Available online at http://isas.org.in/isa/jisas
JOURNAL OF THE INDIAN SOCIETY OF
AGRICULTURAL STATISTICS 77(3) 2023 265–274

A New Two Auxiliary Calibration Estimator of the Population Total in Two Stage Sampling Design using Nonlinear Constraints

Pathi Devendra Kumar[1,2], Kaustav Aditya[2], Tauqueer Ahmad[2], Ankur Biswas[2] and Surya Prakash Tripathi[2]

[1]The Graduate School, ICAR-Indian Agricultural Research Institute, New Delhi
[2]ICAR-Indian Agricultural Statistics Research Institute, New Delhi

Received 01 May 2023; Revised 24 August 2023; Accepted 04 September 2023

SUMMARY

In this study, a two auxiliary calibration estimator is proposed under two stage sampling design using a nonlinear constraint with the assumption of availability of population level auxiliaries at the cluster level and the size of the clusters were assumed unknown. The performance of the proposed estimator was evaluated through a simulation study. The empirical result shows that the developed estimator was performing better than the existing estimators under two stage sampling design when population level auxiliary information were available at the cluster level.

Keywords: Calibration estimation; Two stage sampling; Survey weighted estimates; Non-linear constraints; Model assisted estimator.

1. INTRODUCTION

A survey plays a crucial role in gathering information from a population. Sample surveys are conducted with the aim of drawing conclusions about an entire population based on data collected from a selected sample. These conclusions often involve making estimates, such as predicting the average yield of a crop or the percentage of individuals affected by a certain disease. When it comes to sample surveys, researchers typically prefer using single-stage sampling designs with equal probabilities, as these designs aid in creating new and effective estimation methods. However, real-world surveys tend to be more complex and often involve multiple stages of sampling. Among the various sampling designs, the most widely adopted method for survey estimation worldwide is the stratified multi-stage sampling design. In many practical survey scenarios, a simplified version known as two-stage sampling is commonly employed due to its practicality and ease of implementation.

Auxiliary information is frequently employed to enhance the accuracy of survey estimates. The fundamental method of integrating auxiliary information into survey estimation involves using traditional ratio or regression estimators (Hansen *et al.*, 1953). The Calibration Approach (Deville *et al.*, 1992) stands out as one of the frequently utilized methods for effectively leveraging auxiliary information in survey estimates. This method achieves this by generating a new set of weights through the adjustment of sampling design weights using auxiliary data. Calibration weights were initially introduced by Huang and Fuller (1978) and were referred to as regression weights. What sets calibration estimators apart is their ability to operate without presuming any specific model that links the study and auxiliary variables. In the context of single-stage sampling designs, there exists a group of researchers, including Singh *et al.* (1998), Wu *et al.* (2001), Singh *et al.* (2003, 2004), Tracy *et al.* (2003), Singh *et al.* (2011), Koyuncu *et al.* (2014), Sud *et al.* (2014), Clement *et al.* (2014, 2017), Nidhi *et al.* (2017), and Özgül (2018, 2020), Alam *et al.* (2020, 2021), who have embraced the calibration approach for estimating population parameters during the estimation phase.

Moreover, the utilization of auxiliary information further enhances the accuracy of estimating the total

Corresponding author: Kaustav Aditya
E-mail address: kaustav.aditya@icar.gov.in

population in scenarios involving two-stage designs (Sukhatme *et al.*, 1984). In a two-stage sampling framework with varying probabilities, Sahoo *et al.* (1999) introduced a comprehensive set of estimators for calculating the total population of a finite group. These estimators rely on two auxiliary variables, assuming the availability of auxiliary information at the cluster level. Within this context, they proposed a regression-type estimator that achieves an asymptotic minimum variance bound (MVB) under the two-stage sampling design, capitalizing on the two auxiliary information variables accessible at the cluster level of selection. In the presence of two auxiliary information variables solely at the unit level for the selected Primary Sampling Units (PSUs), Sahoo *et al.* (2011) presented a broader category of MVB ratio estimators within a two-stage sampling setup. Additionally, alongside this expanded array of estimators, several researchers have explored the implementation of the calibration approach under a two-stage design, where auxiliary information is accessible. Notable instances include Aditya *et al.* (2014, 2016a, 2016b), Mourya *et al.* (2016), Aditya *et al.* (2017), Aditya *et al.* (2019), Biswas *et al.* (2020, 2023), and Basak *et al.* (2021). These endeavors aim to enhance estimators within a two-stage sampling design through the integration of auxiliary information using the calibration approach.

To further enhance the performance of the calibration estimator, Singh *et al.* (2003) introduced a connection between the generalized regression (GREG) estimator, obtained through Deville *et al.*'s (1992) calibration technique, and the linear regression estimator as presented by Hansen *et al.* (1953). This linkage was developed in line with Singh *et al.*'s observations (2003, 2004) that the cumulative calibrated weights must align with the sum of the design weights. In addition, they presented the concept of a multi-auxiliary calibration estimator that employs multiple auxiliary variables (Singh *et al.*, 2011) within a single-stage sampling design. Given that contemporary surveys often encompass multiple auxiliary information variables, Rao *et al.* (2012) introduced the notion of a multi-auxiliary calibration estimator for the population mean within a stratified single-stage sampling design. This concept employed information from two auxiliary variables and addressed the challenge of determining optimal calibration weights under various calibration conditions through a Mathematical Programming Problem (MPP). Clement *et al.* (2014) further advanced the field by devising an analytical technique to create a multi-auxiliary calibration estimator using MPP with a Chi-square-type loss function, subject to various calibration constraints. Ozgul (2018) then contributed a new calibration estimator for population mean estimation under stratified sampling, utilizing two auxiliary variables. This novel theory described the estimator and optimized calibration weights through nonlinear constraints involving the two auxiliary variables. In the context of estimating population mean using stratified uni-stage random sampling design, Alam *et al.* (2021) proposed a multi-auxiliary calibration estimator. They incorporated multiple constraints derived from auxiliary variables and introduced a new variance function for the study variable, replacing traditional distance functions. This approach assumed knowledge of the population variance, particularly under Neyman allocation.

