European Journal of Statistics and Probability
Vol.9, No.1, pp, 32-42, 2021
Print ISSN: 2055-0154(Print),
Online ISSN 2055-0162(Online)

CALIBRATION ESTIMATOR FOR POPULATION MEAN IN SMALL SAMPLE SIZE WITH NONRESPONSE

Iseh, Matthew Joshua, PhD[1]
Department of Statistics
Akwa Ibom State University, Mkpat Enin

Bassey, Kufre John, PhD
Output and Socio-economic Survey Unit
Real Sector Statistics Office, Statistics Department
Central Bank of Nigeria

ABSTRACT: *This paper addresses the challenges of estimation of population mean for small or no sample size in the presence of nonresponse and presents a calibration estimator that produces reliable estimates under stratified random sampling from a class of synthetic estimators using calibration approach. Examining the alternative estimator under three distributional assumptions, namely, normal, gamma, and exponential distributions through a simulation analysis with average absolute relative bias, average coefficient of variation, and average mean squared error as evaluation criteria, the results show that it has a consistent estimates of the mean with less bias and greater gain in efficiency. Further validation through the coefficient of variation also shows that the estimator exhibits a more preferred coefficient of variations suitable for small area estimation.*
KEYWORDS: calibration, distance measure, nonresponse, small area estimation, synthetic estimators.

INTRODUCTION

The use of synthetic estimators in small area estimation (SAE) has become one popular technique in small area estimation. This is so because it could produce reliable estimates when there are small or no sample observation in areas of interest as was first examined in 1968 by National Center for Health Statistics (NCHS) of the United States of America. This essential property of the synthetic estimators has made it so attractive in SAE unlike the direct estimators that are based on the sample information obtained from the area of interest which are not reliable due to lack of effective sample size in areas of interest. This indirect method of estimation has also been applied in the estimation of mean income of family, the average production of crops in blocks, and number of unemployed persons in the councils, among others.

Although synthetic estimation technique has been adopted by different authors like Gonzalez (1973), Sarndal (1981), Sarndal, Swensson and Wretman (1992), and Marker (1999) to compensate for the challenges of small sample sizes in SAE, and the use of calibration weights as a means of improving the precision (Rao, 2003; Lundstrom and Sarndal,1999, 2001; Sarndal and Lundstrom, 2005, 2008; Sarndal, 2007; Lehtonen *et al.*, 2003; Kott, 2006; Lehtonen and Veijanen, 2012; Lehtonen and Veijanen, 2015; Pfeffermann, 2013; Rota and Laitila, 2015; Rao and Molina, 2015 and Rota, 2016), the challenge of small or no sample size in the presence of nonresponse still remains a gap in the literature.

Deville and Sarndal (1992) earlier introduced the calibration estimation approach with distance function to account for the auxiliary information in the estimation, a method often refers to as "creating estimators by bench marking the auxiliary information to external controls". Thereafter, Lundstrom and Sarndal (1999, 2001), posited that the distance measure proposed by Deville and Sarndal is not effective in addressing the dual problem of small sample size and nonresponse in a domain of interest (Guisti and Rocco, 2013 advocated for an estimator that could address the dual problem of small sample size in the presence of nonresponse in a domain of interest.). After a study by Hidiroglou and Estavao (2014) showed that calibration estimators performed poorly when sample sizes become very small but more efficient as sample size increased, whereas, synthetic estimator become more effective at domains with small sample sizes, Andersson (2017) proposed a new distance measure for calibration weight to bridge the gap of "how close the calibration weight is to the design weight under nonresponse".

Following the works of Lundstrom and Sarndal (1999, 2001) and Anderson (2017), this paper proposes an alternative synthetic estimator that addresses the limitations in the previous studies using calibration approach.

