ESTIMATING MEASUREMENT ERROR BIAS AND VARIANCE IN TWO-PHASE SAMPLES

Paul P. Biemer, Research Triangle Institute and
Dale Atkinson, National Agricultural Statistics Service

ABSTRACT

Let U denote the population to be surveyed and suppose we wish to estimate the total, M, of some characteristic, μ, of the individuals in the population. Suppose a large sample, S_1, is drawn and the measurement, y, is taken on each sample member where $y = \mu + \epsilon$ and ϵ is a measurement error. Further suppose a subsample S_2 of S_1 is drawn and, through some means (usually more costly to implement than the original survey) μ is measured exactly for each individual in S_2. The samples S_1 and S_2 are referred to as the first and second phase samples, respectively. In this paper, we consider a general methodology for using the data from both phases to estimate the total M as well as the mean and variance of the measurement error distribution associated with the observations, y.

This article is a condensed version of a more detailed proceedings article from the 1995 Annual Research Conference. The reader is referred to that article (Biemer and Atkinson (1995)) for a more complete description of the methodological development.

KEYWORDS Reinterview; Intra-interviewer Correlation; Response Error; Bootstrap

1. INTRODUCTION

Applications of two-phase sampling for evaluating the measurement error in survey data occur quite frequently in large scale national surveys. In some applications, the true values of the characteristics of interest are obtained through a reinterview of a sample of original survey respondents. These types of reinterviews usually employ survey methods which are much improved over those used in the original survey. For example, whereas the original survey may have used telephone interviewing, the reinterview may use face to face interviewing with better trained enumerators, and respondents may be strongly encouraged to consult their household records or other credible sources of information in responding to the survey questions. Further, the responses obtained in the reinterview may be compared with the original survey responses and discrepancies reconciled to produce a single, best response. In other studies, administrative records are used to produce a single, best response.

Because reinterview and record check studies are expensive to implement, the size of the second phase sample is typically quite small - say, 5-10% of the original sample. While these sample sizes may be adequate for estimating biases at the total population level, they may be inadequate for estimates for small subpopulations. In addition, the expense of the studies make it essential that the data from the second phase are exploited to their fullest potential. In the present paper, we present an integrated approach to the analysis of data from two-phase sample studies of measurement error. In particular, a general methodology is proposed for:

- producing improved estimates of measurement bias,
- estimating enumerator variance,
- estimating response reliability,
- producing unbiased estimates of M, and
- identifying enumerators that contribute maximally to enumerator variance.

In the next section, the components of measurement error that are the focus of our study are defined and some of the traditional estimators of these components are considered. In Section 3, the general methodology for estimating these measurement error components using two-phase sampling designs is discussed. This methodology is illustrated in Section 4 for an application to the Agricultural Survey conducted by the National Agricultural Statistics Service.

2. MEASUREMENT ERROR COMPONENTS

In this section, we define the measurement error target parameters to be estimated using a two-phase sampling design. To fix the ideas, we shall consider the case where S_1 and S_2 are selected by simple random sampling without replacement (SRSWOR). Generalizations to stratified random sampling are considered in Section 3.

Suppose the Phase 1 sample is partitioned into J groups of units denoted by G_g, $g = 1,..., J$ corresponding to the J enumerators available for the survey. Thus, G_g denotes the set of units assigned to enumerator g. Then for $i \in G_g$, assume the model

$$y_i = \gamma_0 + \gamma\mu_i + z_{gi} \tag{2.1}$$

where μ_i is the true value of the characteristic, γ_0 and γ are constants, and z_{gi} is a random error term. The parameter γ_0 may be interpreted as a constant or absolute bias that is added to all observations, while γ is a "proportional" bias term. As an example, suppose μ_i is some measure of farm size (i.e., number of acres, cattle, tons of wheat, etc.). The magnitude of the error in y_i is often proportional to size and thus is appropriately modeled by $\gamma\mu_i$. The term z_{gi} is the sum of two random components, d_g and δ_i, where d_g is the "bias" or "enumerator effect" associated with enumerator g and δ_i is an independent unit-level error. We assume that $d_g \sim (0, \sigma_d^2)$ and $\delta_i \sim (0, \sigma_\delta^2 \mu_i^\lambda)$ where λ is a known constant. For the present study, we shall set λ to 0; however, it is possible to estimate λ from the data (see, for example, Wright, 1983).