Most literature in this field has primarily focused on single-stage selection scenarios, although real-world surveys usually involve multistage structures. Multistage sampling introduces complexity due to selection occurring at multiple stages. Addressing this gap, this paper introduces a calibration estimator utilizing two auxiliaries to estimate the population total under a two-stage sampling design. This proposal aligns with the concept pioneered by Ozgul *et al.* (2018), assuming accurate knowledge of cluster-level totals for the two auxiliary variables at the population level. Furthermore, the paper introduces a nonlinear constraint to incorporate cluster-level auxiliary information into the proposed calibration estimator.

The subsequent segments of this paper are structured as follows. The following section outlines the standard notations adopted for the discussion of current calibration estimators within the framework of a two-stage sampling design. This discussion assumes the presence of population-level auxiliary information at the cluster level, in the context of the two-stage sampling design. Section 3 outlines the evolution of two auxiliary calibration estimators under a two-stage sampling design. This development considers the availability of population-level auxiliary information at the cluster level and accommodates instances where the sizes of the Primary Sampling Units (PSUs) are unknown. In Section 4, the outcomes from Monte Carlo simulation studies are detailed. These results serve to evaluate the empirical efficacy of the proposed estimator relative to existing methodologies. Lastly, Section 5 encapsulates the primary concluding remarks drawn from the research.

2. METHODOLOGY

2.1 Notations

Under the two-stage sample design framework, the estimator was created with the presumption that population level auxiliary information is available at the cluster level and the cluster sizes were unknown. Let's divide the population of components U= {1,…, k,…, N_I} into the clusters U_1, U_2,…, U_i,…, U_{N_I}. When there are two stages of selection, the units are referred to as primary stage units (psus) at cluster level and secondary stage units (ssus) at ultimate stage unit level. N_i is used to represent U_i's size.

We have, $U = \bigcup_{i=1}^{N_I} U_i$ and $N = \sum_{i=1}^{N_I} N_i$.

Stage one includes selecting a sample of psus, s_I from U_I in accordance with the design $p_I(.)$ and the inclusion probability π_{Ii} and π_{Iij} at the psu level. The size of s_I was n_I psus. The population components with the labels k= 1,…, N are the ssus.

A sample s_i of size n_i units is drawn from the psu U_i given that U_i was chosen at the psu level, according to a specified design $p_i(.)$ with inclusion probabilities $\pi_{k/i}$ and $\pi_{kl/i}$. There is an invariance and independence property for the second stage sample. The inclusion probabilities at the first stage of selection were given as,

$$\pi_i = Pr\left(i \in s_I\right)$$

$$\pi_{Iij} = \begin{cases} Pr\left(i\, and\; j \in s_I\right), i\, and\; j\;\; belongs\, to \\ \qquad\qquad different\;\; psus \\ \pi_{Ii}, i\, and\;\; j\;\; belongs\, to\, same\, psus. \end{cases}$$

The inclusion probabilities for the second stage of selection were given as,

$$\pi_{k/I} = Pr(k \in s_i | k \in s_I)$$

$$\pi_{kl/i} = \begin{cases} Pr\left(k \text{ and } l \in s_i \;\; |\text{i} \in s_I\right), k\; and\; l \\ \qquad\qquad are\, different \\ \pi_{k/i},\; k \text{ and } l \text{ are same.} \end{cases}$$

Let the study variable be y_k which was observed for $k \in s$. The parameter to estimate was the population total $t_y = \sum_{i=1}^{N} y_k = \sum_{i=1}^{N_I} t_{yi}$ where $t_{yi} = \sum_{i=1}^{N_i} y_k$ i-th psu total.

2.2 Existing estimators under two stage sampling design

The central premise of the study involved the formulation of a novel two auxiliary calibration estimator within a two-stage sampling design. This estimator was designed considering scenarios where population-level auxiliary information is accessible at the cluster or Primary Sampling Unit (PSU) level, and the sizes of the clusters remained unknown, adhering to the bridge constraint established by Singh *et al.* (2011). Given these assumptions, only a limited number of estimators have been developed. Noteworthy among them are the calibration regression-type estimator proposed by Aditya *et al.* (2016), which utilizes single auxiliary information available at the PSU level, and the asymptotic Minimum Variance Bound (MVB) regression-type estimator introduced by Sahoo *et al.* (1999). The latter estimator employs two auxiliary pieces of information accessible at the PSU level. A detailed overview of these existing estimators is presented below.

Estimator 1 (Aditya *et al.*, 2016)

In the context of a two-stage sampling design, a calibration regression-type estimator was introduced. This estimator was developed with the premise that population-level auxiliary information is accessible at the cluster level, and accurate knowledge of the cluster total of the auxiliary information is available. The estimator is governed by two distinct constraints, with one of them being the bridge criteria as defined by Singh *et al.* (2011).

The estimator was given as follows, the Horvitz-Thompson estimator within a two-stage sampling design is considered, with the assumption that population-level auxiliary information (x_{Ii}) is accessible at the cluster level. This auxiliary information (x_{Ii}) is observed for all the sampled clusters, and the accurate value of the summation $\sum_{i=1}^{N_I} x_{1i}$ is available, while the cluster size remains unknown (in accordance with the bridge constraint established by Singh *et al.* (2011)). The formulation of this estimator is provided as follows:

$$\hat{t}_{HT} = \sum_{i=1}^{n_I} \frac{\hat{t}_{yi\pi}}{\pi_{Ii}} = \sum_{i=1}^{n_I} a_{Ii} \hat{t}_{yi\pi} = \sum_{i=1}^{n_I} a_{Ii} \left(\sum_{k=1}^{n_i} \frac{y_k}{\pi_{k/i}} \right)$$

where, $\hat{t}_{yi\pi}$ be the estimator of the cluster total and $a_{Ii} = \frac{1}{\pi_{Ii}}$ is the design weight. After calibration, the proposed estimator will be,

$$\hat{t}^c_{y\pi} = \sum_{i=1}^{n_I} w_{Ii}\hat{t}_{yi\pi}$$

For this purpose, the following chi-square type distance function was minimized

$$\sum_{i=1}^{n_I} \frac{(w_{Ii} - a_{Ii})^2}{a_{Ii}q_{Ii}}$$

subject to the constraints

$$\sum_{i=1}^{n_I} w_{Ii}x_{1i} = \sum_{i=1}^{N_I} x_{1i} \text{ and } \sum_{i=1}^{n_I} w_{Ii} = \sum_{i=1}^{n_I} a_{Ii}.$$