THEORETICAL UNDERPINNING

Consider a finite population $U = 1,2,\ldots,N$ consisting of Nunits. The population can be divided into D -nonoverlapping domains $U_d, d = 1,2,\ldots,N_d$ units such that $\sum_d N_d = N$. Let the population be further partition into G -nonoverlapping groups (considered to be strata) which are considered to be larger than the domains $U_g, g = 1,2,\ldots,G$ consisting of $N_g$units such that $\sum_g N_g = N$. The case considered here occurs when the G - groups cut across the D -domains to form a grid of DGcells denoted by $U_{dg}, d = 1,2,\ldots,D, g = 1,2,\ldots,G$.

Let N_{dg}be the size of the U_{dg}such that $N = \sum_d N_d = \sum_g N_{.g} = \sum_g \sum_d N_{dg}$. The sample s is analogously partitioned into domain subsamples S_d, group subsamples $S_{.g}$ and cells subsamples $S_{dg}$with sizes $n_d, n_{.g}$ and n_{dg} respectively, such that $n = \sum_d n_d = \sum_g n_{.g} = \sum_g \sum_d n_{dg}$. Supposed that Ybe the study variable, which values $y_{dgk}$are known for just the element of a sample, s, of the k^{th}unit in the $(dg)^{th}$cell, where $k = 1,2,\ldots,N_{dg}$ (the number of population units in the $(dg)^{th}$ cell), and X be auxiliary variable, which values $x_{dgk} > 0$ may or may not be known for all units in U. Then, each k^{th} unit has an inclusion probability $\pi_k = P(k \in s)$with design weight $d_k = \pi_k^{-1}$ and a stratum weight $W_{dg} = N_d^{-1} N_{dg}$.

For different reasons, there are missing units in the sample, s. If we further denote the response set by r, and instead of the original sample size, s, we receive complete response for n_r, then, the response probability $P(k \in s\_r \mid k \in s)$, where $s\_r \subset s$, is a responded sample.

Now, let us consider the following estimators for domain estimation in the presence of nonresponse:

a. domain estimator of population mean in the presence of unit nonresponse (direct estimation):

Sarndal, Swensson and Wretman (1992), suggested an estimator under nonresponse in estimating domain population mean $\bar{Y}_d = \frac{1}{N_d}\sum_g \sum_k Y_{dgk}$as:

$$\hat{\bar{y}}_{dr} = \frac{1}{N_d}\sum_g \sum_{k \in s_r} \frac{d_k}{\theta_k} y_{dgk} \qquad (1)$$

where $\theta_k$is the influence probability of inclusion due to nonresponse. Equation (1) is an extension of the basic Horvitz-Thompson estimator to a selection in two phases.

b. calibration estimator of population mean in the presence of unit nonresponse(calibration approach):

Lundstrom and Sarndal (1999, 2001) proposed a single step weighting scheme through calibration approach as an improvement to Equation (1) for estimating the domain population mean in the presence of nonresponse as:

$$\hat{\bar{y}}_d = \sum_{s_r} P_{dg}\, \bar{y}_{dg} \qquad (2)$$

where P_{dg} is the calibration weights formed to be as 'close as possible' to the basic stratum weights W_{dg} in stratified random sampling at the two levels of information which satisfy the calibration equations.

When information is available at the population level of the auxiliary variable, the calibration weight becomes $P_{dg} = W_{dg} v_k$ and the calibration estimator in equation (2) becomes

$$\hat{\bar{y}}_{dU} = \sum_{s_r} W_{dg} v_{kU}\, \bar{y}_{dg} \qquad (3)$$

where $v_{kU} = 1 + q_{dg}\left(\sum_{U_d} \bar{x}_{dg} - \sum_{s_r} W_{dg}\bar{x}_{dg}\right)\left(\sum_{s_r} W_{dg} q_{dg} \bar{x}_{dg}^2\right)^{-1} \bar{x}_{dg}$