Further assume the conditional covariance of the errors for $i \in G_g$ is given by

$$\begin{aligned}\mathrm{Cov}\,(z_{gi}, z_{g'i'} | i) &= \sigma_d^2 + \sigma_\delta^2 \mu_i^\lambda \quad \text{for } i = i' \\ &= \sigma_d^2 \qquad \text{for } i \neq i';\ g' = g \\ &= 0 \qquad \text{for } g' \neq g\end{aligned} \tag{2.2}$$

Under these assumptions, the *measurement bias* in the measure y defined as $\bar{B} = \mathrm{E}(y_i - \mu_i)$, has the form

$$\bar{B} = \gamma_0 + (\gamma - 1)\bar{\mathrm{M}} \tag{2.3}$$

where $\bar{\mathrm{M}} = N^{-1}\sum_{i \in U} \mu_i$. The usual estimator of $\bar{B}$ is the difference estimator (Bureau of the Census, 1985; Hansen, Hurwitz, and Pritzker, 1964) given by either

$$\bar{B}_{12} = \bar{y}_1 - \bar{\mu}_2 \tag{2.4}$$

or

$$\bar{B}_{22} = \bar{y}_2 - \bar{\mu}_2 \tag{2.5}$$

where $\bar{y}_1$ is the average of the observations over S_1, $\bar{y}_2$ and $\bar{\mu}_2$ are averages of observations and true values, respectively, over S_2. Biemer and Atkinson (1993) show that the difference estimators are not efficient in many survey situations and propose several alternative estimators of $\bar{B}$ using a model which is essentially (2.1) setting $\gamma_0 = 0$ and $\sigma_d^2 = 0$. We will consider their estimators again in the context of (2.1) as well as several alternative estimators that arise from this more general formulation.

Let $\mathrm{E}(\bullet | i)$ denote conditional expectation given the unit i over the measurement error distribution and $\mathrm{Var}(\bullet)$ denote the unconditional variance with respect to the sampling distribution. If we assume that G_g, $g = 1,...,J$ are of equal size (say m) and that the finite population correction is ignorable, then Biemer and Stokes (1991) show that $n_1 \mathrm{Var}(\bar{y}_1) = R^{-1}\mathrm{Var}(\gamma\mu_i)[1 + (m-1)\rho_y]$ where R, referred to as the reliability ratio, is

$$R = \frac{\mathrm{Var}\mathrm{E}(y_i | i)}{\mathrm{Var}(y_i)} = \frac{\gamma^2 \sigma_\mu^2}{\sigma_y^2}, \tag{2.6}$$

$\sigma_y^2 = \mathrm{Var}(y_i)$, and the parameter, ρ_y, referred to as the *intra-enumerator correlation coefficient*, is the correlation between two units within an enumerator's assignment. Under model (2.1), ρ_y is given by

$$\rho_y = \frac{\sigma_d^2}{\sigma_y^2}. \tag{2.7}$$

The reliability ratio, R, is the ratio of the variance of the "true score" for the characteristic -- *viz*., $\mathrm{Var}(\gamma_0 + \gamma\mu_i)$ -- to the variance of the observation, y_i and is a widely used measure of data quality. Estimation of R usually requires repeated measurements that are obtained under identical survey conditions and such that the measurement errors associated with each measurement are independent (between measurements) and identically distributed (see Biemer and Stokes, 1991). However, for the present approach we shall estimate R directly from the two-phase sample data.

The intra-enumerator correlation coefficient is the ratio of the variance of an observation due to enumerators to the total variance of the observation. This parameter is widely used in measurement error studies to describe the degree to which the quality of interviewing varies across enumerators (see for example, Groves, 1989). A large estimate of ρ_y indicates that large enumerator biases (d_g) are present in the data. To identify which enumerators are most responsible for the large enumerator variance, an analysis of the d_g is necessary. In what follows, we will provide estimates of ρ_y as well as the d_g associated with the J enumerators for the survey.