The objective function which was minimized using Lagrange multiplier was given as,

$$\phi(w_{Ii},\lambda) = \sum_{i=1}^{n_I} \frac{(w_{Ii} - a_{Ii})^2}{a_{Ii}q_{Ii}} + \lambda_1\left[\sum_{i=1}^{n_I} w_{Ii}x_{1i} - \sum_{i=1}^{N_I} x_{1i}\right] + \lambda_2\left[\sum_{i=1}^{n_I} w_{Ii} - \sum_{i=1}^{n_I} a_{Ii}\right]$$

The new calibration weights were found as,

$$w_{Ii} = a_{Ii} + \frac{a_{Ii}q_{Ii}\left(\sum_{i=1}^{N_I} x_{1i} - \sum_{i=1}^{n_I} a_{Ii}x_{1i}\right)}{\sum_{i=1}^{n_I} a_{Ii}q_{Ii}x^2_{1i} - \frac{\left(\sum_{i=1}^{n_I} a_{Ii}q_{Ii}x_{1i}\right)^2}{\sum_{i=1}^{n_I} a_{Ii}q_{Ii}}}\left[z_i - \frac{\sum_{i=1}^{n_I} a_{Ii}q_{Ii}x_{1i}}{\sum_{i=1}^{n_I} a_{Ii}q_{Ii}}\right]$$

Using the above mentioned calibrated weight, the estimator of the population totalwhen $q_{Ii} = 1$, was given as,

$$\hat{t}^c_{y\pi} = \sum_{i=1}^{n_I} a_{Ii}\hat{t}_{yi\pi} + \sum_{i=1}^{n_I} a_{Ii}\left\{\frac{\left(\sum_{i=1}^{N_I} x_{1i} - \sum_{i=1}^{n_I} a_{Ii}x_{1i}\right)}{\sum_{i=1}^{n_I} a_{Ii}x^2_{1i} - \frac{\left(\sum_{i=1}^{n_I} a_{Ii}x_{1i}\right)^2}{\sum_{i=1}^{n_I} a_{Ii}}}\right\}\left[x_{1i} - \frac{\sum_{i=1}^{n_I} a_{Ii}x_{1i}}{\sum_{i=1}^{n_I} a_{Ii}}\right]\hat{t}_{yi\pi}$$

$$\hat{t}^c_{y\pi} = \hat{t}_{HT} + \hat{b}\left[\sum_{i=1}^{N_I} x_{1i} - \sum_{i=1}^{n_I} a_{Ii}x_{1i}\right]$$

where,

$$\hat{b} = \frac{\sum_{i=1}^{n_I} a_{Ii}\hat{t}_{yi\pi}\left(x_{1i} - \frac{\sum_{i=1}^{n_I} a_{Ii}x_{1i}}{\sum_{i=1}^{n_I} a_{Ii}}\right)}{\left(\sum_{i=1}^{n_I} a_{Ii}x^2_{1i} - \frac{\left(\sum_{i=1}^{n_I} a_{Ii}x_{1i}\right)^2}{\sum_{i=1}^{n_I} a_{Ii}}\right)}$$

Under SRSWOR the expression takes the form,

$$\hat{t}^c_{y\pi} = \frac{N_I}{n_I}\sum_{i=1}^{n_I} \hat{t}_{yi\pi} + \sum_{i=1}^{n_I} \frac{\left(\sum_{i=1}^{N_I} x_{1i} - \frac{N_I}{n_I}\sum_{i=1}^{n_I} x_{1i}\right)}{n_I\sum_{i=1}^{n_I} x_{1i}^2 - \left(\sum_{i=1}^{n_I} x_{1i}\right)^2}\left[n_I x_{1i} - \sum_{i=1}^{n_I} x_{1i}\right]\hat{t}_{yi\pi}$$

$$\hat{t}^c_{y\pi} = \hat{t}_{HT} + \hat{b}\left[\sum_{i=1}^{N_I} x_{1i} - \frac{N_I}{n_I}\sum_{i=1}^{n_I} x_{1i}\right]$$

where,

$$\hat{b} = \frac{\sum_{i=1}^{n_I}\left(n_I x_{1i} - \sum_{i=1}^{n_I} x_{1i}\right)\hat{t}_{yi\pi}}{\left(n_I\sum_{i=1}^{n_I} x_{1i}^2 - \left(\sum_{i=1}^{n_I} x_{1i}\right)^2\right)} \text{ and } \hat{t}_{yi\pi} = \frac{N_i}{n_i}\sum_{k=1}^{n_i} y_k$$

Estimator 2 (Sahoo *et al.*, 1999)

In 1999, Sahoo *et al.* introduced a comprehensive category of estimators within a two-stage sampling design, leveraging information from two auxiliary variables. Their proposed methodology involved an asymptotic Minimum Variance Bound (MVB) regression-type estimator. This estimation approach was built on the assumption of positive correlation between the two auxiliary variables and their availability at the cluster level of selection.

Let Y, $X_1$and X_2 be study and auxiliary variables of interest respectively. The asymptotic MVB regression estimator was given as,

$$\hat{Y}_{RG} = \frac{\left[\tilde{Y}_i - \gamma_i\left(\tilde{X}_{1i} - X_{1i}\right)\right]}{\pi_{Ii}} - \gamma\left(\tilde{X}_{2i} - X_{2i}\right)$$

where $\gamma = \frac{\sigma_{yx_2} + \sum_{i=1}^{N_I}\sigma^2_{ix_2}\left(\beta_{iyx_2} - \beta_{iyx_1}\beta_{ix_1x_2}\right)/\pi_{Ii}}{\sigma^2_{x_2} + \sum_{i=1}^{N_I}\sigma^2_{ix_2}\left(1-\rho^2_{ix_1x_2}\right)/\pi_{Ii}}$,

$\gamma_i = -(\beta_{iyx_1} + \gamma\beta_{ix_1x_2})$, $\beta_{iyx_1} = \sigma_{iyx_1}/\sigma^2_{ix_1}$, $\rho_{ix_1x_2} = \sigma_{ix_1x_2}/\sigma_{ix_1}\sigma_{ix_2}$,

