Available online at www.isas.org.in/jisas
JOURNAL OF THE INDIAN SOCIETY OF
AGRICULTURAL STATISTICS 74(1) 2020 23–32

Product Type Calibration Estimators of Finite Population Total under Two Stage Sampling

Ankur Biswas, Kaustav Aditya, U.C. Sud and Pradip Basak

ICAR-Indian Agricultural Statistics Research Institute, New Delhi

Received 19 December 2018; Revised 10 October 2019; Accepted 18 October 2019

SUMMARY

The Calibration Approach proposed by Deville and Särndal (1992) is a popular technique to efficiently use auxiliary information in survey sampling. In this study, calibration estimators of the finite population total have been developed under two stage sampling design along with variance of the estimator and the corresponding estimator of variance. It is assumed that the population level complex auxiliary information is available at the second stage of selection and the study variable is inversely related to the available auxiliary information. The proposed calibration estimators were evaluated through a simulation study and it was found that all the proposed product type calibration estimators perform better than the Horvitz-Thompson estimator as well as usual product estimator of the population total under two stage sampling design.

Keywords: Calibration approach; Auxiliary information; Product estimator; Simulation.

1. INTRODUCTION

In sample surveys, auxiliary information on the finite population is often used to increase the precision of estimators of unknown finite population parameters of study variable. In the simplest settings, ratio and regression estimators incorporate known finite population parameters of auxiliary variables in estimation of study variable parameters. The Calibration Approach, proposed by Deville and Särndal (1992), is one of the widely used techniques for incorporation of auxiliary information in estimation stages of survey sampling. In fact, the generalized regression estimator (GREG) (Cassel *et al.*, 1976) is a special case of the calibration estimator choosing the Chi-square distance function (Deville and Särndal, 1992). Calibration technique implies that a set of initial weights (usually the sampling design weights) are transformed into a set of new weights, called calibrated weights, which is the product of its initial weight and a calibration factor. In the past few decades, calibration estimation has gained significant attention not only in the field of survey methodology, but also in survey practice. Following Deville and Särndal (1992), a lot work has been carried out in calibration estimation i.e. Singh *et al.* (1998, 1999), Wu and Sitter (2001), Sitter and Wu (2002), Kott (2006) etc. Kim and Park (2010) and Särndal (2007) provided comprehensive review on calibration approach.

In many medium to large scale surveys, it is very often the case that the sampling frame is often unavailable or it could be too expensive to construct one. Also, the population could be spread over a wide area entailing very high operational expenses for personal interviews and supervisions. Two stage sampling serves as a solution in such situations where groups of elements, called primary stage units (PSU), are selected first and, then, a sample of elements, called secondary stage units (SSU), are selected from each selected PSU. For example, in agricultural surveys, villages can be selected as PSU and farmers can be selected as SSU. Estimation of the population parameters in two stage sampling using auxiliary information has been well addressed in survey sampling. Sukhatme *et al.* (1984) suggested regression estimator

Corresponding author: Ankur Biswas
E-mail address: ankur.biswas@icar.gov.in

of the population mean in two-stage sampling. Särndal *et al.* (1992) considered three different situations with respect to availability of complex auxiliary variable under two stage sampling and discussed extensively on ratio and regression estimators under such situations. Aditya *et al.* (2016a, 2016b) and Mourya *et al.* (2016) extended the calibration estimation under different cases of availability of complex auxiliary information under two stage sampling. Sinha *et al.* (2016) proposed calibration estimators for estimating population mean under stratified sampling and stratified double sampling. Aditya *et al.* (2017) attempted to use calibration approach for estimation of crop yields at the district level under two-stage sampling. Basak *et al.* (2017) proposed a calibration estimator of finite population regression coefficient under two-stage sampling design. Veronica *et al.* (2018) considered computation of calibration weights at both the first and second stages of sample selection for estimation of population mean by assuming the population means of auxiliary variables are known at both the stages of sample selection under equal probability two-stage sampling.

It was observed that most of the work related to calibration estimation for the finite population parameters were mostly restricted with the assumption of linear relationship between the study variable and the auxiliary variable. There may be situations when it can be seen that the study variable is inversely related to the auxiliary variable. For instance, an inverse relation, generally, exists between the age of individuals and hours of sleep (Sud *et al.*, 2014a). Again, in household surveys, it is often the case that marketable surplus is inversely related to family consumption for seed, feed etc. In these situations, the product estimator, proposed by Murthy (1964), is a feasible alternative. In that situation the usual methodology for calibration estimation may not fit in. Sud *et al.* (2014a, 2014b) studied the calibration approach for estimation of population total when variable of interest and auxiliary information have inverse relation under uni-stage equal probability sampling. However, multi-stage designs are most prevalent in medium to large scale surveys. Therefore, in this present study, an attempt has been made to develop calibration estimators of finite population total under two stage sampling when study variable is inversely related to the auxiliary variable.

In Section 2, proposed product type calibration estimators of finite population total under two stage sampling has been discussed. In order to study the statistical properties of proposed estimators empirically, a simulation study was carried out. Details of simulation study and discussion on simulation results are given in Section 3 and 4 respectively. Section 5 comprises concluding remarks.

2. PROPOSED CALIBRATION ESTIMATORS UNDER TWO STAGE SAMPLING DESIGN

In this section, two different calibration estimates are proposed under two stage sampling design under the assumption that available auxiliary information is inversely related to the study variable. The proposed estimators were developed with the assumption of availability of auxiliary information at SSU level under two stage sampling. Let, U be the finite population under consideration and Y be the character under study. U is grouped into N different PSUs such that $U_I = \{1,...,i,...,N\}$ and i^{th} PSU consists of M_i SSUs such that $U_i = \{1,...,k,...,M_i\}$, $i \in U_I$. Thus, we have $U = \bigcup_{i=1}^{N} U_i$ and total number of SSUs in the population U is $M_0 = \sum_{i=1}^{N} M_i$. Under two stage sampling, at stage one, a sample of PSUs, s_I, of size n PSUs is selected from U_I according to a specified design $p_I(\cdot)$ with $\pi_{Ii} = P(i \in s_I)$ and $\pi_{Iij} = P(i, j \in s_I)$ as the inclusion probabilities at the PSU level. Given that the PSU U_i is selected at the first stage, a sample s_i of size m_i SSUs is drawn from U_i according to some specified design $p_i(\cdot)$ with inclusion probabilities $\pi_{k/i} = P(k \in s_i / i \in s_I)$ and $\pi_{kl/i} = P(k, l \in s_i / i \in s_I)$ at the SSU level. In the second stage of sampling, invariance and independence property is followed. The entire sample of elements is defined as, $s = \bigcup_{i=1}^{s_I} s_i$. Let, y_{ik} denotes the observation of the study variable from k^{th} SSU in i^{th} PSU and it is observed for all $k \in s$. The parameter of interest is the population total $t_y = \sum_{i=1}^{N}\sum_{k=1}^{M_i} y_{ik} = \sum_{i=1}^{N} t_{yi}$, where $t_{yi} = \sum_{k=1}^{M_i} y_{ik} = i^{th}$ PSU total. An attempt has been made to improve the ordinary Horvitz-Thompson (1952) estimator for population total as given by

$$\hat{t}_{y\pi} = \sum_{i=1}^{n} a_{Ii} \sum_{k=1}^{m_i} \left(a_{k/i} y_{ik}\right) = \sum_{i=1}^{n}\sum_{k=1}^{m_i} a_{ik} y_{ik} \tag{2.1}$$

where, the design weights are given as

$a_{Ii} = 1/\pi_{Ii}$, $\forall i \in s_I$, $a_{k/i} = 1/\pi_{k/i}$, $\forall k \in s_i$, $i \in s_I$ and $a_{ik} = a_{Ii}.a_{k/i}$.