3. ESTIMATION OF MEASUREMENT ERROR BIAS AND VARIANCE

3.1 A General Estimator of Measurement Bias

In this section, the general methodology for estimating the measurement error parameters is described in the context of two-phase, stratified random sampling.

Let $U = \{1,2,...,N\}$ denote the label set for a population consisting of L strata, denoted by U_h, $h = 1,...,L$. A two-phase random sample (with simple random sampling without replacement in both phases) is drawn in each stratum with Phase 1 and Phase 2 stratum sample sizes denoted by n_{1h} and n_{2h}, respectively. Let S_{1h} and S_{2h} denote the Phase 1 and Phase 2 samples in the h-th stratum. Let y_i denote the observation for unit $i \in S_{1h}$ and let μ_i denote the corresponding true value which exists for all $i \in U$ and is known for all units $i \in S_{2h}$.

A general estimator of response bias is proposed which encompasses the estimators considered in Biemer and Atkinson (1993), the general two-phase regression estimators of Särndal, Swensson, and Wretman (1992) and a general class of unweighted model-based estimators. This estimator is $\hat{B}_{\mathrm{G}}/N$ where

$$\hat{B}_{\mathrm{G}} = \hat{Y}_{\mathrm{G}} - \hat{\mathrm{M}}_{\mathrm{G}} \tag{3.1}$$

where $\hat{Y}_{\mathrm{G}}$ is an estimator of the total $Y = \sum_{i \in U} Y_i$, for $Y_i = \mathrm{E}(y_i | i)$ and where $\hat{\mathrm{M}}_{\mathrm{G}}$ is an estimator of $\mathrm{M} = \sum_{i \in U} \mu_i$, the population total of true values. In stratified sampling, these estimators have the form

$$\hat{Y}_{\mathrm{G}} = \sum_{h=1}^{L} \hat{Y}_{\mathrm{G},h} \tag{3.2}$$

and

$$\hat{\mathrm{M}}_{\mathrm{G}} = \sum_{h=1}^{L} \hat{\mathrm{M}}_{\mathrm{G},h}$$

where

$$\hat{Y}_{\mathrm{G},h} = \sum_{i \in S_{1h}} \frac{y_i}{\pi_i} + c_Y \left(\sum_{i \in U_h} \hat{y}_i - \sum_{i \in S_{1h}} \frac{\hat{y}_i}{\pi_i} \right) \tag{3.3}$$

is an estimator of $Y_h = \sum_{i \in U_h} Y_i$ and

$$\hat{M}_{G,h} = \sum_{i\in S_{2h}} \frac{\mu_i}{\pi_i^*} + c_M \left[\sum_{i\in U_h} \hat{\mu}_i^{(1)} - \sum_{i\in S_{1h}} \frac{\hat{\mu}_i^{(1)}}{\pi_i} + \sum_{i\in S_{1h}} \frac{\hat{\mu}_i^{(2)}}{\pi_i} - \sum_{i\in S_{2h}} \frac{\hat{\mu}_i^{(2)}}{\pi_i^*} \right] \quad (3.4)$$

is an estimator of $M_h = \sum_{i\in U_h} \mu_i$. In the above formulations $\pi_i = \Pr[i \in S_{1h}]$, $\pi_i^* = \Pr[i \in S_{2h} \mid i\in S_{1h}] \times \pi_i$, c_Y, c_M are constants in the interval [0,1], and, for $i\in U$, $\hat{y}_i$ is an estimator of Y_i, and $\hat{\mu}_i^{(1)}$ and $\hat{\mu}_i^{(2)}$ are estimators of μ_i.

Note that for the traditional stratified estimators, we let $c_Y = c_M = 0$, $\pi_i = n_{1h}/N_h$, and $\pi_i^* = n_{2h}/N_h$, yielding the familiar estimators $\hat{Y}_{st} = \sum_h N_h \bar{y}_{h1}$ and $\hat{M}_{st} = \sum_h N_h \bar{\mu}_{h2}$. Thus, when $c_Y = c_M = 0$, $\hat{B}_G$ is the stratified difference estimator (Biemer and Atkinson, 1993).