$\beta_{iyx_2} = \sigma_{iyx_2}/\sigma^2_{ix_2}$, $\beta_{ix_1x_2} = \sigma_{ix_1x_2}/\sigma^2_{ix_2}$,

$\tilde{Y}_i = \sum_{s_i} Y_i/\pi_{k/i}$, $\tilde{X}_{1i} = \sum_{s_i} X_{1i}/\pi_{k/i}$,

$\tilde{X}_{2i} = \sum_{s_i} X_{2i}/\pi_{k/i}$, $\sigma^2_{ix_2} = Var\left(\tilde{X}_{2i}\right)$,

$\sigma_{yx_2} = Cov\left(\sum_{s_I}\frac{\tilde{Y}_i}{\pi_{Ii}}, \sum_{s_I}\frac{\tilde{X}_{2i}}{\pi_{Ii}}\right) - \sum_{i=1}^{N_I}\frac{\sigma_{iyx_2}}{\pi_{Ii}}$,

$\sigma^2_{x_2} = Var(\sum_{s_I} X_{2i}/\pi_{Ii})$, $\sigma^2_{ix_1} = Var\left(\tilde{X}_{1i}\right)$,

$\sigma_{ix_1x_2} = Cov\left(\tilde{X}_{1i}, \tilde{X}_{2i}\right)$ and $\sigma_{iyx_1} = Cov\left(\tilde{Y}_i, \tilde{X}_{1i}\right)$.

Under simple random sampling whout replacement the estimator will be given as,

$$\hat{Y}_{RG} = \frac{N_I}{n_I}\left[\tilde{Y}_i - \gamma_i\left(\tilde{X}_{1i} - X_{1i}\right)\right] - \gamma\left(\tilde{X}_{2i} - X_{2i}\right)$$

where, $\tilde{Y}_i = \sum_{s_i} \frac{N_i}{n_i} Y_i, \tilde{X}_{1i} = \sum_{s_i} \frac{N_i}{n_i} X_{1i}$ and

$\tilde{X}_{2i} = \frac{N_i}{n_i}\sum_{s_i} X_{2i}$,

$$\gamma = -\frac{\sigma_{yx_2} + \frac{N_I}{n_I}\sum_{i=1}^{N_I}\sigma^2_{ix_2}\left(\beta_{iyx_2} - \beta_{iyx_1}\beta_{ix_1x_2}\right)}{\sigma^2_{x_2} + \frac{N_I}{n_I}\sum_{i=1}^{N_I}\sigma^2_{ix_2}\left(1 - \rho^2_{ix_1x_2}\right)}$$ and

$\gamma_i = \sigma_{iyx_1} / \sigma^2_{ix_1}$,

$\rho_{ix_1x_2} = \sigma_{ix_1x_2} / \sigma_{ix_1}\sigma_{ix_2},\ \beta_{iyx_2} = \sigma_{iyx_2} / \sigma^2_{ix_2}, \beta_{ix_1x_2} = \sigma_{ix_1x_2} / \sigma^2_{ix_2}$.

All the variance and the covariance term follows the standard form under two stage sampling design when selection at various stages were done using SRSWOR as given in Sukhatme *et al.* (1984).

3. PROPOSED ESTIMATOR

In this research, we examine the parameters Y, X_1 and X_2, which stand for the primary focus of the study and two additional supporting variables, respectively. The specific observed values of these variables are labeled as y, x_1 and x_2. Instead of adopting the conventional linear methods proposed by Clement *et al.* (2014) for the limiting equation, we embrace the methodology introduced by Ozgul (2018). In this approach, we utilize the ratio of cumulative values or averages from the auxiliary variables as a nonlinear restriction. This strategy assumes the existence of accurate population-level cumulative values or averages for the ratio of the supplementary variable.

To illustrate this, let's consider an example where a surveyor needs to estimate the agricultural stock price. In this case, the surveyor may decide to use the price/earnings ratio of the farmers as an auxiliary variable, as the ratio of these two variables is sufficient to provide the required information. By using the ratio instead of using the variables separately, we can simplify the estimator compared to the approach proposed by Clement *et al.* (2014) while minimizing a chi-square type loss function.

Furthermore, we bring forth an additional limitation referred to as the "bridge constraint," following the proposal of Singh *et al.* (2011). This constraint is implemented to enhance the performance of the estimator and acts as a connection between the conventional linear regression estimator and the GREG (Generalized Regression) estimator. Let the proposed estimator be,

$$\hat{t}^{cp}_{y\pi} = \sum_{i=1}^{n_I} w_{Ii}\hat{t}_{yi\pi}$$

For estimating the calibrated weight first the chi-square type loss function was minimized. Let the chi-square type distance function be,

$$\sum_{i=1}^{n_I}\frac{\left(w_{Ii} - a_{Ii}\right)^2}{a_{Ii}q_{Ii}}$$

The above function will be minimized subject to the constraints

$$\sum_{i=1}^{n_I} w_{Ii}\hat{R}_{Ii} = R_{Ii} \text{ and } \sum_{i=1}^{n_I} w_{Ii} = \sum_{i=1}^{n_I} a_{Ii}.$$

where, $\hat{R}_{Ii} = \frac{\sum_{i=1}^{n_I} x_{1i}}{\sum_{i=1}^{n_I} x_{2i}}$ and $R_{Ii} = \frac{\sum_{i=1}^{N_I} X_{1i}}{\sum_{i=1}^{N_I} X_{2i}}$ are sample and population ratios of two auxiliary variables respectively.