Under two stage sampling design, the complex auxiliary information may be available for the PSUs as well as the SSUs within the PSUs (Särndal *et al.*, 1992). In the present study, as per availability of complex auxiliary information at the ultimate stage units following two cases have been considered under two stage sampling design

Case 1: Population level complete auxiliary information is available at the SSU level.

Case 2: Population level auxiliary information is available only for the selected PSUs.

2.1 Case 1: Population level complete auxiliary information is available at SSU level

Under this case, it has been assumed that population level complete auxiliary information is available at the unit (SSU) level i.e. the auxiliary information of k^{th} SSU in i^{th} PSU, x_{ik}, is known for all elements $k \in U$. A correct value of $\sum_{i=1}^{N}\sum_{k=1}^{M_i} x_{ik}^{-1}$ is assumed to be known. In addition, there exist an inverse relationship between the study variable *Y* and the auxiliary variable *X*.

Using the well-known Calibration Approach (Deville and Särndal, 1992), we wish to modify the total design weight of k^{th} SSU of i^{th} PSU, i.e. $a_{ik} = a_{Ii}.a_{k/i}$, as given in the HT estimator of population total in Equation 2.1. The proposed product type calibration estimator of population total under **Case 1** is given by

$$\hat{t}_{yCP1} = \sum_{i=1}^{n}\sum_{k=1}^{m_i} w_{1ik} y_{ik} \tag{2.1.1}$$

where, w_{1ik} is the calibrated weight corresponding to the design weight a_{ik} under **Case 1**.

In order to obtain the calibrated weight w_{1ik}, we minimized the Chi-square type distance $\sum_{i=1}^{n}\sum_{k=1}^{m_i} \frac{\left(w_{1ik} - a_{ik}\right)^2}{a_{ik} q_{ik}}$ subject to the calibration constraint $\sum_{i=1}^{n}\sum_{k=1}^{m_i} w_{1ik} x_{ik}^{-1} = \sum_{i=1}^{N}\sum_{k=1}^{M_i} x_{ik}^{-1}$, where q_{ik} are suitably chosen constants. Using the method of Lagrange multiplier, by minimizing

$$\varphi\left(w_{1ik}, \lambda\right) = \sum_{i=1}^{n}\sum_{k=1}^{m_i} \frac{\left(w_{1ik} - a_{ik}\right)^2}{a_{ik} q_{ik}} - \lambda\left(\sum_{i=1}^{n}\sum_{k=1}^{m_i} w_{1ik} x_{ik}^{-1} - \sum_{i=1}^{N}\sum_{k=1}^{M_i} x_{ik}^{-1}\right)$$

the calibrated weights are obtained as given by

$$w_{1ik} = a_{ik} + a_{ik} q_{ik} x_{ik}^{-1} \left[\frac{\sum_{i=1}^{N}\sum_{k=1}^{M_i} x_{ik}^{-1} - \sum_{i=1}^{n}\sum_{k=1}^{m_i} a_{ik} x_{ik}^{-1}}{\sum_{i=1}^{n}\sum_{k=1}^{m_i} a_{ik} q_{ik} x_{ik}^{-2}}\right],$$

$$\forall k = 1, 2, ..., m_i \text{ and } \forall i = 1, 2, .., n \tag{2.1.2}$$

Using the results of the Equation (2.1.2) in (2.1.1) considering $q_{ik} = x_{ik}$, we have therefore proved the following result.

Theorem 1: Under **Case 1** of two stage sampling, if we consider the calibrated design weights as $w_{1ik} = a_{ik}\left(\sum_{i=1}^{N}\sum_{k=1}^{M_i} x_{ik}^{-1} \Big/ \sum_{i=1}^{n}\sum_{k=1}^{m_i} a_{ik} x_{ik}^{-1}\right)$, $\forall k = 1,2,...,m_i$, then the proposed product type calibration estimator of population total is given as

$$\hat{t}_{yCP1} = \sum_{i=1}^{n}\sum_{k=1}^{m_i} w_{1ik} y_{ik}$$

$$= \left(\sum_{i=1}^{n}\sum_{k=1}^{m_i} a_{ik} y_{ik}\right)\left(\sum_{i=1}^{N}\sum_{k=1}^{M_i} x_{ik}^{-1}\right) \Bigg/ \left(\sum_{i=1}^{n}\sum_{k=1}^{m_i} a_{ik} x_{ik}^{-1}\right). \tag{2.1.3}$$

Corollary 1: Under an equal probability without replacement sampling design (Simple Random Sampling without replacement (SRSWOR)) at both the stages of two stage sampling, the proposed product type calibration estimator under **Case 1** reduces to

$$\hat{t}_{yCP1} = \left(\frac{N}{n}\sum_{i=1}^{n}\frac{M_i}{m_i}\sum_{k=1}^{m_i} y_{ik}\right)\left(\sum_{i=1}^{N}\sum_{k=1}^{M_i} x_{ik}^{-1}\right) \Bigg/ \left(\frac{N}{n}\sum_{i=1}^{n}\frac{M_i}{m_i}\sum_{k=1}^{m_i} x_{ik}^{-1}\right) \tag{2.1.4}$$

The theoretical bias of the proposed product type calibration estimator $\hat{t}_{yCP1}$ is obtained through Taylor series linearization technique as

$$Bias\left(\hat{t}_{yCP1}\right)=\frac{\sum_{i=1}^{N}\sum_{k=1}^{M_i}y_{ik}}{\sum_{i=1}^{N}\sum_{k=1}^{M_i}x_{ik}^{-1}}\left[\frac{Cov\left(\sum_{i=1}^{n}\sum_{k=1}^{m_i}a_{ik}y_{ik},\sum_{i=1}^{n}\sum_{k=1}^{m_i}a_{ik}x_{ik}^{-1}\right)}{\sum_{k=1}^{N}\sum_{k=1}^{M_i}y_{ik}}+\frac{V\left(\sum_{i=1}^{n}\sum_{k=1}^{m_i}a_{ik}x_{ik}^{-1}\right)}{\sum_{i=1}^{N}\sum_{k=1}^{M_i}x_{ik}^{-1}}\right] \quad (2.1.5)$$