The Särndal, Swensson, and Wretman forms of $\hat{Y}_{G,h}$ and $\hat{M}_{G,h}$ are obtained from (3.3) and (3.4) by defining π_i and π_i^* as for the traditional stratified estimator, setting $c_Y = c_M = 1$, and using a linear regression equation to predict the unobserved Y_i's and unknown μ_i's. As an example, the estimator $\hat{B}_{SSW}$ considered by Biemer and Atkinson (1993) is obtained by assuming the model

$$y_i = \beta x_i + \varepsilon_i \quad (3.5)$$

for Y_i where β is a constant, x is an auxiliary variable known for all $i \in U$, and ϵ_i is an independent error term having zero mean and variance $\sigma_\varepsilon^2 x_i$. This yields the predictions $\hat{y}_i = x_i \bar{y}_{1st}/\bar{x}_{1st}$ which, when substituted into (3.3), produces their general estimator of Y. To estimate M they assume the models

$$y_i = \gamma\mu_i + \delta_i \quad (3.6)$$

for $i \in S_1{\sim}S_2$, and

$$\mu_i = \alpha x_i + \xi_i \quad (3.7)$$

for $i \notin S_1$, where γ and α are constants, $\delta_i \sim (0, \sigma_\delta^2 \mu_i)$ and $\xi_i \sim (0, \sigma_\xi^2 x_i)$. Thus, $\hat{\mu}_i^{(2)}$ in (3.4) is $\hat{\gamma}^{-1} y_i$ where $\hat{\gamma} = \bar{y}_{2st}/\bar{\mu}_{2st}$. Similarly, $\hat{\mu}_i^{(1)}$ in (3.4) is $\hat{\alpha} x_i$ where $\hat{\alpha} = \bar{\mu}_{2st}/\bar{x}_{2st}$. Thus, substituting these predictions in (3.4) yields their general estimator of M.

Finally, to obtain the purely model-based forms of these estimators we set $\pi_i = \pi_i^* = 1$ and $c_Y = c_M = 1$. As an example, the estimator $\hat{B}_M$ considered in Biemer and Atkinson (1993) is obtained under these assumptions using the models in (3.5), (3.6), and (3.7).

In the current work mixed models are used to predict the unobserved y's and μ's (i.e. $\hat{y}_i$, $\hat{\mu}_i^{(1)}$, and $\hat{\mu}_i^{(2)}$). Details of generating the predictions can be found in Biemer and Atkinson (1995).

3.2 Estimators of Enumerator Variance, Reliability and Enumerator Effects

Estimates of enumerator variance, reliability and the enumerator effects are obtained as a direct result of the modeling process in general, and specifically the model to predict μ_i with y_i in S_1. For example the estimator of ρ_y defined in (2.7) is

$$\hat{\rho}_y = \frac{\hat{\sigma}_d^2}{\hat{\sigma}_y^2} \quad (3.8)$$

where $\hat{\sigma}_d^2$ is the estimate of σ_d^2 obtained from fitting model (2.1),

$$\hat{\sigma}_y^2 = \frac{1}{N}\sum_h^L \frac{N_h}{n_{1h}} \sum_{i\in S_{1h}} y_i^2 - \bar{y}_{1st}^2 + v(\bar{y}_{1st}) \quad , (3.9)$$

where

$$v(\bar{y}_{1st}) = \frac{1}{N^2}\sum_h^L N_h(N_h - n_{1h}) \frac{s_{yh}^2}{n_{1h}}, \quad (3.10)$$

and

$$s_{yh}^2 = \sum_{i\in S_{1h}} (y_i - \bar{y}_{1h})^2/(n_{1h}-1). \quad (3.11)$$

We estimate R by

$$\hat{R} = \frac{\hat{\gamma}\,\hat{\sigma}_\mu^2}{\hat{\sigma}_y^2} \quad (3.12)$$

where

$$\hat{\sigma}_\mu^2 = \frac{1}{N}\sum_h^L \frac{N_h}{n_{2h}} \sum_{i\in S_{2h}} \mu_i^2 - \bar{\mu}_{2st}^2 + v(\bar{\mu}_{2st}), \quad (3.13)$$

$$v(\bar{\mu}_{2st}) = \frac{1}{N^2}\sum_h^L N_h(N_h - n_{2h}) \frac{s_{\mu h}^2}{n_{2h}}, \quad (3.14)$$

$$s_{\mu h}^2 = \sum_{i\in S_{2h}} (\mu_i - \bar{\mu}_{1h})^2/(n_{2h}-1), \quad (3.15)$$

and $\hat{\gamma}$ is the estimator of γ, again from fitting the model (2.1) to predict μ_i with y_i in S_1.

