Cooperative Project on Harvest Survey Sampling and Estimation: weighting adjustments for sampling, measurement-error, and non-response

Kingsford Jones
M.S. Candidate
Department of Mathematical Sciences
Montana State University
Pricipal Advisor: John Borkowski

July 20, 2010

1 Introduction

This report is a continuation of the work described in Jones and Borkowski (2007), and background for the cooperative agreement can be found in that paper and in Gude 2007. Detailed information on FWP's post-season hunter surveys is supplied in Gude et al. 2006. We refer the reader to those papers for more detailed background information than is given below.

1.1 Background

A primary goal of collecting the post-season phone survey data is to produce estimates of population parameters quantifying annual big game harvest in Montana. Parameters of interest include harvest within the following categories:

- Hunting District
- Age-Sex Class
- Antler Point Class
- Time Period
- Method (Bow/Rifle)

Ideally, the data produced by the phone surveys will allow for precise, unbiased estimation of each of these parameters at the state and regional levels, as well as within strata[1]. However, it is likely that biases exist in the estimates. As described in Gude 2007, two likely sources of bias in the current harvest estimators are measurement error and nonresponse. In the first report we derived estimators that incorporated adjustments for measurement error. In this report we include revisions of those methods, and we derive new methods that adjust for survey nonresponse.

[1]Each stratum is formed by combining residency-status and license-permit type of each hunter

1.2 Measurement Error

Measurement error in the big game harvest survey occurs when attributes of hunting and harvest activities recorded for a particular hunter are incorrect; such errors naturally result in biased parameter estimates. Major sources of measurement error bias include

- incorrect recall by the hunter (Chu et al. 1992)
- reporting of more "prestigious" categories of animals harvested (Atwood 1956, Wright 1978)
- data entry mistakes.

Data entry mistakes can be minimized with appropriate survey software and training programs. The other sources of measurement error are often confounded and difficult to remove from raw survey responses. To deal with this source of bias, correction factors can be built into estimators for population parameters (Geissler 1990).

1.3 Nonresponse

The problem of survey nonresponse bias is depicted clearly by a study commissioned by New Mexico Department of Fish and Game. Harrod and Lesser (2006) report that game managers with the NMDFG had long been concerned about the effect of nonresponse on annual elk harvest estimates derived from hunter questionnaires (the questionnaires are returned voluntarily, resulting in a self-selected sample that typically includes about 1 of every 3 hunters). The research, which involved contacting a sample of non-responders in each of two years, revealed that the nonresponse effect artificially inflated harvest estimates by about 30%. This was due to the fact that hunters who responded to the survey were far more likely to have harvested an elk than those who did not.

Fortunately, FWP survey protocols are designed to select a random sample of hunters to be included in post-season surveys and the effects of nonresponse should therefore be less drastic than they were in the New Mexico study. Nevertheless, there is generally a 30-40% nonresponse rate in the post-season phone survey (and this year may have as high as 60% nonresponse), and to the extent that those who respond to the phone calls are more or less likely to have harvested an animal than the population of hunters as a whole, the current harvest estimates will be biased. There are many plausible scenarios that could lead to this situation. For example, response rates may be dependent on harvest success due to a causal relationship (e.g., if successful hunters are more eager to talk about their hunting than unsuccessful hunters), or an indirect relationship (e.g., if non-residents were both harder to reach and more or less likely to have been successful than residents). Therefore, there is a clear need for nonresponse adjustments to be incorporated into the harvest estimators.

2 Methods

2.1 Estimator Derivations

2.1.1 Revised measurement error adjustments

In this section, we describe methods for adjusting survey counts for measurement bias. These are easiest to understand if we focus on just one variable of interest and organize the data in tables. We use elk age-sex class as an example, assuming calculations for other variables will follow similarly from this model.

Table 1 contains counts of all hunters who passed through check stations and are later phoned in the post-season harvest survey.

Table 1: Check station (row) and phone survey (column) elk sex-age class harvest frequencies for all hunters who passed through check stations. For example, A_{13} represents the number of hunters who harvested a cow according to check station records but claimed it was bull in the phone survey.

	Phone Survey				
Check Station	Cow	Calf	Bull	None	Row Total
Cow	A_{11}	A_{12}	A_{13}	A_{14}	$A_{1.}$
Calf	A_{21}	A_{22}	A_{23}	A_{24}	$A_{2.}$
Bull	A_{31}	A_{32}	A_{33}	A_{34}	$A_{3.}$
Column Total	$A_{.1}$	$A_{.2}$	$A_{.3}$	$A_{.4}$	$A_{..}$

Because all values in Table 1 are known we can use them to estimate bias due to differences between phone survey records and the truth (measurement bias). It is worth noting that an important row is missing from this table: hunters who harvested nothing but claimed they did harvest something. This row is not observable because a hunter who has nothing at the time of a check station visit may still harvest an animal at a later time.