The objective function be defined as,

$$L = \sum_{i=1}^{n_I}\frac{\left(w_{Ii} - a_{Ii}\right)^2}{a_{Ii}q_{Ii}} - 2\lambda_2\left(\sum_{i=1}^{n_I} w_{Ii}\hat{R}_{Ii} - R_{Ii}\right) - 2\lambda_1\left(\sum_{i=1}^{n_I} w_{Ii} - \sum_{i=1}^{n_I} a_{Ii}\right)$$

which was minimized by using the method of Lagrange multiplier to obtain the calibrated weight w_{Ii}. After minimization the new weight was found as,

$$w_{Ii} = a_{Ii} + a_{Ii}q_{Ii}\left(\lambda_1 + \lambda_2\hat{R}_{Ii}\right) \tag{1}$$

Now taking summation on both sides for equation (1),

$$\lambda_1\sum_{i=1}^{n_I} a_{Ii}q_{Ii} + \lambda_2\sum_{i=1}^{n_I} a_{Ii}q_{Ii}\hat{R}_{Ii} = 0 \left(\text{since } \sum_{i=1}^{n_I} w_{Ii} = \sum_{i=1}^{n_I} a_{Ii}\right) \tag{2}$$

Now multiplying equation (1) with $\hat{R}_{Ii}$ and then taking summation and putting in the constraint equation gives,

$$\lambda_1\sum_{i=1}^{n_I} a_{Ii}q_{Ii}\hat{R}_{Ii} + \lambda_2\sum_{i=1}^{n_I} a_{Ii}q_{Ii}\hat{R}^2_{Ii} = R_{Ii} - \sum_{i=1}^{n_I} a_{Ii}\hat{R}_{Ii} \tag{3}$$

From equation (2) and (3) we can write,

$$\begin{bmatrix} \sum_{i=1}^{n_I} a_{Ii}q_{Ii} & \sum_{i=1}^{n_I} a_{Ii}q_{Ii}\hat{R}_{Ii} \\ \sum_{i=1}^{n_I} a_{Ii}q_{Ii}\hat{R}_{Ii} & \sum_{i=1}^{n_I} a_{Ii}q_{Ii}\hat{R}_{Ii}^2 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix} = \begin{bmatrix} 0 \\ R_{Ii} - \sum_{i=1}^{n_I} a_{Ii}\hat{R}_{Ii} \end{bmatrix} \quad (4)$$

Now by solving eq. (4) using system of linear equations, the values of λ_1 and λ_2 are obtained as $\lambda_1 = \frac{\Delta_1}{A}$ and $\lambda_2 = \frac{\Delta_2}{A}$

where, $\Delta_1 = -\left(R_{Ii} - \sum_{i=1}^{n_I} a_{Ii}\hat{R}_{Ii}\right)\left(\sum_{i=1}^{n_I} a_{Ii}q_{Ii}\hat{R}_{Ii}\right)$,

$$\Delta_2 = \left(R_{Ii} - \sum_{i=1}^{n_I} a_{Ii}\hat{R}_{Ii}\right)\left(\sum_{i=1}^{n_I} a_{Ii}q_{Ii}\right),$$

$$A = \left(\sum_{i=1}^{n_I} a_{Ii}q_{Ii}\right)\left(\sum_{i=1}^{n_I} a_{Ii}q_{Ii}\hat{R}_{Ii}^2\right) - \left(\sum_{i=1}^{n_I} a_{Ii}q_{Ii}\hat{R}_{Ii}\right)^2.$$

After solving the above equations, the new calibrated weight will be given as,

$$w_{Ii} = a_{Ii} + a_{Ii}q_{Ii}\left(\frac{\Delta_1 + \Delta_2\hat{R}_{Ii}}{A}\right)$$

Using the new calibrated weight, the proposed estimator will be given as,

$$\hat{t}_{y\pi}^{cp} = \sum_{i=1}^{n_I} w_{Ii}\hat{t}_{yi\pi} = \sum_{i=1}^{n_I}\left(a_{Ii} + a_{Ii}q_{Ii}\left(\frac{\Delta_1 + \Delta_2\hat{R}_{Ii}}{A}\right)\right)\hat{t}_{yi\pi}$$

$$= \hat{t}_{HT} + \hat{\beta}\left(R_{Ii} - \sum_{i=1}^{n_I} a_{Ii}\hat{R}_{Ii}\right)$$

where

$$\hat{\beta} = \left(\frac{\left(\sum_{i=1}^{n_I} a_{Ii}q_{Ii}\hat{t}_{yi\pi}\hat{R}_{Ii}\right)\left(\sum_{i=1}^{n_I} a_{Ii}q_{Ii}\right) - \left(\sum_{i=1}^{n_I} a_{Ii}q_{Ii}\hat{t}_{yi\pi}\right)\left(\sum_{i=1}^{n_I} a_{Ii}q_{Ii}\hat{R}_{Ii}\right)}{\left(\sum_{i=1}^{n_I} a_{Ii}q_{Ii}\right)\left(\sum_{i=1}^{n_I} a_{Ii}\hat{R}_{Ii}^2\right) - \left(\sum_{i=1}^{n_I} a_{Ii}\hat{R}_{Ii}\right)^2}\right)$$

The expression of the proposed estimator under simple random sampling without replacement sampling scheme will be given as,

$$\hat{t}_{y\pi}^{c} = \hat{t}_{HT} + \hat{\beta}\left(R_{Ii} - \frac{N_I}{n_I}\sum_{i=1}^{n_I}\hat{R}_{Ii}\right)$$

Where, $\hat{\beta} = \dfrac{n_I\left(\sum_{i=1}^{n_I}\hat{t}_{yi\pi}\hat{R}_{Ii}\right) - \left(\sum_{i=1}^{n_I}\hat{t}_{yi\pi}\right)\left(\sum_{i=1}^{n_I}\hat{R}_{Ii}\right)}{n_I\left(\sum_{i=1}^{n_I}\hat{R}_{Ii}^2\right) - \left(\sum_{i=1}^{n_I}\hat{R}_{Ii}\right)^2}$ and

$$\hat{t}_{yi\pi} = \frac{N_i}{n_i}\sum_{k=1}^{n_i} y_k$$

3.1 Variance of variance of proposed estimator

The approximate variance of proposed estimator, derived using the Taylor series linearization technique following Sarndal *et al.* (1992)was given as,

$$V\left(\hat{t}_{y\pi}^{cp}\right) = V\left(\hat{t}_{HT} + \hat{\beta}\left(R_{Ii} - \sum_{i=1}^{n_I} a_{Ii}\hat{R}_{Ii}\right)\right)$$

$$== \sum_{i=1}^{N_I}\sum_{j=1}^{N_I}\Delta_{Iij}\frac{U_{Ii}}{\pi_{Ii}}\frac{U_{Ij}}{\pi_{Ij}} + \sum_{i=1}^{N_I}\frac{1}{\pi_{Ii}}\sum_{k=1}^{N_i}\sum_{l=1}^{N_i}\Delta_{k|l}\frac{y_k}{\pi_{k|i}}\frac{y_l}{\pi_{l|i}}$$

where, $U_{Ii} = \sum_{i=1}^{N_I}\sum_{k=1}^{N_i} y_k - \beta R_{Ii}$, $\Delta_{Iij} = \left(\pi_{Iij} - \pi_{Ii}\pi_{Ij}\right)$,