Finally, the enumerator effects are estimated by $\hat{d}_g$, the estimator of d_g, once again from fitting model (2.1).

4. APPLICATION TO THE AGRICULTURAL SURVEY

4.1 Description of the Survey

The National Agricultural Statistics Service (NASS) annually conducts a series of surveys which are collectively referred to as the Agricultural Survey Program. The purpose of these surveys is to collect data for specific agricultural commodities at the state and national levels in support of the agency's estimation program. Each December in the years 1988-1990, reinterview studies designed to assess the measurement bias in the data collected by Computer Assisted Telephone Interviewing (CATI) were conducted in six states : Indiana, Iowa, Minnesota, Nebraska, Ohio, and Pennsylvania. The reinterview techniques employed in these three studies are very similar to those used by the U.S. Census Bureau (see, for example, Forsman and Schreiner, 1991). However, unlike the Census Bureau's program, the major objective in the NASS studies is the estimation of measurement bias rather than interviewer performance evaluation.

As noted above, only Agricultural Survey responding units whose original interview was conducted by CATI were eligible for selection into the reinterview sample. The reasons for this restriction on sampling were primarily cost, timing, and convenience. However, a large proportion of the Agricultural Survey is conducted by CATI and, thus, information regarding Agricultural Survey measurement bias for this group would provide important information for the entire Agricultural Survey program.

For the NASS reinterview studies, the interviewing staff consisted of a mix of field supervisors and experienced field interviewers. This interviewing staff, which was a separate corps of interviewers from those used for CATI, conducted face-to-face reinterviews in a subsample of Agricultural Survey units for a subset of Agricultural Survey items. To minimize any problems that respondents may have with recall, the reinterviews were conducted within 10 days of the original interview. Differences between the original Agricultural Survey and reinterview responses were reconciled to determine the "true" value. Considerable effort was expended in procedural development, training, and supervision of the reinterview process to ensure that the final reconciled response was as accurate as possible. For the most part, the wording of the subset of Agricultural Survey questions asked in the reinterview was identical to that of the parent survey. The reinterviewers attempted to contact the most knowledgeable respondent in order to ensure the accuracy of the reconciled values.

In this report, only the 1990 data are analyzed. To conserve space, our presentation here is confined to five variables: cropland acreage, grain storage capacity, total land in farm, total hog/pig inventory, and winter wheat seedings. These variables adequately demonstrate the range of results observed for the entire set of variables we investigated. Table 3 presents the population sizes as well as the parent and reinterview sample sizes. Standard errors where computed using the two-phase bootstrap variance estimation method for two-phase samples described in Biemer and Atkinson (1993). As previously indicated, the variance-related parameter λ was set to 0 in all the analyses.

4.2 Estimation and Analysis

Table 4 shows the results of fitting both the mixed measurement error models described in this paper and the "no intercept" models proposed by Biemer and Atkinson (1993) to these data. For comparison purposes, the second and third columns of this table give the results for the traditional difference estimator. Note that the difference estimator does not utilize an explicit model. Thus, the rows labeled "mixed model" and "no intercept model" do not apply to the difference estimator. For the mixed model, we defined the group-term in the model as the enumerator assignment.

As was shown in Biemer and Atkinson (1993), substantial gains in precision over the traditional difference estimators are possible using a model prediction approach to bias estimation. The SSW estimators (both with and without an intercept) produced the estimates with the lowest standard errors for all five items studied. Furthermore, for all items except winter wheat seedings the two SSW estimators produced very comparable estimates.

Note the huge bias estimates (and standard errors) that resulted from the mixed model-based approach for total hogs and pigs, winter wheat seedings and grain storage capacity. These are all items which are fairly rare (i.e., a low percentage of positive reports). Given the size of the estimates for these items relative to those produced by the other estimators, we believe that they reflect a model bias. However, further work is needed to fully understand these results.