Now consider only those hunters who did not pass through a check station but were contacted in the post-season phone survey. Table 2 summarizes the set of possible frequencies.

Table 2: True (row) and phone survey (column) harvest frequencies for hunters who did *not* pass through check stations but did provide responses to the phone survey. For example B_{21} represents the number of survey responders who actually harvested a calf but reported a cow. The only known values in this table are the column totals.

	Phone Survey				
Truth	Cow	Calf	Bull	None	Row Total
Cow	B_{11}	B_{12}	B_{13}	B_{14}	$B_{1.}$
Calf	B_{21}	B_{22}	B_{23}	B_{24}	$B_{2.}$
Bull	B_{31}	B_{32}	B_{33}	B_{34}	$B_{3.}$
None	B_{41}	B_{42}	B_{43}	B_{44}	$B_{4.}$
Column Total	$\mathbf{B}_{.1}$	$\mathbf{B}_{.2}$	$\mathbf{B}_{.3}$	$\mathbf{B}_{.4}$	$\mathbf{B}_{..}$

The values in this table will be important because the goal is to estimate measurement error for all hunters, not just those who pass through check stations. However, the only known values are the boldface values in the bottom row (i.e., the number of phone survey harvests reported in each age-sex class). Also note that unlike Table 1 this table does contain a row of nonzero frequencies for the hunters that did not harvest anything but claimed they did (the second to last row).

The two types of errors that may occur are random or prestige. A *random error* is an error that occurs due to data entry mistakes and incorrect recall by the hunter when participating in the survey. A *prestige error* is a nonrandom error that occurs when a hunter intentionally reports a more "prestigious" harvest category when giving a survey response. For example, when considering elk sex-age classes, the order of harvest prestige from lowest to highest is None → Calf → Cow → Bull.

Before we can estimate the two types of measurement error that may occur, we need to assert the following assumptions:

1. No errors occurred with the recorded classifications of hunters who passed through the check stations. That is, every harvest type reported at a check station is in the correct row in Table 1.

2. The phone survey reporting bias is approximately the same for hunters who passed through the check stations and responded and all hunters surveyed.

3. When a surveyed hunter reports a harvest type that is lower in prestige than the true harvest type, then it is purely a random error.

4. When a surveyed hunter reports a harvest type that is higher in prestige than the true harvest type, then it can be either a random error or a prestige error.

Given these assumptions, in Table 1 and Table 2:

- The diagonal cells A_{ii} for $i = 1, 2, 3$ in Table 1 and B_{ii} for $i = 1, 2, 3, 4$ in Table 2 represent correct matches between the true harvest and the harvest reported from the phone survey.

- The $A_{12}, A_{14}, A_{24}, A_{31}, A_{32}, A_{34}$ cells in Table 1 and the $B_{12}, B_{14}, B_{24}, B_{31}, B_{32}, B_{34}$ cells in Table 2 represent those cases when only random misclassification errors (but not prestige misclassification errors) have occurred.

- The A_{13}, A_{21}, A_{23} cells in Table 1 and the $B_{13}, B_{21}, B_{23}, B_{41}, B_{42}, B_{43}$ cells in Table 2 represent those cases when random or prestige misclassification errors have occurred.

Given the assumptions (1)-(4), we can estimate the random and prestige errors based on the frequencies observed in Table 1 and apply them to estimate the missing values in Table 2.

The estimate of the random misclassification error rate (denoted m_r) is the proportion of total survey responses that fall into cells associated with only random misclassification errors:

$$m_r = \frac{A_{12} + A_{14} + A_{24} + A_{31} + A_{32} + A_{34}}{A_{..}} \tag{1}$$

The estimate of the random plus prestige misclassification error rate (denoted m_{r+p}) is the proportion of survey responses that fall into cells associated with a random or prestige misclassification error for those harvest types (Cow and Calf) in Table 1 for which a prestige error could possibly occur:

$$m_{r+p} = \frac{A_{13} + A_{21} + A_{23}}{A_{1.} + A_{2.}} \tag{2}$$

Because a prestige error cannot occur in the *Bull* row and because there are all zero counts in the *None* row, we consider only the *Cow* and *Calf* row totals in the denominator $A_{1.} + A_{2.}$.

Note that for the *None* row in Table 2, $(1 - m_r)$ is an estimate of the proportion of the *None* phone call survey responses ($B_{.4}$) who, in truth, did not harvest an elk. Therefore, we can next estimate B_{44} by:

$$\widehat{B}_{44} = (1 - m_r)B_{.4} \tag{3}$$

$\widehat{B}_{44}$ is the estimated number of surveyed hunters that reported they did not harvest anything and actually did not. This is the estimated number of correct matches between the survey and the truth for no harvest.