Under SRSWOR at both the stages we obtain the bias using Taylor series linearization as given by

$$Bias(\hat{t}_{yCP1})=t_y\left[\left(\frac{1}{n}-\frac{1}{N}\right)\left(\rho_b C_{by}C_{bx}+C_{bx}^2\right)+\frac{1}{n}\left(\bar{\rho}_w C_{wy}C_{wx}+C_{wx}^2\right)\right] \quad (2.1.6)$$

where,

$$\rho_b=\frac{S_{bxy}}{S_{bx}S_{by}},\quad C_{by}^2=\frac{S_{by}^2}{\bar{Y}_{N.}^2},\quad C_{bx}^2=\frac{S_{bx}^2}{\bar{X}_{N.}^2},$$

$$\bar{\rho}_w=\frac{\bar{S}_{wxy}}{\bar{S}_{wx}\bar{S}_{wy}},\quad C_{wy}^2=\frac{\bar{S}_{wy}^2}{\bar{Y}_{N.}^2},\quad C_{wx}^2=\frac{\bar{S}_{wx}^2}{\bar{X}_{N.}^2},$$

$$\bar{X}_{N.}=\frac{1}{N}\sum_{i=1}^{N}M_i\bar{X}_{i.},\quad \bar{Y}_{N.}=\frac{1}{N}\sum_{i=1}^{N}M_i\bar{Y}_{i.},$$

$$\bar{X}_{i.}=\frac{1}{M_i}\sum_{k=1}^{M_i}x_{ik},\quad \bar{Y}_{i.}=\frac{1}{M_i}\sum_{k=1}^{M_i}y_{ik},$$

$$S_{bxy}=\frac{1}{N-1}\sum_{i=1}^{N}\left(M_i\bar{Y}_{i.}-\bar{Y}_{N.}\right)\left(M_i\bar{X}_{i.}-\bar{X}_{N.}\right),$$

$$S_{by}^2=\frac{1}{N-1}\sum_{i=1}^{N}\left(M_i\bar{Y}_{i.}-\bar{Y}_{N.}\right)^2,$$

$$S_{bx}^2=\frac{1}{N-1}\sum_{i=1}^{N}\left(M_i\bar{X}_{i.}-\bar{X}_{N.}\right)^2,$$

$$\bar{S}_{wxy}=\frac{1}{N}\sum_{i=1}^{N}M_i^2\left(\frac{1}{m_i}-\frac{1}{M_i}\right)S_{ixy},$$

$$\bar{S}_{wy}^2=\frac{1}{N}\sum_{i=1}^{N}M_i^2\left(\frac{1}{m_i}-\frac{1}{M_i}\right)S_{iy}^2,$$

$$\bar{S}_{wx}^2=\frac{1}{N}\sum_{i=1}^{N}M_i^2\left(\frac{1}{m_i}-\frac{1}{M_i}\right)S_{ix}^2,$$

$$S_{ixy}=\frac{1}{M_i-1}\sum_{k=1}^{M_i}\left(x_{ik}-\bar{X}_{i.}\right)\left(y_{ik}-\bar{Y}_{i.}\right),$$

$$S_{iy}^2=\frac{1}{M_i-1}\sum_{k=1}^{M_i}\left(y_{ik}-\bar{Y}_{i.}\right)^2$$

$$S_{ix}^2=\frac{1}{M_i-1}\sum_{k=1}^{M_i}\left(x_{ik}-\bar{X}_{i.}\right)^2.$$

Usual product estimator under **Case 1** of two stage sampling considering SRSWOR at both the stages is given by

$$\hat{t}_{yP1}=\left(\frac{N}{n}\sum_{i=1}^{n}\frac{M_i}{m_i}\sum_{k=1}^{m_i}y_{ik}\right)\left(\frac{N}{n}\sum_{i=1}^{n}\frac{M_i}{m_i}\sum_{k=1}^{m_i}x_{ik}\right)\Bigg/\left(\sum_{k=1}^{N}\sum_{k=1}^{M_i}x_{ik}\right) \quad (2.1.7)$$

and, its bias is given as

$$Bias(\hat{t}_{yP1})=t_y\left[\left(\frac{1}{n}-\frac{1}{N}\right)\rho_b C_{by}C_{bx}+\frac{1}{n}\bar{\rho}_w C_{wy}C_{wx}\right] \quad (2.1.8)$$

It has been found that under SRSWOR at both the stages of a two stage sampling design under **Case 1**, product estimator $\left(\hat{t}_{yP1}\right)$ is better than usual HT estimator $\left(\hat{t}_{y\pi}\right)$ if $\rho_b\frac{C_{by}}{C_{bx}}<-\frac{1}{2}$ and $\bar{\rho}_w\frac{C_{wy}}{C_{wx}}<-\frac{1}{2}$. Under these conditions in two stage sampling design, it can be seen that

$$\left|Bias(\hat{t}_{yCP1})\right|\le\left|Bias(\hat{t}_{yP1})\right|.$$

Following Deville and Särndal (1992) and Särndal *et al.* (1992), the approximate variance of the proposed product type calibration estimator under **Case 1** by first order Taylor series linearization technique was obtained as

$$AV(\hat{t}_{yCP1})=\sum_{i=1}^{N}\sum_{j=1}^{N}\Delta_{Iij}\frac{t_{E_{i1}}}{\pi_{Ii}}\frac{t_{E_{j1}}}{\pi_{Ij}}+\sum_{i=1}^{N}\frac{1}{\pi_{Ii}}\sum_{k=1}^{M_i}\sum_{l=1}^{M_i}\Delta_{kl/i}\frac{E_{k/i}}{\pi_{k/i}}\frac{E_{l/i}}{\pi_{l/i}}, \quad (2.1.9)$$

where,

$$t_{E_{i1}}=\sum_{k=1}^{M_i}E_{k/i},\quad E_{k/i}=y_{ik}-\left(\sum_{i=1}^{N}\sum_{k=1}^{M_i}y_{ik}\Bigg/\sum_{i=1}^{N}\sum_{k=1}^{M_i}x_{ik}^{-1}\right)x_{ik}^{-1},$$

$$\Delta_{Iij}=(\pi_{Iij}-\pi_{Ii}\pi_{Ij}),\quad \Delta_{kl/i}=\pi_{kl/i}-\pi_{k/i}\pi_{l/i}.$$