$\Delta_{kl|i} = \left(\pi_{kl|i} - \pi_{k|i}\pi_{l|i}\right)$ and

$$\beta = \frac{\sum_{i=1}^{N_I}\sum_{k=1}^{N_i} y_k R_{Ii} - \frac{1}{N_I}\left(\sum_{i=1}^{N_I}\sum_{k=1}^{N_i} y_k\right)\left(\sum_{i=1}^{n_I} R_{Ii}\right)}{\left(\sum_{i=1}^{N_I} R_{Ii}^2\right) - N_I\left(\sum_{i=1}^{N_I} R_{Ii}\right)^2}$$

For determining the approximate estimator of variance of the proposed estimator suitable re-sampling technique can be employed in future studies.

4. EMPIRICAL EVALUATIONS

In this section, we present the findings from a simulation study designed to assess the practical efficacy of the recommended estimator in comparison to two alternative estimators. The first alternative is the calibration regression estimator introduced by Aditya *et al.* (2016), while the second is the regression estimator proposed by Sahoo *et al.* (1999). This evaluation was conducted within a two-stage sampling framework, assuming that auxiliary information at the cluster stage is available. To gauge the performance of the suggested estimator, we generated an artificial population using model-based Monte Carlo simulation. Subsequently, we drew a total of 5000 samples from this synthetic population to thoroughly evaluate the performance of the proposed estimator. Previous scholarly work has consistently demonstrated that the incorporation of auxiliary information notably enhances the accuracy of estimators, yielding superior outcomes compared to estimators that lack such supplementary variables. Consequently, we compared our proposed estimator to the existing methods introduced by Aditya *et al.* (2016), which utilize two constraint equations and a single auxiliary variable at the cluster level, and by Sahoo *et al.* (1999), which employ two auxiliary variables at the PSU level.

This comparative analysis was based on two primary criteria: the percentage relative bias (%RB) and the percentage root mean square error (%RMSE).

The performance measures used for efficiency comparison were,

$$\%RB = \frac{1}{R}\sum_{r=1}^{R}\frac{\left(\hat{T}_r - T\right)}{T}\times 100$$

$$\%RRMSE = \sqrt{\frac{1}{R}\sum_{r=1}^{R}\frac{\left(\hat{T}_r - T\right)^2}{T}}\times 100$$

where, T is the actual value of the population total, $\hat{T}_r$ is the calculated value of the estimator for the r^{th} run and R is total the number of simulations.

A population of size N=5000 was generated. Under two stage sampling case, population is divided into primary stage units (psus) and secondary stage units (ssus). In simulation study, we fixed the number of psus as $N_I = 100$ and with each psus of size $N_i = 50\left(i = 1,..,100\right)$. Here out of N_I psus we have selected a sample n_I psus with varying sizes i.e.10, 15, 20 and 25. Within each psus, out of N_i units we have selected samples of $n_i\left(i = 1,..,n_I\right)$ units. For each value of n_I, we considered three choices for ssus n_i as, $n_i = p\times N_i$, where p represents the proportion of ssus selected in the sample from each sample psu. We choose three values of p as 0.20, 0.30 and 0.40. This led three values for n_i as 10, 15 and 20. So, we have total twelve combinations of sample sizes. For each case, a simple random sample without replacement (SRSWOR) sample of size n_I psus were first drawn and then from sample psus a sample of n_i ssus were drawn by SRSWOR. Subsequently, the estimation of population total was carried out. In particular, we repeated the simulation process R= 10000 times and calculated the estimates of population total. First the auxiliary variable x_1 and x_2 is generated independently from a normal distribution with mean *5* and variance σ_x^2 (1) i.e., $x_1 \sim N\left(5,1\right)$ and $x_2 \sim N\left(3,1\right)$. After generating both $x_1$and x_2, variable under study y was generated from the model

$$y_i = \alpha_0 + \alpha_1 x_{1i} + \alpha_2 x_{2i} + e_i ; i = 1,2,...,N\ ,$$

where, errors $e_i\left(i = 1,2,...,N\right)$ are generated from standard normal distribution with mean 0 and variance σ_e^2 i.e. $e_i \sim N\left(0,\sigma_e^2\right)$ where σ_e^2 is considered as 1. Here, the values of $\alpha_0 = 70$ and $\alpha_1 = 4, \alpha_2 = 5$ has been chosen randomly and fixed throughout the simulations. The value of the correlation coefficient considered between $x_1$and x_2 were in the tune of around 0.71 and the correlation between the study variable and the auxiliary variables were considered to be around 0.81 and 0.87 respectively. The proposed estimator is compared with the existing estimators based on %RB and %RRMSE. Table 1 presents the various combinations of psu and ssu sample sizes which are used for simulation studies. Table 2 represents the various estimators that were considered for comparison. Table 3 and Table 4 presents %RB and %RRMSE of the proposed and existing estimators respectively.

Table 1. Different combinations of sample sizes used in simulation studies

Set	n_I	n_i	Total Sample size
1	10	10	100
2	10	15	150
3	10	20	200
4	15	10	150
5	15	15	225
6	15	20	300
7	20	10	200
8	20	15	300
9	20	20	400
10	25	10	250
11	25	15	375
12	25	20	500

Table 2. Estimators considered under simulation study

SI. No.	Estimators	Form of the Estimators
1.	$\hat{T}^{(1)}$	$\hat{t}_{y\pi}^{c} = \hat{t}_{HT} + \hat{b}\left(\sum_{i=1}^{N_I} x_{1i} - \sum_{i=1}^{n_I} a_{Ii} x_{1i}\right)$
2.	$\hat{T}^{(2)}$	$\hat{Y}_{RG} = \dfrac{\left[\tilde{Y}_i - \gamma_i\left(\tilde{X}_{1i} - X_{1i}\right)\right]}{\pi_{Ii}} - \gamma\left(\tilde{X}_{2i} - X_{2i}\right)$
3.	$\hat{T}^{(3)}$	$\hat{t}_{y\pi}^{c} = \hat{t}_{HT} + \hat{\beta}\left(R_{Ii} - \dfrac{N_I}{n_I}\sum_{i=1}^{n_I}\hat{R}_{Ii}\right)$

Table 3 and Table 4 presents %RB and %RRMSE of the proposed and existing estimators respectively.