The next relationship to consider is between $B_{4.}$ and B_{44}. Consider the Truth=*None* row in Table 1. Note that B_{41}, B_{42}, and B_{43} represent misclassification errors that are either random or prestige errors while B_{44} represents the *None* survey responses that were not subject to random or prestige errors. Therefore,

$$B_{44} = (1 - m_{r+p})B_{4.} \implies B_{4.} = \frac{B_{44}}{1 - m_{r+p}}. \tag{4}$$

Substituting $\widehat{B}_{44}$ from equation (3) into equation (4) provides our estimate of $B_{4.}$:

$$\widehat{B}_{4.} = \frac{\widehat{B}_{44}}{1 - m_{r+p}} = \frac{(1 - m_r)B_{.4}}{1 - m_{r+p}}. \tag{5}$$

Under the given assumptions and using the $\widehat{B}_{44}$ estimate, the estimated total number of incorrect matches of surveyed hunters who claim cow, calf, or bull when they did not harvest anything (the first three columns of the *None* row) is $m_{r+p}\widehat{B}_{4.}$. Given our assumptions, this estimate can then be partitioned into estimates for cow, calf, and bull based on proportionality of those harvest types in the survey. That is, for columns $j = 1, 2, 3$ (i.e., Cow, Calf, and Bull survey responses), the estimates of B_{4j} are:

$$\widehat{B}_{4j} = m_{r+p}\,\widehat{B}_{4.}\frac{B_{.j}}{B_{.1} + B_{.2} + B_{.3}} = \frac{m_{r+p}B_{4.}B_{.j}}{B_{..} - B_{.4}} \tag{6}$$

Now that we have estimates for the entries in the *None* row in Table 2, we can apply the combined check station and survey data summarized in Table 1. Given the assumption that the phone survey reporting bias is approximately the same for hunters who passed through the check stations and those who did not, for the i^{th} row and j^{th} column of Table 2 ($i = 1, 2, 3$ and $j = 1, 2, 3, 4$), we have the estimating equation

$$\frac{A_{ij}}{A_{.j}} = \frac{\widehat{B}_{ij}}{\sum_{i=1}^{3}\widehat{B}_{ij}} \implies \widehat{B}_{ij} = \left(\frac{A_{ij}}{A_{.j}}\right)\left(B_{.j} - \widehat{B}_{4j}\right). \tag{7}$$

By setting $j = 1$, we can calculate $\widehat{B}_{11}$, $\widehat{B}_{21}$, and $\widehat{B}_{31}$ to fill in all of the remaining values in the first column of Table 2 with estimates. By repeating the process described above for columns $j = 2, 3, 4$, we can fill in all remaining unknown values in Table 2 with estimates (as shown in Table 3). The method to fill in Table 3 with estimates will be applied in Section 2.1.3 to derive estimates of harvest adjusted for measurement-bias.

Table 3: Estimated true (row) and (column) harvest frequencies calculated from Equations (3) to (7).

	Phone Survey			
Truth	Cow	Calf	Bull	None
Cow	$\widehat{w}_{11}B_{.1}$	$\widehat{w}_{12}B_{.2}$	$\widehat{w}_{13}B_{.3}$	$\widehat{w}_{14}B_{.4}$
Calf	$\widehat{w}_{21}B_{.1}$	$\widehat{w}_{22}B_{.2}$	$\widehat{w}_{23}B_{.3}$	$\widehat{w}_{24}B_{.4}$
Bull	$\widehat{w}_{31}B_{.1}$	$\widehat{w}_{32}B_{.2}$	$\widehat{w}_{33}B_{.3}$	$\widehat{w}_{34}B_{.4}$
None	$\widehat{w}_{41}B_{.1}$	$\widehat{w}_{42}B_{.2}$	$\widehat{w}_{43}B_{.3}$	$\widehat{w}_{44}B_{.4}$
Column Total	$\mathbf{B}_{.1}$	$\mathbf{B}_{.2}$	$\mathbf{B}_{.3}$	$\mathbf{B}_{.4}$

We now combine phone survey data and estimates from hunters who did (Table 1) and did not (Table 2) go through check stations. It is convenient to represent these data as a table of proportions where the ij^{th} element ($i = 1, 2, 3, 4$ and $j = 1, 2, 3, 4$) is calculated as

$$\widehat{w}_{ij} = \frac{A_{ij} + \widehat{B}_{ij}}{A_{.j} + B_{.j}} \tag{8}$$

Now that we have estimates of the random error rate m_r and the random or prestige error rate m_{r+p}, we want to apply them to existing or future survey data to make adjustments for expected misclassifications. Unfortunately, it is not a simple matter of taking a weighted combination of the existing survey responses. The main obstacle is that the survey data do not represent a random sample from the population of hunters. There are two reasons:

(i) The initial sample taken was a stratified simple random sample. This is easily adjusted for because we know how the samples within each stratum were determined.