Under SRSWOR design at both the stages the approximate variance reduces to

$$AV(\hat{t}_{yCP1}) = N^2\left(\frac{1}{n}-\frac{1}{N}\right)\left(S_{by}^2 + R_1^2 S_{bx^{-1}}^2 - 2R_1 S_{byx^{-1}}\right) + \frac{N}{n}\sum_{i=1}^{N} M_i^2\left(\frac{1}{m_i}-\frac{1}{M_i}\right)\left(S_{iy}^2 + R_1^2 S_{ix^{-1}}^2 - 2R_1 S_{iyx^{-1}}\right) \quad (2.1.10)$$

where,

$$R_1 = \left(\sum_{i=1}^{N}\sum_{k=1}^{M_i} y_{ik}\right) \Bigg/ \left(\sum_{i=1}^{N}\sum_{k=1}^{M_i} x_{ik}^{-1}\right),$$

$$\bar{X}_{(-1)i.} = \frac{1}{M_i}\sum_{k=1}^{M_i} x_{ik}^{-1}, \quad \bar{X}_{(-1)\mathrm{N.}} = \frac{1}{N}\sum_{i=1}^{N} M_i \bar{X}_{(-1)i.},$$

$$S_{byx^{-1}} = \frac{1}{N-1}\sum_{i=1}^{N}\left(M_i\bar{Y}_{i.} - \bar{Y}_{N.}\right)\left(M_i\bar{X}_{(-1)i.} - \bar{X}_{(-1)\mathrm{N.}}\right),$$

$$S_{bx^{-1}}^2 = \frac{1}{N-1}\sum_{i=1}^{N}\left(M_i\bar{X}_{(-1)i.} - \bar{X}_{(-1)\mathrm{N.}}\right)^2,$$

$$S_{iyx^{-1}} = \frac{1}{M_i-1}\sum_{i=1}^{M_i}\left(y_{ik} - \bar{Y}_{i.}\right)\left(x_{ik}^{-1} - \bar{X}_{(-1)i.}\right),$$

$$S_{ix^{-1}}^2 = \frac{1}{M_i-1}\sum_{i=1}^{M_i}\left(x_{ik}^{-1} - \bar{X}_{(-1)i.}\right)^2.$$

Following Särndal *et al.* (1992), the Yates–Grundy form of estimator of variance (Yates and Grundy, 1953) of the proposed product type calibration estimator under **Case 1** is given by

$$\hat{V}_{YG}(\hat{t}_{yCP1}) = \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} d_{Iij}\left(\frac{\hat{t}_{E_{i1}}}{\pi_{Ii}} - \frac{\hat{t}_{E_{j1}}}{\pi_{Ij}}\right)^2 + \frac{1}{2}\sum_{j=1}^{n}\frac{1}{\pi_{Ii}}\sum_{k=1}^{m_i}\sum_{l=1}^{m_i} d_{kl/i}(w_{1ik}e_{k/i} - w_{1il}e_{l/i})^2 \quad (2.1.11)$$

where,

$$\hat{t}_{E_{i1}} = \sum_{k=1}^{m_i}\frac{e_{k/i}}{\pi_{k/i}},$$

$$e_{k\ i} = y_{ik} - \left(\sum_{i=1}^{n}\sum_{k=1}^{m_i} a_{ik}y_{ik} \Bigg/ \sum_{i=1}^{n}\sum_{k=1}^{m_i} a_{ik}x_{ik}^{-1}\right)x_{ik}^{-1},$$

$$d_{Iij} = \frac{(\pi_{Ii}\pi_{Ij} - \pi_{Iij})}{\pi_{Iij}} \text{ and } d_{kl/i} = \frac{\pi_{k/i}\pi_{l/i} - \pi_{kl/i}}{\pi_{kl/i}}.$$

Under SRSWOR design at both the stages the estimator of variance reduces to

$$\hat{V}(\hat{t}_{yCP1}) = N^2\left(\frac{1}{n}-\frac{1}{N}\right)\left(\hat{S}_{by}^2 + \hat{R}_1^2 \hat{S}_{bx^{-1}}^2 - 2\hat{R}_1 \hat{S}_{byx^{-1}}\right) + \frac{N}{n}\sum_{i=1}^{n} M_i^2\left(\frac{1}{m_i}-\frac{1}{M_i}\right)\left(\hat{S}_{iy}^2 + \hat{R}_1^2 \hat{S}_{ix^{-1}}^2 - 2\hat{R}_1 \hat{S}_{iyx^{-1}}\right) \quad (2.1.12)$$

where,

$$\hat{R}_1 = \left(\frac{N}{n}\sum_{i=1}^{n}\frac{M_i}{m_i}\sum_{k=1}^{m_i} y_{ik}\right) \Bigg/ \left(\frac{N}{n}\sum_{i=1}^{n}\frac{M_i}{m_i}\sum_{k=1}^{m_i} x_{ik}^{-1}\right),$$

$$\hat{S}_{bx^{-1}}^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(M_i\bar{x}_{(-1)i.} - \bar{x}_{(-1)n.}\right)^2,$$

$$\hat{S}_{by}^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(M_i\bar{y}_{i.} - \bar{y}_{n.}\right)^2,$$

$$\hat{S}_{byx^{-1}} = \frac{1}{n-1}\sum_{i=1}^{n}\left(M_i\bar{y}_{i.} - \bar{y}_{n.}\right)\left(M_i\bar{x}_{(-1)i.} - \bar{x}_{(-1)n.}\right),$$

$$\hat{S}_{iy}^2 = \frac{1}{m_i-1}\sum_{i=1}^{m_i}\left(y_{ik} - \bar{y}_{i.}\right)^2,$$

$$\hat{S}_{ix^{-1}}^2 = \frac{1}{m_i-1}\sum_{i=1}^{m_i}\left(x_{ik}^{-1} - \bar{x}_{(-1)i.}\right)^2,$$

$$\hat{S}_{iyx^{-1}} = \frac{1}{m_i-1}\sum_{i=1}^{m_i}\left(y_{ik} - \bar{y}_{i.}\right)\left(x_{ik}^{-1} - \bar{x}_{(-1)i.}\right),$$

$$\bar{y}_{i.} = \frac{1}{m_i}\sum_{k=1}^{m_i} y_{ik}, \quad \bar{y}_{n.} = \frac{1}{n}\sum_{i=1}^{n} M_i\bar{y}_{i.}, \quad \bar{x}_{(-1)i.} = \frac{1}{m_i}\sum_{k=1}^{m_i} x_{ik}^{-1},$$

$$\bar{x}_{(-1)n.} = \frac{1}{n}\sum_{i=1}^{n} M_i\bar{x}_{(-1)i}.$$

2.2 Case 2: Population level auxiliary information is available only for the selected PSUs

In this case, it has been assumed that the population level auxiliary information is available at the SSU level only for the selected PSUs i.e. the auxiliary information is known for all the SSUs within the PSU $i \in s_I$. The correct value of $\sum_{k=1}^{M_i} x_{ik}^{-1}$ is assumed to be available for each i^{th} sampled PSU. Suppose, there exist inverse relationship between the study variable Y and the auxiliary variable X. Using well-known Calibration Approach (Deville and Särndal, 1992), the design weight at the second stage $a_{k/i}$ has been revised. The proposed product type calibration estimator of population total under **Case 2** is given by

$$\hat{t}_{yCP2} = \sum_{i=1}^{n} a_{Ii} \sum_{k=1}^{m_i} w_{2ik}\, y_{ik} \quad (2.2.1)$$

where, w_{2ik} is the calibrated weight corresponding to the design weight $a_{k/i}$.