Table 3. Values of % RB of the proposed ($\hat{\boldsymbol{T}}^{(3)}$) and existing estimators ($\hat{\boldsymbol{T}}^{(1)}$ & $\widehat{\boldsymbol{T}}^{(2)}$)

No. of PSUS Selected (n_I)	No. of SSUS Selected (n_i)	$\hat{T}^{(1)}$	$\hat{T}^{(2)}$	$\hat{T}^{(3)}$
10				
	10	0.222	0.170	0.163
	15	0.116	0.109	0.107
	20	0.152	0.141	0.133
15				
	10	0.117	0.112	0.097
	15	0.136	0.125	0.122
	20	0.104	0.099	0.095
20				
	10	0.145	0.132	0.129
	15	0.103	0.097	0.095
	20	0.088	0.081	0.080
25				
	10	0.136	0.123	0.118
	15	0.086	0.080	0.079
	20	0.059	0.053	0.053

Table 4. Values of % RRMSE of the proposed ($\hat{\boldsymbol{T}}^{(3)}$) and existing estimators ($\hat{\boldsymbol{T}}^{(1)}$ & $\hat{\boldsymbol{T}}^{(2)}$)

No. of PSUS Selected (n_I)	No. of SSUS Selected (n_i)	$\hat{T}^{(1)}$	$\hat{T}^{(2)}$	$\hat{T}^{(3)}$
10				
	10	1.199	1.150	1.136
	15	0.962	0.651	0.645
	20	0.664	0.446	0.435
15				
	10	1.241	0.585	0.567
	15	0.668	0.480	0.469
	20	0.752	0.521	0.509
20				
	10	0.661	0.522	0.505
	15	0.664	0.506	0.499
	20	0.539	0.501	0.493
25				
	10	0.646	0.428	0.421
	15	0.552	0.422	0.419
	20	0.435	0.301	0.299

Upon careful examination of Table 3, it is evident that the proposed estimator exhibits a lower percentage relative bias (%RB) compared to the existing estimators under a two-stage sampling design, considering various combinations of PSU and SSU sample sizes. The highest %RB value of 0.163 is observed for n_I =10 and n_I =10 (overall sample size of 100), while the lowest %RB value of 0.053 is observed for n_I =25 and n_i =20 (overall sample size of 500). The proposed estimator demonstrates an improved precision in terms of relative bias compared to the existing estimators as the sample size increases.

Furthermore, it is apparent that the proposed estimator outperforms the Sahoo *et al.* (1999) regression type estimator with two auxiliary variables under a two-stage sampling design when population-level auxiliary information is available at the PSU level. This superiority of the proposed estimator is consistent across various sample sizes of n_I, ranging from 10 to 20, with lower %RB values compared to both existing estimators.

Additionally, it can be observed that the proposed estimator performs better than the Sahoo *et al.* (1999) regression type estimator with two auxiliary variables under a two-stage sampling design when population-level auxiliary information is available at the PSU level in most sample sizes. For sample sizes of n_I ≥25 and n_i ≥20, both estimators with two auxiliary variables exhibit similar performance in terms of %RB. Since sample sizes typically range around 15-20% of the population, it can be concluded that the proposed estimator is superior to the Sahoo *et al.* (1999) regression type estimator in terms of %RB.

Moreover, the proposed estimator also demonstrates better performance than the Aditya *et al.* (2016) calibration estimator with two constraint equations, similar to the proposed estimators that include the bridge constraint of Singh *et al.* (2011). This implies that the proposed estimator outperforms both the existing calibration regression type estimator and the MVB regression type estimator under a two-stage sampling design when population-level auxiliary information is available at the cluster level, considering the criterion of %RB.

From table 4, it can be seen that, the % RRMSE of the proposed estimator is less than the existing estimators under two stage sampling design for various combinations of psu as well as ssu sample sizes. It is observed that the value of the percentage relative root mean squared error is highest 1.136 for n_I =10 and n_i =10 (overall sample size of 100) and it is lowest 0.299 for n_I = 25 and n_i = 20 (overall sample size of 500). With increase in the sample size there is significant gain in precision from the point of view of %RRMSE of the proposed estimator. Also, it is observed that there is decrease in percentage relative root mean squared error with increase in the number of ssus for selected psus. From the above results, it is clear that the %RRMSE for selected psus, decreases with increase in number of ssus selected under each psus for both the proposed and existing estimators. Further, it can also be seen across various sample sizes of n_I, when n_i varies between 10 to 20, the proposed estimator shows lesser %RRMSE w.r.t. the existing estimators. Further, with increase in sample sizes beyond $n_I \geq 25$ and $n_i \geq 20$, both the Sahoo *et al.* (1999) regression type estimator and proposed calibration estimator with two auxiliary information were found to be performing at per w.r.t. %RRMSE with little improvement in the results from the proposed estimator. Hence, it can be concluded that the proposed estimator is performing better than both the existing estimators under two stage sampling design when population level auxiliary information was available at the cluster level.

Table 4 provides further insights into the performance of the proposed estimator by examining the percentage relative root mean squared error (%RRMSE) in comparison to the existing estimators under a two-stage sampling design, considering various combinations of PSU and SSU sample sizes. The highest % RRMSE value is 1.136, observed for n_I =10 and n_i =10 (overall sample size of 100), while the lowest % RRMSE value is 0.299, observed for n_I = 25 and n_i = 20 (overall sample size of 500). As the sample size increases, there is a significant improvement in precision in terms of % RRMSE for the proposed estimator.