(ii) Within this initial sample, there will be a sizable number of non-responders to the survey when contacted. Therefore, we expect a non-response bias to exist. The method for adjusting for this nonresponse bias will be described in the next section.

2.1.2 Nonresponse bias adjustments to survey count data

Figure 1 depicts the hierarchical structure of the population of interest for the current or any future phone survey. It also introduces the following notation that will be used in this section:

- N_P is the number of units in the population of interest.
- N_S is the number of population units selected by the sampling scheme to be surveyed.
- N_E is the number of remaining population units that are excluded from the sample. Thus, $N_S + N_E = N_P$
- N_R is the number of respondents (or the number of actual responses received) from the N_S sampling units.
- N_M are the number of non-respondents (or the number of missing responses) from the N_S sampling units.

Because auxiliary information exists for the N_S hunters selected to be in the survey sample, we can develop a model for predicting whether or not a hunter would respond based on the auxiliary information. The goal is to take the auxiliary data collected from the N_R responders and draw valid inferences about all N_P units in the population. Potentially important auxiliary variables include:

- License Permit Type
- Geographic region
- Age
- Gender
- Residency status
- Years of residency

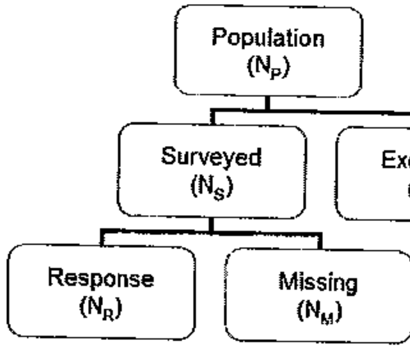

Figure 1: Hierarchical structure of the population of interest.

As mentioned earlier, the problem is that the units that respond are not likely to be a representative sample of the entire population and we therefore need estimators that account for this. Here we present the approach of developing and using a model for predicting whether or not a hunter will respond to the survey that will produce model-assisted estimators adjusted for nonresponse bias. For example:

- We could fit a logistic regression model with a binary response of 1 for a hunter who responded and a response of 0 for a hunter who did not respond. This binary response would be modeled using the hunter's auxiliary information.
- We could create a classification tree for classifying whether or not a hunter would respond based on the auxiliary information.

Once a model is developed from the N_S hunters selected by the sampling frame, we consider the predicted probabilities of responding for only those N_R hunters who actually did respond. These predicted probabilities will be converted to weights that, when applied, will provide estimated stratum totals and population totals. We now need to introduce additional notation:

- $\pi_S(i)$ is the known (inclusion) sampling probability that hunter i is included in the survey.
- $\pi_{R|S}(i)$ is the probability that hunter i $(i = 1, 2, \ldots, N_R)$ will respond to the survey given hunter i is included in the selected survey sample. This probability will have to be estimated.
- $\pi_R(i) = \pi_{R|S}(i) \times \pi_S(i)$ is the probability that hunter i will provide a response to the survey. This probability will also have to be estimated given that $\pi_{R|S}(i)$ must be estimated.

Note that:

- Hunters who have low $\pi_{R|S}(i)$ probabilities associated with responding to the survey (but actually did respond) will be assigned a large weight. The goal is to inflate their contribution when calculating estimated because hunters with similar auxiliary information are underrepresented among the N_R responders.
- Hunters who have high $\widehat{\pi}_{R|S}(i)$ probabilities associated with responding to the survey (and actually did respond) will be assigned a low weight. The goal is to deflate their contribution when calculating estimated because hunters with similar auxiliary information are overrepresented among the N_R responders.

Because the sampling protocol was a stratified random sample (and not a simple random sample), we have unequal ($\pi_S(i)$) sampling inclusion probabilities across strata, but which are equal within a stratum. Suppose hunter i was sampled from stratum h, and let N_h be the total number of LPT holders from stratum h. Based on the current sampling protocol, the value of $\pi_S(i)$ is

1	if $N_h \leq 375$	(All hunters in stratum h are in the survey.
.55	if $N_h \geq 1000$	(55% of all hunters in stratum $h$are in the survey.)
$1.27 - .00072N_h$	if $375 < N_h \leq 1000$	(based on a linear decrease between 100% and 55%.)

