\sum a_{gi}y_{gi}\right)/\left(\sum_i a_{gi}\right)\right]^2+\left[\sum_g\sum_i\left(w_{ig}^I-w_{ig}\right)y_{ig}\right]^2$$

With:

$w_{ig}^I=$ integerized fully adjusted weight for farm *i* in group *g.*

$w_{ig}=$ nonintegerized fully adjusted weight for farm *i* in group *g.*

$\hat{V}\left(\hat{N}_g\right)$ is guestimated based on the assumption that $CV\left(\hat{N}\right)$ does not exceed 7% for economically small farm adjustment groups and 1% for the economically large farm adjustment groups .

$q_{gi}=\hat{N}_g/\sum_{i\in g}a_{gi}$, the coverage adjustment.

$e_{gi}=y_{gi}-\left(\sum_i a_{gi}y_{gi}\right)/\sum_i a_{gi}$, the "error" from the group mean for farm *i* in group *g.*

The first set of terms gives the standard variance with the coverage adjustment accounted for. The second set of terms gives the estimated variance contributed by the coverage target (which is only an estimate of the total coverage farms.) The third set of terms gives the squared bias due to integerizing the weights.