In this situation, we minimized the Chi-square type distance function $\sum_{k=1}^{m_i} \frac{(w_{2ik} - a_{k/i})^2}{a_{k/i}\, q_{ik}}$ subject to $\sum_{k=1}^{m_i} w_{2ik}\, x_{ik}^{-1} = \sum_{k=1}^{M_i} x_{ik}^{-1}$, where q_{ik} are suitably chosen constants. Using Lagrange multiplier technique, by minimizing

$$\varphi(w_{2ik}, \lambda) = \sum_{k=1}^{m_i} \frac{(w_{2ik} - a_{k/i})^2}{a_{k/i}\, q_{ik}} - \lambda \left(\sum_{k=1}^{m_i} w_{2ik}\, x_{ik}^{-1} - \sum_{k=1}^{M_i} x_{ik}^{-1} \right),$$

the new set of calibrated weights is obtained as

$$w_{2ik} = a_{k/i} + a_{k/i}\, q_{ik} x_{ik}^{-1} \left[\frac{\sum_{k=1}^{M_i} x_{ik}^{-1} - \sum_{k=1}^{m_i} a_{k/i}\, x_{ik}^{-1}}{\sum_{k=1}^{m_i} a_{k/i}\, q_{ik} x_{ik}^{-2}} \right]$$

$$\forall\, k = 1, 2, \ldots, m_i. \quad (2.2.2)$$

Using the results of the Equation (2.2.2) in (2.2.1) considering $q_{ik} = x_{ik}$, we have therefore proved the following result.

Theorem 2: Under **Case 2** of two stage sampling, if we consider the calibrated design weights as

$$w_{2ik} = a_{k/i} \left(\sum_{k=1}^{M_i} x_{ik}^{-1} \bigg/ \sum_{k=1}^{m_i} a_{k/i}\, x_{ik}^{-1} \right), \qquad \forall\, k = 1, 2, \ldots, m_i$$

then the proposed product type calibration estimator of population total is given as

$$\hat{t}_{yCP2} = \sum_{i=1}^{n} a_{Ii} \sum_{k=1}^{m_i} w_{2ik}\, y_{ik}$$

$$= \sum_{i=1}^{n} a_{Ii} \frac{\left(\sum_{k=1}^{m_i} a_{k/i}\, y_{ik} \right) \left(\sum_{i=1}^{M_i} x_{ik}^{-1} \right)}{\left(\sum_{i=1}^{m_i} a_{k/i}\, x_{ik}^{-1} \right)}. \quad (2.2.3)$$

Corollary 2: Under SRSWOR at both the stages of two stage sampling, the proposed product type calibration estimator under **Case 2** reduces to

$$\hat{t}_{yCP2} = \frac{N}{n} \sum_{i=1}^{n} \left[\left(\frac{M_i}{m_i} \sum_{k=1}^{m_i} y_{ik} \right) \left(\sum_{i=1}^{M_i} x_{ik}^{-1} \right) \bigg/ \left(\frac{M_i}{m_i} \sum_{k=1}^{m_i} x_{ik}^{-1} \right) \right]. \quad (2.2.4)$$

Using Taylor series linearization technique, its bias is obtained as

$$Bias(\hat{t}_{yCP2}) = t_y \left[\frac{1}{n} \left(\bar{\rho}_w C_{wy} C_{wx} + C_{wx}^2 \right) \right] \quad (2.2.5)$$

where, the terms are as defined in **Case 1** (Eqn. 2.1.6).

Usual product estimator under **Case 2** of two stage sampling with SRSWOR at both the stages is given by

$$\hat{t}_{yP2} = \left(\frac{N}{n} \sum_{i=1}^{n} \frac{M_i}{m_i} \sum_{k=1}^{m_i} y_{ik} \right) \left(\frac{N}{n} \sum_{i=1}^{n} \frac{M_i}{m_i} \sum_{k=1}^{m_i} x_{ik} \right) \bigg/ \left(\frac{N}{n} \sum_{i=1}^{n} \sum_{k=1}^{M_i} x_{ik} \right) \quad (2.2.6)$$

and, using Taylor series linearization technique, its bias is given by

$$Bias(\hat{t}_{yP2}) = t_y \left[\frac{1}{n} \bar{\rho}_w C_{wy} C_{wx} \right] \quad (2.2.7)$$

It has been found that under SRSWOR at both the stages of two stage sampling design under **Case 2**, product estimator $(\hat{t}_{yP2})$ is better than usual HT estimator $(\hat{t}_{y\pi})$ if $\bar{\rho}_w \frac{C_{wy}}{C_{wx}} < -\frac{1}{2}$. Under this condition in two stage sampling design, it can be seen that

$$\left| Bias(\hat{t}_{yCP2}) \right| \le \left| Bias(\hat{t}_{yP2}) \right|.$$

Following Särndal *et al.* (1992) the approximate variance of the proposed product type calibration estimator under **Case 2** by first order Taylor series linearization technique was obtained as

$$AV(\hat{t}_{yCP2}) = \sum_{i=1}^{N}\sum_{j=1}^{N} \Delta_{Iij} \frac{t_{E_{i2}}}{\pi_{Ii}} \frac{t_{E_{j2}}}{\pi_{Ij}} + \sum_{i=1}^{N} \frac{1}{\pi_{Ii}} \sum_{k=1}^{M_i}\sum_{l=1}^{M_i} \Delta_{kl/i} \frac{E_{k/i}}{\pi_{k/i}} \frac{E_{l/i}}{\pi_{l/i}}, \quad (2.2.8)$$

where,

$$t_{E_{i2}} = \sum_{k=1}^{M_i} E_{k/i}\ ,\ E_{k/i} = y_{ik} - R_i x_{ik}^{-1}, .$$

$$R_i = \left(\sum_{i=1}^{M_i} y_{ik} \Big/ \sum_{i=1}^{M_i} x_{ik}^{-1} \right).$$

Under SRSWOR design at both the stages, it reduces to

$$AV(\hat{t}_{yCP2}) = N^2 \left(\frac{1}{n} - \frac{1}{N} \right) S_{by}^2 + \frac{N}{n} \sum_{i=1}^{N} M_i^2 \left(\frac{1}{m_i} - \frac{1}{M_i} \right) \left\{ S_{iy}^2 + R_i^2 S_{ix^{-1}}^2 - 2R_i S_{iyx^{-1}} \right\} \quad (2.2.9)$$

where, the terms are as defined in **Case 1** (Eqn. 2.1.6 and 2.1.10).