Although $\pi_{R|S}(i)$ is unknown, we will have an estimate $\widehat{\pi}_{R|S}(i)$ from the aforementioned prediction model (e.g., a logistic regression model or a classification tree). Therefore, our estimate of $\pi_R(i)$ is

$$\widehat{\pi}_R(i) \;=\; \widehat{\pi}_{R|S}(i) \times \pi_S(i).$$

The next step is to apply the $\widehat{\pi}_R(i)$ values. To see how these probabilities will be applied, consider the following example. Suppose we represent summary data from a future elk harvest phone survey by Table 4 .

Table 4: Summary count data for a future elk harvest phone survey

Phone Survey				
Cow	Calf	Bull	None	Total
$C_{.1}$	$C_{.2}$	$C_{.3}$	$C_{.4}$	$C_{..}$

where $C_{.j}$ are the total number of survey respondents who reported harvest category j (j = Cow, Calf, Bull, None). We know that each $C_{.j}$ is not representative of the responses for the population of hunters because of unequal sampling probabilities and nonresponse bias. Therefore, these counts need to be adjusted by taking into account the probabilities of receiving these responses given the sampling protocol and the potential nonresponse biases.

The adjusted count of the observed survey count $C_{.j}$ for harvest category i, denoted $C^*_{.j}$, is

$$C^*_{.j} \;=\; \sum_{k=1}^{N_S} \frac{I_j(k)}{\widehat{\pi}_R(k)} \tag{9}$$

where $I_j(k) = 1$ if hunter k's survey response is category j, and is 0 otherwise. That is, $I_j(k) = 0$ if hunter k does not respond or the response is some category $j' \neq j$. Therefore, nonrespondents have no impact on the adjusted value $C^*_{.j}$.

If we define weight $r_k = 1/\widehat{\pi}_R(k)$, we can rewrite $C^*_{.j}$ as

$$C^*_{.j} \;=\; \sum_{k=1}^{N_S} I_j(k)\, r_k \tag{10}$$

for category j = Cow, Calf, Bull, None. We now have survey counts adjusted for non-response bias for each response category. Table 5 is a summary of the adjusted counts.

In the next section, a method is described for applying adjustments due to nonresponse bias and the inclusion probabilities associated with the stratified sampling plan for selecting the sample of hunters to be in the phone survey.

Table 5: Summary count data for a future elk harvest phone survey adjusted for nonresponse bias

Phone Survey				
Cow	Calf	Bull	None	Total
$C^*_{.1}$	$C^*_{.2}$	$C^*_{.3}$	$C^*_{.4}$	$C^*_{..}$

2.1.3 Applying nonresponse bias adjustments and estimated response error rates

Because paired check station and phone survey data will be collected over the first two years of this study, these data can be aggregated and saved for creating the $\widehat{w}_{ij}$ values (defined in Section 2.1.1) for use in future years. After they are calculated they can be incorporated into each cell of any table having the same structure as Table 2.

We now summarize how to incorporate the random and prestige response bias estimates, the nonresponse bias adjustments, and the stratified sampling weights to estimate population and stratum total harvest. The first step is to generate a table of response error adjusted and non-response bias adjusted estimates of counts for each combination of "Truth" and "Reported" harvest types.

For example, consider a future elk harvest phone survey. Let $C_{.1}$, $C_{.2}$, $C_{.3}$, and $C_{.4}$ be the number of hunters who responded harvest types Cow, Calf, Bull, and None, respectively. Then:

1. Using the auxiliary variable information for each respondent and nonrespondent in the survey, fit a model to generate estimated probabilities $\widehat{\pi}_{R|S}(i)$ that hunter i ($i = 1, 2, \ldots, N_R$) will respond to the survey given hunter i is included in the survey.

2. Calculate the stratified sampling weight $\pi_S(i)$ which is the known (inclusion) sampling probability that hunter i is included in the survey..

3. Calculate $\widehat{\pi}_R(i) \;=\; \pi_{R|S}(i) \times \pi_S(i)$, the estimated probability that hunter i will provide a response to the survey.

4. Calculate the non-response adjusted survey harvest counts $C^*_{.1}$, $C^*_{.2}$, $C^*_{.3}$, and $C^*_{.4}$ where

$$C^*_{.j} \;=\; \sum_{k=1}^{N_S} I_j(k)\, r_k$$

   for category j = Cow, Calf, Bull, and None.

5. Apply the response error estimates m_r and m_{r+p} and the estimation formulas in Equations (1) to (7) in Section 2.1.1 to $C^*_{.1}$, $C^*_{.2}$, $C^*_{.3}$, and $C^*_{.4}$ to create a table of the nonresponse bias adjusted and response error adjusted estimates C^*_{ij}. These estimates are summarized in Table 6.