Equation (1) was compared to Equation (3) by assuming that $v_{kU} = \frac{1}{\theta_k}$. However, when information on the population of the auxiliary variable is unknown, the calibration estimator was obtained as:

$$\hat{\bar{y}}_{ds} = \sum_{s_r} W_{dg} v_{ks}\, \bar{y}_{dg} \qquad (4)$$

where $v_{ks} = 1 + q_{dg}\left(\sum_{s} W_{dg}\bar{x}_{dg} - \sum_{s_r} W_{dg}\bar{x}_{dg}\right)\left(\sum_{s_r} W_{dg} q_{dg} \bar{x}_{dg}^2\right)^{-1} \bar{x}_{dg}$

Print ISSN: 2055-0154(Print),
Online ISSN 2055-0162(Online)

The estimators in equations (3) and (4) are very unstable with small sample size. More so when the domain of interest has no sample unit it becomes difficult (if not impossible) to be computed given that they are modified direct estimators.

METHODOLOGY

Let us start by considering a synthetic estimator using the Lundstrom and Sarndal (1999, 2001).

Lemma: Supposed that preference is given to the groups as a powerful factor in explaining individual variation of elements within groupsg's $(\mathrm{g}=1,2,\ldots,\mathrm{G})$ considered to be homogeneous for small area d's $(\mathrm{d}=1,2,\ldots,\mathrm{D})$ under stratified sampling. Let the groupsg's $(\mathrm{g}=1,2,\ldots,\mathrm{G})$ be similar for small area d's $(\mathrm{d}=1,2,\ldots,\mathrm{D})$
under stratified sampling. Then, the Lundstrom and Sarndal (1999, 2001) could be modified to obtain a calibration synthetic estimator for population mean in small area in the presence of nonresponse as follows:

(i) $\hat{\bar{y}}^{*}_{dCU}=\sum_{s_r} W_{dg}\vartheta_{kU}\,\bar{y}_{.g}$ *(when information from the auxiliary variable is available at the population level), where* $\vartheta_{kU}=1+q_{dg}\left(\sum_{U_d}\bar{x}_{.g}-\sum_{s_r}W_{dg}\bar{x}_{.g}\right)\left(\sum_{s_r}W_{dg}q_{dg}\bar{x}^{2}_{.g}\right)^{-1}\bar{x}_{.g}$, for $g\in s_r$
and,

(ii) $\hat{\bar{y}}^{*}_{dCS}=\sum_{s_r} W_{dg}\vartheta_{ks}\,\bar{y}_{.g}$ *(when information from the auxiliary variable is only available at the sample level), where* $\vartheta_{ks}=1+q_{dg}\left(\sum_{s}W_{dg}\bar{x}_{.g}-\sum_{s_r}W_{dg}\bar{x}_{.g}\right)\left(\sum_{s_r}W_{dg}q_{dg}\bar{x}^{2}_{.g}\right)^{-1}\bar{x}_{.g}$

Proof: Let the calibration synthetic estimator $\hat{\bar{y}}^{*}_{dC}$ be given as

$$\hat{\bar{y}}^{*}_{dC}=\sum_{s_r} W^{*}_{dg}\,\bar{y}_{.g} \qquad (5)$$

where $\bar{y}_{.g}=\sum_d\sum_k\frac{y_{dgk}}{n_{.g}}$, $n_{.g}=\sum_d n_{dg}$and W^{*}_{dg} the chosen calibration weights such that the chi-square type distance measure:

$$\Phi=\sum_{s_r}\frac{\left(W^{*}_{dg}-W_{dg}\right)^2}{W_{dg}q_{dg}} \qquad (6)$$

is minimized, while satisfying the calibration constraints:

$$\sum_{s_r}W^{*}_{dg}\bar{x}_{.g}=\sum_{U_d}\bar{x}_{.g} \qquad (7)$$

and,

$$\sum_{s_r}W^{*}_{dg}\bar{x}_{.g}=\sum_{s}W_{dg}\bar{x}_{.g} \qquad (8)$$

Case 1: Availability of information for auxiliary variable at the population level

Assume that information from the auxiliary variable is available at the population level, Info-U: then minimizing the distance function in (6) subject to the calibration constraint in (7) will give the following calibration weights;