The Yates–Grundy form of estimator of variance of the proposed product type calibration estimator under **Case 2** is given by

$$\hat{V}_{YG}(\hat{t}_{yCP2}) = \frac{1}{2} \sum_{i=1}^{n}\sum_{j=1}^{n} d_{Iij} \left(\frac{\hat{t}_{E_{i2}}}{\pi_{Ii}} - \frac{\hat{t}_{E_{j2}}}{\pi_{Ij}} \right)^2 + \frac{1}{2} \sum_{j=1}^{n} \frac{1}{\pi_{Ii}} \sum_{k=1}^{m_i}\sum_{l=1}^{m_i} d_{kl/i} (w_{2ik} e_{k/i} - w_{2il} e_{l/i})^2 \quad (2.2.10)$$

where,

$$\hat{t}_{E_{i2}} = \sum_{k=1}^{n_i} \frac{e_{k/i}}{\pi_{k/i}}, ,\ e_{k/i} = y_{ik} - \hat{R}_i x_{ik}^{-1},$$

$$\hat{R}_i = \left(\sum_{k=1}^{n_i} a_{k/i} y_{ik} \right) \Bigg/ \left(\sum_{i=1}^{n_i} a_{k/i} x_{ik}^{-1} \right),$$

$d_{Iij} = (\pi_{Ii}\pi_{Ij} - \pi_{Iij}) / \pi_{Iij}$ and

$d_{kl/i} = \left(\pi_{k/i}\pi_{l/i} - \pi_{kl/i} \right) / \pi_{kl/i}$.

Under SRSWOR design at both the stages it reduces to

$$\hat{V}(\hat{t}_{yCP2}) = N^2 \left(\frac{1}{n} - \frac{1}{N} \right) \hat{S}_{by}^2 + \frac{N}{n} \sum_{i=1}^{n} M_i^2 \left(\frac{1}{m_i} - \frac{1}{M_i} \right) \left\{ \hat{S}_{iy}^2 + \hat{R}_i^2 \hat{S}_{ix^{-1}}^2 - 2\hat{R}_i \hat{S}_{iyx^{-1}} \right\} \quad (2.2.11)$$

where, the terms are as defined in **Case 1** (Eqn. 2.1.12).

3. SIMULATION STUDY

In order to evaluate the statistical performance of proposed product type calibration estimators, a simulation study was carried out. We have considered the case of two stage sampling where sample selection at each stage is governed by SRSWOR for the situation that the size of the PSU and the corresponding SSUs were fixed. For the simulation study, a finite population of 5000 units considering, number of PSU, N=50 and PSU size, M_i=100, was generated from $y_k = \beta x_k^{-1} + e_k$, $k = 1, ..., M_0$, where $M_0 = \sum_{i=1}^{N} M_i$. The auxiliary variable was generated from normal distribution with mean 5 and variance 1 i.e. $x_k \sim N(5,\ 1)$ and the errors, e_k, $k = 1,...,M_0$, from normal distribution with mean 0 and variance $\sigma^2 x_k^{-1}$ i.e. $e_k \sim N(0,\ \sigma^2 x_k^{-1})$. We have fixed the value of β = 20 and chosen four different values for σ^2 as 0.25, 1.0, 2.0 and 5.0. In this way, we generated four sets of population, denoted as **Set 1**, **Set 2**, **Set 3** and **Set 4**, with different correlation coefficient values between study variable Y and auxiliary variable X as -0.91, -0.85, -0.78 and -0.64 respectively. The value of left hand side of the Condition 1 and Condition 2 i.e. $\rho_A \frac{C_{by}}{C_{bx}} < -\frac{1}{2}$ and $\bar{\rho}_w \frac{C_{wy}}{C_{wx}} < -\frac{1}{2}$ are lesser than -0.5 in all the population sets which can be seen in the following table:

Set	Set 1	Set 2	Set 3	Set 4
Condition 1	-1.1	-1.17	-1.23	-1.34
Condition 2	-1.11	-1.1	-1.1	-1.08

Then, from each of the study population sets, we have selected a total of 10000 different samples of following sizes using SRSWOR at both the stages of the two stage sampling design and calculated different estimates of population total under **Case 1** and **2**:

n =10, m_i =20	n =15, m_i =25	n =20, m_i=30	n =25, m_i=40
n =10, m_i =25	n =15, m_i =30	n =20, m_i=40	n =25, m_i=50

Developed product type calibration estimators as well as all other usual estimators of population total under two stage sampling were evaluated on the basis of two measures viz. percentage Relative Bias (%RB) and percentage Relative Root Mean Squared Error (%RRMSE) of any estimator of the population parameter θ as given by

$$RB(\hat{\theta}) = \frac{1}{S}\sum_{i=1}^{S}\left(\frac{\hat{\theta}_i - \theta}{\theta}\right)\times 100 \text{ and}$$

$$RRMSE(\hat{\theta}) = \sqrt{\frac{1}{S}\sum_{i=1}^{S}\left(\frac{\hat{\theta}_i - \theta}{\theta}\right)^2}\times 100$$

where, $\hat{\theta}_i$ are the estimates of population parameter θ for the character under study obtained at i^{th} sample in the simulation study.

4. RESULTS AND DISCUSSION

Table 1 shows the %RB of the HT estimators ($\hat{t}_{y\pi}$), product estimators ($\hat{t}_{yP1}$ and $\hat{t}_{yP2}$), ratio estimators ($\hat{t}_{yR1}$ and $\hat{t}_{yR2}$), linear regression estimators ($\hat{t}_{ylr1}$ and $\hat{t}_{ylr2}$) (as in Särndal *et al.*, (1992), pp-323) and proposed product type calibration estimators ($\hat{t}_{yCP1}$ and $\hat{t}_{yCP2}$) of population total under both the **Case 1** and **2** respectively when available auxiliary variable is inversely related with the study variable. Table 2 presents comparison of performance of all the estimators for all the population Sets on the basis of %RRMSE.

From Table 1 it can be seen that, the proposed product type calibration estimators of the population total for both the **Case 1** and **2** of availability of auxiliary information were giving consistently least amount %RB in all the sets compared to their usual linear regression, product, ratio and HT estimators under two stage sampling design when available auxiliary variable is inversely related with the study variable. It is evident that ratio estimator is not at all suitable for this situation.

A close look of Table 2 reveals that, the product type calibration estimators of the population total developed under two stage sampling design under **Case 1** and **2** were always more efficient than the respective linear regression, product, ratio and HT estimators in all the population sets with respect to %RRMSE. The %RRMSE of both the proposed product type calibration estimators of the population total under **Case 1** and **2** were decreasing with the increase of sample sizes. With the increase of negative correlation between the study and auxiliary variable, %RRMSE of both the proposed product type calibration estimators of the population total under **Case 1** and **2** were decreasing. The proposed product type calibration estimators of the population total developed under **Case 1** of two stage sampling design was producing least %RRMSE in all Sets. Therefore, for the situations of availability of population level complete auxiliary information at SSU level i.e. **Case 1**, performance of the proposed product type calibration estimator is best among all other competitors. On the other hand, for more practical situation of availability of population level auxiliary information only for selected PSUs i.e. **Case 2**, proposed product type calibration estimator can be preferred over usual HT, product and linear regression estimators of population total.