We can now calculate bias-adjusted estimates of hunter harvest within any of the levels of interest for this set of respondents. For example, to estimate true total cow harvest τ_1 at a statewide level we have

$$\widehat{\tau}_1 = \frac{N}{C^*_{..}}\,\left[\widehat{w}_{11}C^*_{.1} + \widehat{w}_{12}C^*_{.2} + \widehat{w}_{13}C^*_{.3} + \widehat{w}_{14}C^*_{.4}\right]$$

where $\widehat{\tau}_1$ is the adjusted estimate of cow harvest, N is the known total number of elk hunters (based on numbers of elk LPTs sold), the $C^*_{.j}$'s are the response error and nonresponse bias adjusted phone survey reported harvests of type j, $C^*_{..} = \sum_{j=1}^{4} C^*_{.j}$, and the $\widehat{w}_{1j}$'s are the estimated proportions of respondents who report a harvest of type j and actually harvested a cow.

Table 6: Estimated true (row) and reported (column) harvest frequencies in a future phone survey adjusted for nonresponse bias and response error.

	Phone Survey			
Truth	Cow	Calf	Bull	None
Cow	$\widehat{w}_{11}C^*_{.1}$	$\widehat{w}_{12}C^*_{.2}$	$\widehat{w}_{13}C^*_{.3}$	$\widehat{w}_{14}C^*_{.4}$
Calf	$\widehat{w}_{21}C^*_{.1}$	$\widehat{w}_{22}C^*_{.2}$	$\widehat{w}_{23}C^*_{.3}$	$\widehat{w}_{24}C^*_{.4}$
Bull	$\widehat{w}_{31}C^*_{.1}$	$\widehat{w}_{32}C^*_{.2}$	$\widehat{w}_{33}C^*_{.3}$	$\widehat{w}_{34}C^*_{.4}$
None	$\widehat{w}_{41}C^*_{.1}$	$\widehat{w}_{42}C^*_{.2}$	$\widehat{w}_{43}C^*_{.3}$	$\widehat{w}_{44}C^*_{.4}$
Column Total	$C^*_{.1}$	$C^*_{.2}$	$C^*_{.3}$	$C^*_{.4}$

In matrix notation,

$$\widehat{\boldsymbol{\tau}} = \begin{bmatrix} \widehat{\tau}_{cow} \\ \widehat{\tau}_{calf} \\ \widehat{\tau}_{bull} \\ \widehat{\tau}_{none} \end{bmatrix} = \frac{N}{C^*_{..}} \widehat{\mathbf{W}}\, \mathbf{c}^*_R = \frac{N}{C^*_{..}} \begin{bmatrix} \widehat{w}_{11} & \widehat{w}_{12} & \widehat{w}_{13} & \widehat{w}_{14} \\ \widehat{w}_{21} & \widehat{w}_{22} & \widehat{w}_{23} & \widehat{w}_{24} \\ \widehat{w}_{31} & \widehat{w}_{32} & \widehat{w}_{33} & \widehat{w}_{34} \\ \widehat{w}_{41} & \widehat{w}_{42} & \widehat{w}_{43} & \widehat{w}_{44} \end{bmatrix} \begin{bmatrix} C^*_{.1} \\ C^*_{.2} \\ C^*_{.3} \\ C^*_{.4} \end{bmatrix}$$

More generally, we can work with data at the level of the sampling strata and write the equations in matrix notation

$$\widehat{\boldsymbol{\tau}}_h = \frac{N_h}{C^*_{h..}} \widehat{\mathbf{W}}\, \mathbf{c}^*_{h_R}$$

where $\widehat{\boldsymbol{\tau}}_h$ is a vector containing the adjusted estimated harvest values for each level of a variable of interest in the h^{th} stratum, $\frac{N_h}{C^*_{h..}}$ is the scalar ratio of the total number of hunters in stratum h to the number from that stratum who responded to the phone survey, $\widehat{\mathbf{W}}$ is the square matrix of estimated proportions of hunters who actually harvest type i but report type j (see equation 8), and $\mathbf{c}^*_{h_R}$ is the vector of response error and nonresponse bias adjusted phone survey reported harvests for each level of the variable of interest. (Note that we are using the subscript R to identify data coming from phone survey responders, as opposed to the subscript M which we will later use for missing data).