$W_{dg}^{*} = W_{dg}\vartheta_{kU}$
(9)
substituting (9) in (5) will give
$\hat{\bar{y}}_{dCU}^{*} = \sum_{s_r} W_{dg}\vartheta_{kU}\,\bar{y}_{.g}$
(10)
where ϑ_{kU} is as earlier defined $\vartheta_{kU} = 1 + q_{dg}\left(\sum_{U_d}\bar{x}_{.g} - \sum_{s_r} W_{dg}\bar{x}_{.g}\right)\left(\sum_{s_r} W_{dg}q_{dg}\bar{x}_{.g}^{2}\right)^{-1}\bar{x}_{.g}$

Case 2: Non-availability of information for the auxiliary at the population level of the domain
Suppose that there is no information on the population mean of the auxiliary variable in the domain, calibration can be done on the unbiased estimate $\sum_{s} W_{dg}\bar{x}_{.g}$. Here, minimizing the distance measure in (6) subject to the calibration constraint in (8) will result in the calibration weights:
$W_{dg}^{*} = W_{dg}\vartheta_{ks}$
(11)
and substituting (11) in (5) will result in a new estimator under Info-S as follows:
$\hat{\bar{y}}_{dCS}^{*} = \sum_{s_r} W_{dg}\vartheta_{ks}\,\bar{y}_{.g}$
(12)
where ϑ_{sk} is as earlier defined as $\vartheta_{ks} = 1 + q_{dg}\left(\sum_{s} W_{dg}\bar{x}_{.g} - \sum_{s_r} W_{dg}\bar{x}_{.g}\right)\left(\sum_{s_r} W_{dg}q_{dg}\bar{x}_{.g}^{2}\right)^{-1}\bar{x}_{.g}$

Note: Athough the estimators in equations (10) and (12) are useful in areas where there are small/no sample sizes, they are biased when there is small sample size and nonresponse.

New Calibration Estimator with Alternative Weights for Small Area in the presence of Nonresponse.
Here, we proposed a new estimator with alternative distance measure and a new design weight d_k^{*} (which is the product of the original design weight $d_k$and the inverse of the nonresponse influence probability $\hat{\varphi}$) for the estimation of population mean $\bar{Y}_d$ to resolve the challenges of biasness and higher mean square error due to small sample size and nonresponse in small area estimation.

Proposition: Let the estimator of population mean for small area in the presence of nonresponse be$\hat{\bar{y}}_{dr} = \frac{1}{N_d}\sum_{g}\sum_{k\in s_r}\frac{d_k}{\theta_k}y_{dgk}$,then, an alternative estimator $\hat{\bar{y}}_{dr}^{o} = \bar{X}_d^{T}\hat{B}_{drc}$ that can produce reliable estimates under stratified random sampling can be obtained by defining a new design weight and calibrating on an alternative distance measure $\sum_{s_r}\frac{\left(W_{dg}^{o}-W_{dg}\right)^{2}}{W_{dg}(W_{dg}-1)}$.

Proof: Recall the influence probability$\theta_k$in equation (1), and let the estimate of its inverse be given as $\hat{\varphi} = \frac{1}{\hat{\theta}_k} = \frac{n_{dg}}{n_r}$, then one can obtain a new design weight under nonresponse for the distance minimization as:
$d_k^{*} = d_k\hat{\varphi}$ (13)
and under stratified sampling, equation (1) can be written as:

$$\hat{\bar{y}}^*_{dr} = \frac{1}{N_d}\sum_g \sum_{k\in s_r} d^*_k y_{dgk}$$

$\hat{\bar{y}}^*_{dr} = \sum_{s_r} W_{dg}\bar{y}_{.g}$

(14) Thus, the estimator for the population mean using calibration approach is given as:

$\hat{\bar{y}}^0_{dr} = \sum_{s_r} W^0_{dg}\bar{y}_{.g}$

(15)

where W^o_{dg} is the chosen calibration weight such that the distance function:

$$\Phi = \sum_{s_r} \frac{\left(W^o_{dg} - W_{dg}\right)^2}{W_{dg}(W_{dg}-1)} \qquad (16)$$

is minimized subject to the calibration constraint:

$\sum_{s_r} W^o_{dg}\bar{x}_{.g} = \sum_{s} W_{dg}\bar{x}_{.g}$

(17)

and the calibration weights obtained as:

$W^o_{dg} = W_{dg}\bar{x}^2_{.g}\left(\sum_{s_r} W_{dg}\,\bar{x}^2_{.g}\right)^{-1}\bar{X}_d$

(18)

given that $\sum_s W_{dg}\bar{x}_{.g} = \bar{X}_d$.

Substituting (18) in (15) gives a new calibration estimator for small area in the presence of nonresponse as

$$\hat{\bar{y}}^o_{dr} = \bar{X}_d \hat{B}_{drc} \qquad (19)$$

Equation (19) takes the form of the global regression-synthetic estimator of the population mean for small area, (see, for example, Rao, 2003), where

$$\hat{B}_{drc} = \frac{\sum_{s_r} W_{dg}\bar{x}_{.g}\bar{y}_{.g}}{\sum_{s_r} W_{dg}\bar{x}^2_{.g}} \qquad (20)$$

FINDINGS AND DISCUSSION

In this section, empirical investigation is carried out using simulation analysis in R. The procedure of population generation and sample selection for different sample settings in the simulation analysis is adopted from Hidiroglou and Estevao (2014). However, three different probability distributions are considered, namely; normal, gamma, and exponential distributions.

Findings:

First, bivariate observations (x_{ij}, y_{ij}) were generated which comprised finite population of size 4950 units. A population, U, was created by generating data for three separate subpopulations termed groups (strata) with different intercepts and slopes. Group1 was split into ten cells $U_{11}, \dots, U_{110}$; Group 2 into ten cells, $U_{21}, \dots, U_{210}$; and Group 3 into ten cells, $U_{31}, \dots, U_{310}$. The population were further partitioned into ten domains $U_1, \dots, U_{10}$ with the groups larger than the domains and cut across the domains, forming a total of thirty grids that were mutually exclusive and exhaustive. The number of units in each cell, N_{dg} were allocated in a monotonic manner: cell U_{11} with 20 units; cell U_{12} with 30 units; and cell U_{310} with 310 units.

The values of x in each group were generated from three different distributions: Gamma $(\alpha = 5, \beta = 10)$, Normal (5,1) and Exponential (1.5) distributions. Then the simulation for the variable of interest y was obtained using the model $y_{dk} = \beta_{og} + \beta_{1g}x_{dk} + v_d + e_{dk}$, where $d = 1,2,\ldots,30; k = 1,2,\ldots,N$ and $g = 1,2,3$; $e_{dk} \sim N(0, C_{dk}^2\sigma_e^2), v_d \sim N(0, \sigma_v^2)$.

It was assumed that $\sigma_e^2 = \sigma_v^2 = 20^2 = 400$ for the gamma distribution; and $\sigma_e^2 = \sigma_v^2 = 1^2 = 1$ for normal and exponential distributions, respectively. To reflect the heterogeneity of the model errors for the synthetic and calibration estimators we set $c_{dk} = x_{dk}$. The summary of the representation of units in each group across the domains and the results of the evaluation under nonresponse are as presented in the Tables (Appendix A).

DISCUSSION

Table 1 shows how the population was split into three groups with the respective values of intercepts and slopes for gamma, normal and exponential distributions. Table 2 presents the population of a broad domain under study divided into sub domains and further partitioned into groups that are larger than the domains but cut across the domains to form grids that are mutually exclusive and exhaustive. Each simulation run in Table 2 involves the selection of $R = 100{,}000$, using R software for independent samples and the computation of various estimates for sample of sizes $n = 248(5\%)$, $n = 495(10\%)$, $n = 744(15\%)$, $n = 990(20\%)$, and $n = 1239(25\%)$ drawn using SRSWOR from U. An assumption of 89% for the response rate has been considered for all sample settings.