The estimator performance comparisons suggest that model-assisted predictions are generally better than the traditional approach. The model-based approach can be better, but tends to perform poorly for rare items. There were no uniformly best estimators employing either mixed or no-intercept models. However, for the variables considered in our analysis, the SSW estimator using a mixed model with random enumerator effects was frequently the most efficient estimator and seldom produced very unstable (as indicated by extremely large standard errors) or biased (as indicated by unusually large estimates of bias) estimates.

Table 4. Comparison of the Estimators of *B* Under No Intercept and Mixed Prediction Models

Item	Difference		SSW		Model-based	
	% Bias	S.E.	% Bias	S.E.	% Bias	S.E.
Total Land in Farm						
Mixed Model	0.4	4.54	-1.1	1.10	-1.9	2.65
No Intercept Model			-0.9	0.82	-5.2	1.95
Total Hogs and Pigs						
Mixed Model	-0.2	5.90	-5.2	2.22	53.3	11.06
No Intercept Model			-5.3	2.39	-5.4	3.54
Winter Wheat Seedings						
Mixed Model	6.6	8.03	-3.2	4.84	44.4	27.48
No Intercept Model			-2.9	2.47	9.5	7.81
Cropland Acreage						
Mixed Model	-3.4	2.26	-0.3	1.27	2.2	2.24
No Intercept Model			-0.1	1.10	-6.2	1.63
Grain Storage Capacity						
Mixed Model	-2.4	3.26	-2.0	3.94	23.6	18.09
No Intercept Model			1.2	2.81	-6.0	2.91

In Table 5, we display our best estimate of the relative bias for the five agricultural variables and estimates of ρ_y and R which were obtained from the two-phase sample data using model (2.1).

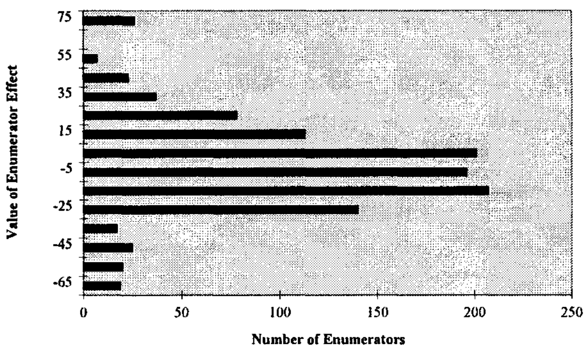
Table 5. Estimates of Bias, Enumerator Variance, and Reliability

Item	Relative Bias (%)	ρ_y	R
Total Land in Farm	-1.1	0.001	.68
Total Hogs and Pigs	-5.2	0.000	.89
Winter Wheat Seedings	-2.9	0.001	.58
Cropland Acres	-0.3	0.006	.98
Grain Storage Capacity	-2.0	0.000	.90

There was very little evidence of enumerator variance in any of the estimates; however, cropland acres exhibited the largest intra-enumerator correlation coefficient (ρ_y=.006). Total land in farm and winter wheat seedings exhibited the lowest reliabilities at .68 and .58, respectively. Further investigation is needed to understand the source of the confusion for these items. However, if the enumerators are contributing to the instability of these estimates, they are doing so consistently since ρ_y is very small. The largest relative bias was found for total hogs and pigs (-5.2%); however, the enumerator variance is small and reliability is relatively high. This would indicate that systematic measurement error is more of a problem than random error for this item. Had a large ρ_y been estimated for some item in our study, the next step might be to examine the estimates of enumerator effects to determine the source of the enumerator confusion.

We illustrate this approach using the estimates of enumerator effects for cropland acres, the item having the largest estimate of ρ_y. Figure 1 is a bar chart showing the frequency and range of enumerator effects for this variable, denoted by d_g in model 2.1. These estimates can be used to identify enumerator assignments most responsible for enumerator variance; for example, the enumerators associated with values of d_g at the tails of the distribution in Figure 1. In this way, efforts to reduce enumerator variance and measurement bias can be directed to areas of greatest need.