To make this more clear it is useful to look at the calculation for estimating elk age-sex class harvests, substituting the i-subscripts on the τ's with their class level names

$$\widehat{\boldsymbol{\tau}}_h = \begin{bmatrix} \widehat{\tau}_{h_{cow}} \\ \widehat{\tau}_{h_{calf}} \\ \widehat{\tau}_{h_{bull}} \\ \widehat{\tau}_{h_{none}} \end{bmatrix} = \frac{N_h}{C^*_{h..}} \widehat{\mathbf{W}}\, \mathbf{c}^*_{h_R} = \frac{N_h}{C^*_{h..}} \begin{bmatrix} \widehat{w}_{11} & \widehat{w}_{12} & \widehat{w}_{13} & \widehat{w}_{14} \\ \widehat{w}_{21} & \widehat{w}_{22} & \widehat{w}_{23} & \widehat{w}_{24} \\ \widehat{w}_{31} & \widehat{w}_{32} & \widehat{w}_{33} & \widehat{w}_{34} \\ \widehat{w}_{41} & \widehat{w}_{42} & \widehat{w}_{43} & \widehat{w}_{44} \end{bmatrix} \begin{bmatrix} C^*_{h.1} \\ C^*_{h.2} \\ C^*_{h.3} \\ C^*_{h.4} \end{bmatrix}$$

Linear combinations of the strata-level estimators can then be used to get estimates at the resolution of interest.

2.2 Variance estimates

At this point we have estimators for the harvest parameters of interest, but without associated measures of precision any future estimates will be of little use. Derivations of closed-form estimators for the variances are tenuous due to unknown non-zero covariances between the levels of the variables of interest. Therefore we recommend resampling methods for variance estimation. Either the infinitesimal jackknife (Davison and Hinkley 1997, sections 2.7 and 3.2) or bootstrap methods could be used. Bootstrap methods should be less computationally intensive and that is the approach we describe here.

A potential problem with using the bootstrap technique is that the traditional bootstrap does not account for the reduction in variance due to sampling without replacement from a finite population. In a sample such as the FWP post-season harvest survey a relatively large proportion of the total number of hunters may be sampled, resulting in an important reduction in estimator variance. To account for this finite population issue we recommend using a finite population bootstrap technique outlined in Canty and Davison 1999. Using this technique, rather than taking resamples of size n (the original sample size) with replacement, resamples of size n are taken *without* replacement from a bootstrap population of size N (the size of the true population of interest), created by concatenating $\frac{1}{f}$ copies of the original sample of size n, where f is the proportion of hunters sampled from the population of size N.

Below we outline an algorithm for estimating the variance of the bias corrected estimators by applying the finite population bootstrapping technique to the data contained in the tables described above. Once again we motivate the method by using the age-sex class harvest estimates as a model. The procedure is as follows:

1. *Create a bootstrap population of the individuals that make up Table 1* (which is to be saved for use in future years). This is done by concatenating $\frac{1}{f_A}$copies of the data-set, where f_A is the proportion of the total population size (N) included in that data-set. When this value is not an integer make up the difference with a random sample. See Canty and Davison (1999) for details.

2. *Create a bootstrap population of the individuals that make up Table 2* (these data are also to be saved for use in future years). This is done in the same manner as described in Step 1.

3. *Create a bootstrap population of the N_R individuals that make up Table 4* (all responders in the year of interest). Again, use the method described in Step 1.

4. *Adjust the bootstrap replication in Step 3 that make make up Table 5 of nonresponse bias adjusted values.* Use the steps described in Section 2.1.1.

5. *Using the bootstrap data from Steps 1 and 2, follow the steps in section 2.1.1* to calculate a bootstrap replicate of the measurement error weights matrix (call it $\widehat{\mathbf{W}}^*$).

6. *Using the adjusted bootstrap data from Step 4, follow the steps in section 2.1.1* to calculate $\mathbf{c}^*_{h_R}$, and *follow the steps in section 2.1.2* to re-estimate the parameters of the chosen predictive model.

7. *Use the values from Steps 5 and 6 to calculate $\widehat{\tau}^*$'s of interest, and save the values.*

8. *Repeat these steps many times (e.g., 1000), and save the values each time.* The variance of the saved $\widehat{\tau}^*$ values is the bootstrap estimate of the variance of $\widehat{\tau}$.

2.3 Summary

Although this process is complex it can depicted by a relatively simple schematic (Figure 2). Notice there are two main components, one starting with the selected sample of individuals for the year of harvest estimation, and the other starting with the saved data for responders to the 2007 and 2008 surveys. Working downward from each of these, the figure depicts where each estimate is made and how they are combined to derive the harvest estimates. To calculate variances the entire process is repeated many times, each time using data collected from finite population bootstraps conducted at the levels that are shown in gray in the figure.