Table 3 (Appendix A) presents the results of the evaluation under nonresponse, where the percent average absolute relative bias of the optimum calibration estimator $\hat{\hat{y}}_{dr}^0$ was smallest in all sample settings with 6% for normal distribution, 13.1% to 13.2% for gamma and 13.1% to 13.3% for exponential distributions, making it a more reliable calibration estimator for small area. The values are better than other calibration synthetic estimator $\hat{\hat{y}}_{drSC}^*$ under nonresponse with constant $\overline{\%ARB}$ of 52.8%, 61.0% and 63.5% for gamma, normal and exponential probability distributions, respectively, and negligible compared to that of the existing estimator $\hat{\hat{y}}_{ds}$ with constant $\overline{\%ARB}$ of 115.5% and 114.9% for gamma and normal distributions, respectively, as well as 117.1% to 117.4% for exponential distribution. The performance of the new calibration estimator is an improvement of the Andersson (2017) alternative distance measure over that of Lundstrom and Sarndal (1999, 2001).

As expected, this result agrees with the argument of Andersson (2017) on the choice of weights in the presence of nonresponse. The effective reduction in bias of $\hat{\hat{y}}_{dr}^0$ further justifies the suggestion by Guisti and Rocco (2013) on addressing both small area problems and nonresponse adjustment as it has smaller average MSE under nonresponse for normal and exponential distributions. Although existing calibration estimator $\hat{\hat{y}}_{ds}$ outperformed the new estimators in terms of average

MSE under gamma distribution in all the sample settings considered in this study, the new estimator is shown to be consistently more precise under the normal distribution, followed by exponential and the gamma distributions. Other than weight adjustments, this result is in tandem with the suggestion of Sarndal, *et. al.* (1992) that partitioning the elements perceived to belong to the response homogeneous groups (RHGs) help in reducing the variance of the interest variable.

Furthermore, the new estimator has $\overline{\%CV}$ of 6.1% to 6.4% for normal distribution which is less than 10% and makes the bias to be negligible, 13.4% to 14.0% for gamma and 14.7% to 22.0% for exponential distributions for different sample settings. These values fall within the benchmark of 25% proposed by Molina and Rao (2010) for small area estimators and found to be very suitable for small area estimation in the presence of nonresponse.

Implication to Research

This paper considered the use of calibration weighting scheme in small area estimation under stratified sampling design to produce reliable synthetic estimators of population mean. The paper also presented a calibration estimator with alternative weighting scheme that exhibits smaller relative biases, gain in efficiencies and highly preferred coefficient of variations suitable for small area estimation. In terms of the probability distributions, the proposed estimator is more consistent in performance with Normal distributions. This supports the idea of calibration technique using the assumed constraint on synthetic estimators under stratified sampling design for greatly improving on the precision of estimators even in areas where there are smaller/no sample observation. The result of the coefficient of variation also suggests that when nonresponse occurs, it corresponds to an additional phase of sampling determined by original sample design in line with Guisti and Rocco (2013).

CONCLUSION

In conclusion, the concept of calibration on new design weights and an alternative distance measure is a contribution towards the development of an ultimate estimator for small area estimation under nonresponse and has yielded a fruitful result compared to the Lundstrom and Sarndal (1999, 2001) approach. This result will further enhance the disaggregation of national data and effective estimation in small areas with negligible biases in areas where there are small/no sample observation for proper policy implementations.

REFERENCES

Anderson, P. G. (2017). Optimal calibration weights under nonresponse in survey sampling. https://statswiki.unece.org. Retrieved February 21, 2018.

Deville, J. C. & Sarndal, C. E. (1992). Calibration estimators in survey sampling. *Journal of the American Statistical Association, 87,* 376-382.

Gonzales, M. E. (1973). Use and evaluation of synthetic estimator. *Proceedings of the Section on Social Statistics, American Statistical Association,* 33-36.