Figure 1. Enumerator Effects for Cropland Acres



5. SUMMARY AND CONCLUSIONS

In this paper, a general estimator of measurement bias was proposed for two-phase sampling which encompasses the traditional difference estimators as well as a large class of model-assisted and model-based estimators, including those considered by Biemer and Atkinson (1993). For the general estimator, the user specifies models for predicting the observations y_i for the population units that were not sampled in Phase 1 and for the true values, μ_i for the units not sampled in Phase 2. Virtually any model may be specified so long as it is based upon information that is available for the units to be predicted. For predicting Y_i, we used auxiliary information available from the sampling frame and either enumerator or stratum indicator variables. The models for predicting μ_i were similar except that units in the Phase 1 sample also incorporated the observed y_i's into the predictions of μ_i. Our approach for estimating measurement bias also produces, as a by-product of the estimation procedure, estimates of reliability, enumerator variance, and fixed enumerator effects. Thus, hypotheses regarding the magnitude of measurement bias as well as the major sources of the bias can be specified and tested using the two-phase sample data alone.

In an application to the Agricultural Survey, we estimated these components for a number of agricultural survey characteristics and presented the results for five of these. Standard errors were estimated using two-phase bootstrap variance estimation (Biemer and Atkinson, 1993). In these analyses, the model prediction approach out-performed the traditional difference estimator. However, among model prediction estimators, there was no clear winner. The SSW, mixed model estimator performed very well generally and appeared to be quite stable over the characteristics we examined. Further, the mixed model approach provides estimates of reliability, enumerator variance, and enumerator fixed effects as a by-product of the estimation process.

In analyzing estimates of measurement error parameters, characteristics are typically one of the following types: a) little or no measurement bias, little or no enumerator variance, and good reliability; b) moderate or large measurement bias, little or no enumerator variance, and poor reliability; and c) moderate or large measurement bias, moderate or large enumerator variance, and poor reliability. For case (a), there is little or no evidence of measurement error in the data. For case (b), there is evidence of measurement bias that is most likely caused by the questionnaire or the respondent but is not influenced by the enumerator. In case (c), there is evidence of measurement bias that may be caused by the questionnaire or the respondent and that *is* influenced by the enumerator. In the latter case, an examination of the individual enumerator effects would reveal which enumerators are most responsible for or whose assignments are most affected by enumerator variance. Thus, an analysis of the fixed enumerator effects can be used to direct efforts to reduce the measurement bias.

In a subsequent paper, we plan to consider alternative prediction models for estimating measurement bias. We also plan to provide estimates of the standard errors for the intra-enumerator correlation coefficient, the reliability ratio, and the enumerator effects using the two-phase bootstrap.

REFERENCES

Biemer, P. and Atkinson, D. (1993). "Estimation of Measurement Bias Using a Model Prediction Approach," *Survey Methodology*, Vol. 19, No. 2, pp. 127-136.

Biemer, P. and Atkinson, D. (1995). "An Integrated Approach for Estimating Measurement Error Bias and Variance in Two-Phase Samples," *Proceedings of the 1995 Annual Research Conference*, pp. 355-357.

Biemer, P. and S.L. Stokes (1991). "Approaches to the Modeling of Measurement Errors." In P.P. Biemer, R.M. Groves, L.E Lyberg, N.A. Mathiowetz, S. Sudman (Eds.) *Measurement Errors in Surveys.* New York: John Wiley & Sons, pp. 487-516.

Forsman, G. and Schreiner, I. (1991). The Design and Analysis of Reinterview: An Overview," in *Measurement Errors in Surveys*, P. Biemer, et al. (eds), John Wiley & Sons, N.Y.

Groves, R.M. (1989). *Survey Errors and Survey Costs*, John Wiley & Sons, N.Y.

Hansen, M.H., Hurwitz, W.N., and Pritzker, L., "The Estimation and Interpretation of Gross Differences and the Simple Response Variance," in C.R. Rao (ed.), *Contributions to Statistics*, Calcutta: Pergamon Press Ltd.

Särndal, C.-E., Swensson, B., and Wretman, J. (1992). *Model-Assisted Survey Sampling.* Springer-Verlag, N.Y.

U.S. Bureau of the Census (1985). *Evaluation of Censuses of Population and Housing*, STD-ISP-TR-5, Washington, D.C., Government Printing Office.

Wright, R.L. (1983). "Finite Population Sampling with Multivariate Auxiliary Information," *Journal of the American Statistical Association,* 78, pp. 879-884.