5. CONCLUSIONS

In this study, following the calibration approach (Deville and Särndal, 1992), we proposed product type calibration estimators of the finite population total under two stage sampling design when the available auxiliary variable is inversely related to the study variable. Here, two different cases under two stage sampling viz. "***Case 1****: population level complete auxiliary information is available at the SSU level*" and "***Case 2****: population level auxiliary information is available only for the selected PSUs*" have been considered. In order to study the statistical performance of proposed product type calibration estimators as compared to existing estimators of population total of study variable, a simulation study was carried out. The simulation results show that the proposed product type calibration estimator of the population total were performing better than usual linear regression, product and HT estimators under two stage sampling design when available auxiliary variable is inversely related with the study variable. The proposed product type calibration estimators of the population total developed under **Case 1** performs better than that of **Case 2**, since more auxiliary information was available under **Case 1**.

ACKNOWLEDGEMENT

Authors are profoundly thankful to the reviewers for their productive comments for improvements.

REFERENCES

Aditya, K., Sud, U.C., Chandra, H. and Biswas, A. (2016a). Calibration Based Regression Type Estimator of the Population Total under Two Stage Sampling Design. *Journal of Indian Society of Agriculture Statistics*, **70 (1)**, 19-24.

Aditya, K., Sud, U.C. and Chandra, H. (2016b). Calibration Approach based Estimation of Finite Population Total under Two Stage Sampling. *J. Ind. Soc. Agril. Statist.*, **70(3)**, 219-226.

Aditya K., Biswas A., Gupta, A.K. and Chandra, H. (2017). District-level crop yield estimation using calibration approach. *Current Science*, **112(9)**, 1927-1931.

Basak P., Sud, U.C. and Chandra, H. (2017). Calibration Estimation of Regression Coefficient for Two-stage Sampling Design. *J. Ind. Soc. Agril. Statist.*, **71(1)**, 1-6.

Cassel, C.M., Sarndal, C.E. and Wretman, J.H. (1976). Some results on generalized difference estimation and generalized regression estimation for finite population. *Biometrika*, **63**, 615-620.

Deville, J.C. and Sarndal, C.E. (1992). Calibration estimators in survey sampling. *J. Amer. Statist. Assoc.*, **87**, 376-382.

Horvitz, D.G. and Thompson, D.J. (1952). A Generalization of Sampling without Replacement from a Finite Universe. *J. Amer. Statist. Assoc.*, **47,** 663-685.

Kim, J.K. and Park, M. (2010). Calibration estimation in survey sampling. *International Statistical Review*, **78**, 21-39.

Kott, P.S. (2006). Using calibration weighting to adjust for nonresponse and coverage errors. *Survey Methodology*, **32**, 133142.

Mourya, K.K., Sisodia, B.V.S. and Chandra, H. (2016). Calibration approach for estimating finite population parameter in two stage sampling. *Journal of Statistical Theory and Practice*, **10(3)**, 550-562.

Murthy, M.N. (1964). Product method of estimation. *Sankhya*, **26A**, 69-74.

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

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

Singh, S., Horn, S. and Yu, F. (1998). Estimation of variance of the general regression estimator: Higher level calibration approach. *Survey Methodology*, **24**, 41-50.

Singh, S., Horn, S., Choudhury, S. and Yu, F. (1999). Calibration of the estimators of variance, *Australian and New Zealand Journal of Statistics,* **41(2)**, 199-212.

Sinha N., Sisodia, B.V.S., Singh, S. and Singh, S.K. (2016). Calibration Approach Estimation of Mean in Stratified Sampling and Stratified Double Sampling, *Communications in Statistics - Theory and Methods*, **46(10)**, DOI: 10.1080/03610926.2015.1091083.

Sitter, R.R. and Wu, C. (2002). Efficient estimation of quadratic finite population functions. *J. Amer. Statist. Assoc.*, **97**, 535-543.

Sud, U.C., Chandra, H. and Gupta, V.K. (2014a). Calibration based product estimator in single and two phase sampling. *Journal of Statistical Theory and Practice*, **8(1)**, 1-11.

Sud, U.C., Chandra, H. and Gupta, V.K. (2014b). Calibration approach based regression type estimator for inverse relationship between study and auxiliary variable. *Journal of Statistical Theory and Practice*, **8(4)**, 707-721.

Sukhatme, P.V., Sukhatme, B.V., Sukhatme, S. and Asok. C. (1984). *Sampling Theory with Applications*. Indian Society of Agricultural Statistics, New Delhi.

Veronica I.S., Stephen A.S. and Sarjinder Singh. (2018). Calibrated estimators in two-stage sampling. *Communications in Statistics - Theory and Methods*, DOI: 10.1080/03610926.2018.1433850.

Wu, C. and Sitter, R.R. (2001). A model calibration approach to using complete auxiliary information from survey data. *J. Amer. Statist. Assoc.*, **96**, 185-193.

Yates, F. and Grundy, P.M. (1953). Selection without replacement from within strata with probability proportional to size. *J. Roy. Statist. Soc., **B** 9*, 223-261.

Table 1. Comparison of all the estimators under Case 1 and 2 with respect to %RB in case of all four population sets when available auxiliary variable is inversely related with the study variable

Set	Sample Size (n_m_i)	$\hat{t}_{y\pi}$	Case 1				Case 2			
			$\hat{t}_{yCP1}$	$\hat{t}_{yP1}$	$\hat{t}_{ylr1}$	$\hat{t}_{yR1}$	$\hat{t}_{yCP2}$	$\hat{t}_{yP2}$	$\hat{t}_{ylr2}$	$\hat{t}_{yR2}$
Set 1 (ρ= -0.91)	10_20	-0.016	0.001	-0.026	-0.057	0.013	-0.005	-0.029	-0.058	0.014
	10_25	0.015	-0.001	-0.011	-0.035	0.055	0.000	-0.009	-0.031	0.050
	15_25	0.000	0.002	-0.009	-0.024	0.019	0.000	-0.006	-0.020	0.015
	15_30	-0.005	0.002	-0.015	-0.026	0.012	0.002	-0.011	-0.021	0.007
	20_30	0.013	-0.001	-0.005	-0.016	0.037	0.001	-0.003	-0.013	0.034
	20_40	0.010	0.003	0.000	-0.007	0.024	0.006	0.003	-0.003	0.020
	25_40	-0.001	-0.001	-0.008	-0.013	0.009	-0.009	-0.013	-0.018	0.013
	25_50	-0.007	-0.002	-0.005	-0.008	-0.007	-0.007	-0.009	-0.012	-0.004
Set 2 (ρ= -0.85)	10_20	-0.004	0.001	-0.024	-0.054	0.034	0.006	-0.016	-0.044	0.024
	10_25	0.006	0.000	-0.018	-0.041	0.044	-0.006	-0.020	-0.042	0.042
	15_25	-0.026	-0.011	-0.025	-0.039	-0.018	-0.011	-0.026	-0.039	-0.018
	15_30	-0.015	-0.004	-0.020	-0.030	-0.002	-0.009	-0.021	-0.032	-0.002
	20_30	-0.024	-0.006	-0.019	-0.026	-0.024	-0.004	-0.017	-0.024	-0.026
	20_40	0.003	0.001	-0.008	-0.014	0.019	-0.003	-0.009	-0.015	0.019
	25_40	0.024	0.003	0.001	-0.006	0.050	0.003	0.002	-0.005	0.048
	25_50	-0.010	-0.001	-0.004	-0.007	-0.012	-0.002	-0.006	-0.008	-0.012