Guisti, C. & Rocco, E. (2013). *Small area estimation in the presence of nonresponse.* www.ds.unifi.it. Retrieved December 29, 2017.

Hidiroglou, M. A. & Estevao, V. M. (2014). A comparison of small area and calibration estimators via simulation. *Statistics in Transition*, *New Series, 17*(1), 133-154.

Horvitz, D. G. & Thompson, D. J. (1952). A generalization of sampling without replacement from a finite universe. *Journal of the American Statistical Association, 47*(260), 663-687.

Kott, P. S. (2006). Using calibration weighting to adjust for nonresponse and coverage errors. *Survey Methodology, 32,* 133-142.

Lehtonen, R., Sarndal, C. E. & Veijanen, A. (2003). The effect of model choice in estimation for domains, including small domains. *Survey Methodology, 1,* 33-44.

Lehtonen, R., & Veijanen, A. (2012). Small area poverty estimation by model calibration. *Journal of the Indian Society of Agricultural Statistics, 66,* 125-133.

Lehtonen, R., & Veijanen, A. (2015). *Small area estimation by calibration methods.* world statistics conference 2015, Rio de Janeiro.

Lundstrom, S. & Sarndal, C. E. (1999). Calibration as a standard method for treating nonresponse. *Journal of Official Statistics, 15*(2), 305-327.

Lundstrom, S. & Sarndal, C. E. (2001). *Estimation in the presence of nonresponse and frame imperfections.* Statistics Sweden.

Marker, D. S. (1999). Organization of small area estimation using generalized linear regression framework. *Journal of Official Statistics*, *15*(1), 1-24.

Molina, I. & Rao, J. N. K. (2010). Small area estimation poverty indicators. *Canadian Journal of statistics, 38,* 369-385.

National Center for Health Statistics (1968). *Synthetic state estimates of disability*. Public Health Service Publication. (No. 1759), Washington: U. S. Government Printing Office.

Pfeffermann, D. (2013), "New Important Developments in Small Area Estimation" *Statistical Sciences*, 28, 40-68.

Rao, J. N. K. (2003). *Small area estimation*. New York: John Wiley

Rao, J. N. K. & Molina, I. (2015). *Small area*. Second edition. New York: John Wiley

Rota, B. J. & Laitila, T. (2015). Comparison of some weighting methods for non-response adjustment. *Lithuanian Journal of Statistics*, *54*(1), 69-83.

Rota, B. J. (2016). *Calibration adjustment for nonresponse in sample surveys*. www.oru.se/publikationer-avhandlingar. Retrieved February 25, 2018.

Sarndal, C. E., Swensson. B. & Wretman, J. (1992). *Model-assisted surveys*. New York: Springer-Verlag.

Sarndal, C. E. (1981). When robust estimation is not an obvious answer: The case of the synthetic estimator versus alternatives for small areas. *Proceedings of the American Statistical Association, Survey Research Section*, 53-59.

Sarndal, C. E. & Lundstrom, S. (2005). *Estimation in survey with nonresponse*. New York: John Wiley.

Sarndal, C. E. (2007). The calibration approach in survey theory and practice. *Survey Methodology*, *33*, 99-119.

Sarndal, C. E. & Lundstrom, S. (2008). Assessing auxiliary vectors for control of nonresponse bias in the calibration estimator. *Journal of Official Statistics*, *24*, 167-191.

Appendix A: Empirical Evaluation Results

DISTRIBUTIONS		**GAMMA**		**NORMAL & EXPONENTIAL**	
GROUP $(\boldsymbol{g})$	Cells in groups	$\boldsymbol{\beta_{0g}}$	$\boldsymbol{\beta_{1g}}$	$\boldsymbol{\beta_{0g}}$	$\boldsymbol{\beta_{1g}}$
1	U_{d1} for $k = 1,2,\ldots,10$	200	30	5	1.5
2	U_{d2} for $k = 11,12,\ldots,20$	300	20	10	1.0
3	U_{d3} for $k = 21,22,\ldots,30$	400	10	15	0.5