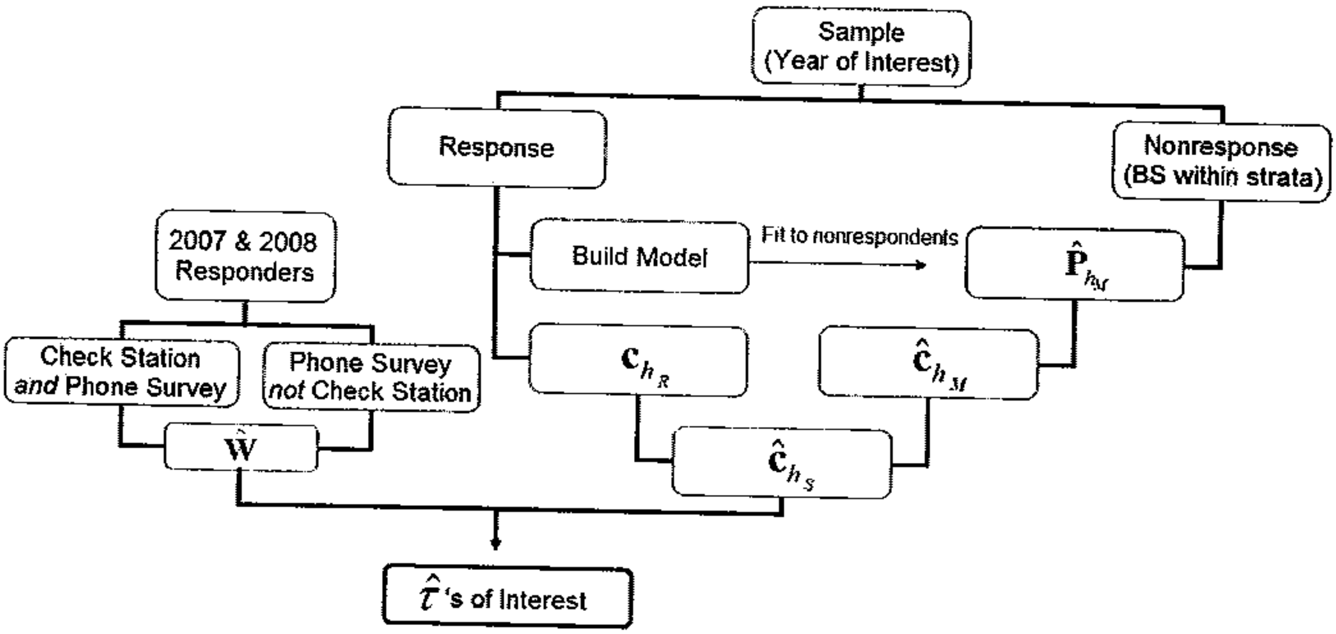

Figure 2: Steps involved in calculating harvest estimates while accounting for the effects of measurement error, nonresponse, and stratified sampling from a finite population. Grey fill indicates levels at which the finite population bootstrap sample is taken in each iteration of the variance estimation process.

3 Discussion

The goal in analyzing the FWP post-season harvest survey data is to extract valuable information from an observational study of a stratified, finite population with missing data and measurement error. In this report we have described methods for calculating harvest estimates while adjusting for biases stemming from measurement error and nonresponse, while accounting for stratified finite population sampling. As time allows, potential additions to this work include incorporating propensity score weights into the estimators, outlining methods for conducting sensitivity analyses on the effects of the various adjustments and violations of assumptions, and writing R or C code for the computations.

References

[1] Atwood, EL. 1956. Validity of mail survey data on bagged waterfowl. Journal of Wildlife Management 20:1-16.

[2] Canty, A. J., and A. C. Davison. 1999. Resampling-based variance estimation for labour force surveys. The Statistician 48:379-391.

[3] Chu, A, D Eisenhower, M Hay, D Morganstein, J Neter, and J Waksberg. 1992. Measuring the recall error in self-reported hunting activities. Journal of Official Statistics 8:19-39.

[4] Davison, A. C., and D. V. Hinkley. 1997. Bootstrap Methods and Their Application. Cambridge: Cambridge University Press.

[5] Geissler, PH. 1990. Estimation of confidence intervals for federal waterfowl harvest surveys. Journal of Wildlife Management 54:201-205.

[6] Gude, J. 2007. Improving harvest survey sampling and analysis techniques: project proposal. MT Fish Wildlife and Parks, unpublished proposal.

[7] Gude, J, C Hinz, B Beardslee, and K Aune. 2006. Elk hunting and harvest estimates, license year 2005. Research and Technical Services Bureau Wildlife Division, MT Fish Wildlife and Parks.

[8] Harrod, L.A., and V.M. Lesser. 2006. The use of propensity scores to adjust for nonignorable nonresponse bias. Proceedings of the Survey Research Section, American Statistical Association Meetings.

[9] Jones, K., and J Borkowski. 2007. Cooperative project on harvest survey sampling and estimation. Study plan I: Assessing measurement error and deriving correction methods. Submitted in partial fulfillment of Cooperative Agreement FWP #040076.

[10] Wright, VL. 1978. Causes and effects of biases on waterfowl harvest estimates. Journal of Wildlife Management 42:251-262.