Set 3 (ρ= -0.78)	10_20	-0.014	-0.005	-0.022	-0.047	0.012	-0.011	-0.023	-0.047	0.011
	10_25	-0.008	0.000	-0.023	-0.044	0.021	0.000	-0.019	-0.039	0.015
	15_25	0.016	0.000	-0.010	-0.027	0.052	0.001	-0.007	-0.022	0.047
	15_30	-0.001	0.006	-0.003	-0.013	0.009	0.006	0.001	-0.009	0.004
	20_30	-0.011	-0.015	-0.019	-0.029	0.002	-0.014	-0.020	-0.029	0.002
	20_40	-0.008	0.004	-0.004	-0.008	-0.009	0.000	-0.007	-0.011	-0.007
	25_40	0.004	0.004	0.001	-0.003	0.009	0.002	0.000	-0.004	0.010
	25_50	-0.001	-0.005	-0.008	-0.011	0.007	-0.002	-0.005	-0.007	0.004
Set 4 (ρ= -0.64)	10_20	0.005	0.002	-0.021	-0.049	0.048	0.007	-0.017	-0.043	0.042
	10_25	0.031	0.030	0.016	-0.003	0.061	0.031	0.020	0.003	0.054
	15_25	-0.014	-0.014	-0.023	-0.036	0.005	-0.018	-0.023	-0.036	0.004
	15_30	0.018	0.016	0.003	-0.008	0.041	0.012	0.004	-0.007	0.039
	20_30	-0.007	-0.014	-0.017	-0.025	0.009	-0.019	-0.023	-0.031	0.014
	20_40	-0.008	-0.007	-0.013	-0.018	0.001	-0.009	-0.012	-0.017	0.000
	25_40	0.006	-0.003	-0.007	-0.012	0.022	-0.002	-0.005	-0.009	0.018
	25_50	-0.008	-0.010	-0.011	-0.014	-0.003	-0.011	-0.011	-0.014	-0.003

Table 2. Comparison of all the estimators under Case 1 and 2 with respect to %RRMSE in case of all four population sets of all the estimators when available auxiliary variable is inversely related with the study variable

Set	Sample Size (n_m_i)	$\hat{t}_{y\pi}$	Case 1				Case 2			
			$\hat{t}_{yCP1}$	$\hat{t}_{yP1}$	$\hat{t}_{ylr1}$	$\hat{t}_{yR1}$	$\hat{t}_{yCP2}$	$\hat{t}_{yP2}$	$\hat{t}_{ylr2}$	$\hat{t}_{yR2}$
Set 1 (ρ= -0.91)	10_20	1.677	0.377	0.726	0.714	2.975	0.690	0.899	0.890	2.841
	10_25	1.469	0.333	0.638	0.630	2.613	0.677	0.827	0.819	2.461
	15_25	1.199	0.271	0.524	0.513	2.112	0.521	0.659	0.651	2.011
	15_30	1.059	0.246	0.466	0.457	1.873	0.516	0.619	0.613	1.750
	20_30	0.921	0.212	0.404	0.395	1.626	0.420	0.522	0.516	1.535
	20_40	0.770	0.176	0.343	0.336	1.355	0.401	0.473	0.469	1.246
	25_40	0.663	0.154	0.299	0.294	1.170	0.332	0.399	0.397	1.091
	25_50	0.576	0.133	0.258	0.253	1.013	0.322	0.369	0.365	0.917
Set 2 (ρ= -0.85)	10_20	1.781	0.755	0.967	0.958	3.002	0.971	1.116	1.111	2.876
	10_25	1.562	0.668	0.848	0.841	2.639	0.892	1.011	1.006	2.486
	15_25	1.260	0.536	0.682	0.676	2.128	0.708	0.804	0.801	2.018
	15_30	1.139	0.485	0.615	0.610	1.931	0.666	0.746	0.744	1.812
	20_30	0.983	0.420	0.537	0.532	1.656	0.556	0.635	0.633	1.567
	20_40	0.829	0.356	0.454	0.448	1.392	0.518	0.574	0.572	1.279
	25_40	0.716	0.307	0.389	0.385	1.209	0.429	0.478	0.476	1.128
	25_50	0.620	0.267	0.337	0.333	1.045	0.398	0.434	0.432	0.949
Set 3 (ρ= -0.78)	10_20	1.907	1.081	1.220	1.215	3.048	1.229	1.343	1.341	2.924
	10_25	1.708	0.954	1.076	1.071	2.744	1.131	1.218	1.217	2.601
	15_25	1.372	0.763	0.864	0.860	2.209	0.888	0.964	0.961	2.102
	15_30	1.246	0.700	0.794	0.789	1.990	0.840	0.909	0.908	1.869
	20_30	1.073	0.588	0.667	0.662	1.731	0.696	0.751	0.749	1.644
	20_40	0.894	0.497	0.561	0.557	1.439	0.621	0.664	0.663	1.329
	25_40	0.787	0.435	0.493	0.489	1.263	0.532	0.572	0.570	1.179
	25_50	0.672	0.380	0.427	0.424	1.073	0.484	0.513	0.512	0.979
Set 4 (ρ= -0.64)	10_20	2.316	1.685	1.767	1.767	3.353	1.807	1.876	1.877	3.222
	10_25	2.063	1.510	1.576	1.574	2.982	1.635	1.690	1.690	2.840
	15_25	1.665	1.220	1.286	1.284	2.394	1.314	1.366	1.366	2.288
	15_30	1.505	1.090	1.143	1.139	2.177	1.192	1.239	1.237	2.055
	20_30	1.271	0.931	0.975	0.972	1.837	1.007	1.048	1.047	1.749
	20_40	1.072	0.777	0.810	0.807	1.553	0.869	0.895	0.895	1.445
	25_40	0.947	0.697	0.725	0.722	1.364	0.767	0.789	0.788	1.283
	25_50	0.822	0.603	0.626	0.623	1.180	0.679	0.699	0.698	1.084