Table 1: Partitioning the Population into Groups with their respective slopes and intercepts for Gamma, Normal and Exponential Distributions

	Domains (d)			
Groups (g)	1	2	3	U_d
1	U_{11}	U_{12}	U_{13}	U_1
2	U_{21}	U_{22}	U_{23}	U_2
3	U_{31}	U_{32}	U_{33}	U_3
4	U_{41}	U_{42}	U_{43}	U_4
5	U_{51}	U_{52}	U_{53}	U_5
6	U_{61}	U_{62}	U_{63}	U_6
7	U_{71}	U_{72}	U_{73}	U_7
8	U_{81}	U_{82}	U_{83}	U_8
9	U_{91}	U_{92}	U_{93}	U_9
10	U_{101}	U_{102}	U_{103}	U_{10}
$\boldsymbol{U_{.g}}$	$U_{.1}$	$U_{.2}$	$U_{.3}$	U

Table 2: Summary Representation of the Units in each Group across the Domains

Print ISSN: 2055-0154(Print),
Online ISSN 2055-0162(Online)

		$\overline{\%ARB}$			$\overline{MSE}$			$\overline{\%CV}$		
SAM	DIS	$\hat{\bar{y}}_{ds}$	$\hat{\bar{y}}^*_{drSC}$	$\hat{\bar{y}}^0_{dr}$	$\hat{\bar{y}}_{ds}$	$\hat{\bar{y}}^*_{drSC}$	$\hat{\bar{y}}^0_{dr}$	$\hat{\bar{y}}_{ds}$	$\hat{\bar{y}}^*_{drSC}$	$\hat{\bar{y}}^0_{dr}$
5%	Gam	115.5	52.8	13.1	2.0×10^9	3.8×10^{10}	2.3×10^9	115.6	52.8	14.0
	Norm	114.9	61.0	**6.0**	3.3×10^9	1.0×10^7	1.2×10^5	114.9	61.0	**6.4**
	Exp	117.4	63.5	13.3	2.1×10^9	6.8×10^6	3.7×10^5	118.0	63.9	16.0
10%	Gam	115.5	52.8	13.2	2.0×10^9	3.8×10^{10}	2.3×10^9	115.6	52.8	13.6
	Norm	114.8	61.0	**6.0**	3.3×10^9	1.0×10^7	1.2×10^5	114.9	61.0	**6.2**
	Exp	117.2	63.6	13.2	2.0×10^9	6.8×10^6	3.7×10^5	117.5	63.6	14.7
15%	Gam	115.5	52.8	13.2	2.0×10^9	3.8×10^{10}	2.3×10^9	115.5	52.8	13.5
	Norm	114.8	61.0	**6.0**	3.3×10^9	1.0×10^7	1.2×10^5	114.9	61.0	**6.1**
	Exp	117.1	63.6	13.2	2.0×10^9	6.8×10^6	3.6×10^5	117.3	63.6	14.8
20%	Gam	115.5	52.8	13.2	2.0×10^9	3.8×10^{10}	2.3×10^9	115.5	52.8	13.4
	Norm	114.8	61.0	**6.0**	3.3×10^9	1.0×10^7	1.2×10^5	114.8	61.0	**6.1**
	Exp	117.1	63.6	13.1	2.0×10^9	6.8×10^6	4.0×10^5	117.2	63.6	22.0
25%	Gam	115.5	52.8	13.2	2.0×10^9	3.8×10^{10}	2.3×10^9	115.5	52.8	13.4
	Norm	114.8	61.0	**6.0**	3.3×10^9	1.0×10^7	1.2×10^5	114.8	61.0	**6.1**
	Exp	117.1	63.6	13.3	2.0×10^9	6.8×10^6	4.0×10^5	117.2	63.6	13.7

Table 3: $\overline{\%ARB}$, $\overline{MSE}$ and $\overline{\%CV}$ for Gamma (5,10), Norm (5,1) and Exp (1.5) for 89% Response Rate.

.

N/B: The R-code is on request