Additionally, the % RRMSE decreases with an increase in the number of SSUs selected for each PSU, indicating that increased SSU selection leads to decreased percentage relative root mean squared error for both the proposed and existing estimators. Across various sample sizes of n_I, when n_i varies between 10 to 20, the proposed estimator consistently exhibits a lower % RRMSE compared to the existing estimators.

Furthermore, for sample sizes beyond $n_I \geq 25$ and $n_i \geq 20$, both the Sahoo *et al.* (1999) regression type estimator and the proposed calibration estimator with two auxiliary variables demonstrate similar performance in terms of % RRMSE, with a slight improvement in the results from the proposed estimator. Consequently, it can be concluded that the proposed estimator outperforms both the existing estimators under a two-stage sampling design when population-level auxiliary information is available at the cluster level, considering the criterion of % RRMSE.

5. CONCLUSIONS

In this study a new type of calibration estimator was introduced which employs two auxiliary variables to estimate the population total within a two-stage sampling design context. The estimator employs a non-linear constraint function, leveraging auxiliary information at the cluster level from the population. Additionally, the estimator integrates the bridge constraint methodology introduced by Singh *et al.* (2011) in its formulation. Through extensive Monte Carlo simulations conducted on synthetic datasets, the proposed estimator exhibited remarkable performance when contrasted with the existing methods: the Aditya *et al.* (2016) calibration estimator incorporating two constraint equations and the Sahoo *et al.* (1999) regression-type estimator using two auxiliary variables. This assessment was conducted using the evaluation criteria of percentage relative bias (% RB) and percentage root mean square error (% RRMSE).

Moreover, the incorporation of the bridge constraint from Singh *et al.* (2011) facilitated the asymptotic convergence of the proposed estimator to the classical linear regression estimator developed by Hansen *et al.* (1953). The performance of the newly proposed estimator showcased improvement with increasing sample sizes at both the Primary Sampling Unit (PSU) and Secondary Sampling Unit (SSU) levels. Furthermore, the proposed estimator consistently outperformed the existing estimators in terms of %RB and %RRMSE, especially as the overall sample size (combining PSU and SSU) expanded.

Based on the findings, it can be concluded that the proposed two-auxiliary calibration estimator, designed for a two-stage sampling design scenario with available population-level auxiliary information at the cluster level and unknown cluster sizes, presents a robust approach for accurately estimating the population total.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the valuable comments and suggestions provided by the anonymous referee. The comments have improved the paper significantly.

REFERENCES

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.

Aditya, K., Sud, U.C. and Chandra, H. (2014). *Some Calibration Estimators under Two-Stage Sampling Design*. Project report. ICAR-IASRI, New Delhi.

Aditya, K., Sud, U.C. and Chandra, H. (2016). Calibration approach-based estimation of finite population total under two stage sampling. *Journal of the Indian Society of Agricultural Statistics*, **70**(3), 219-226.

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

Aditya K, Chandra H, Kumar S. and Das S. (2019). Higher Order Calibration Estimator of Finite Population Total under Two Stage Sampling Design when Population Level Auxiliary Information is Available at Unit Level. *Journal of the Indian Society of Agricultural Statistics*, **73**(2), 99-103.

Alam, S., Singh, S. and Shabbir, J. (2020). Calibrated estimators using non-linear calibration

constraints. *Journal of Statistical Computation and Simulation*, **90**(3), 489-514.

Alam, S., Singh, S. and Shabbir, J. (2021). Optimal calibrated weights while minimizing a variance function. *Communications in Statistics-Theory and Methods*, **52**(211), 1-18.

Biswas, A., Aditya, K., Sud, U.C. and Basak, P. (2020). Product type calibration estimation of finite population total under two stage sampling. *Journal of the Indian Society of Agricultural Statistics*, **74**(1), 23-32.

Biswas, A., Aditya, K., Sud, U.C. and Basak, P. (2023).Calibration Estimator in Two Stage Sampling Using Double Sampling Approach when Study Variable is Inversely Related to Auxiliary Variable. *Statistics and Applications*, **21**(1), 11-22

Basak, P., Aditya, K., Kumari, V. and Singh, D. (2021).Two step calibration for estimation of finite population total under two-stage sampling design. *Journal of the Indian Society of Agricultural Statistics*, **75**(3), 235-243.

Clement, E.P. and Enang, E.I. (2014). Multivariate calibration estimation for domain in stratified random sampling. *International Journal of Modern Mathematical Sciences*, **13**(2), 187-197.

Clement, E. (2017). A new ratio estimator of mean in survey sampling by calibration estimation. *Elixir International Journal: Statistics*, **106**, 46461-46465.

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

Hansen, M.H., Hurwitz, W.N., & Madow, W.G. (1953). *Sample survey methods and theory. Vol. 1. Methods and applications.* John Wiley.

Koyuncu, N. and Kadilar, C. (2014). A new calibration estimator in stratified double sampling. *Hacettepe Journal of Mathematics and Statistics*, **43**(2), 337-346.

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

Nidhi, Sisodia, B.V.S., Singh, S. and Singh, S.K. (2017). Calibration approach estimation of the mean in stratified sampling and stratified double sampling. *Communications in Statistics-Theory and Methods*, **46**(10), 4932-4942.

Ozgul N. (2018). New calibration estimator based on two auxiliary variables in stratified sampling, *Communications in Statistics–Theory and Methods*, **48**(6),1481-92. DOI: 10.1080/03610926.2018.1433852.

Ozgul, N. (2020). New improved calibration estimator based on two auxiliary variables in stratified two-phase sampling. *Journal of Statistical Computation and Simulation*, **91**(6), 1243-1256.

Rao, D., Khan, M.G.M. and Khan, S. (2012). Mathematical programming on multivariate calibration estimation in stratified sampling. *World Academy of Science, Engineering and Technology*, **72**, 78-82.

Sahoo, L.N. and Panda, P. (1999). A class of estimators using auxiliary information in two-stage sampling. *Australian & New Zealand Journal of Statistics*, **41**, 405-410.

Sahoo, L.N., Sahoo R.K., Senapati S.C. and Mangaraj A.K. (2011). A general class of estimatorsin two-stage sampling with two auxiliary variables. *Hacettepe Journal of Mathematics and Statistic,* **40**(5), 757-765.

Singh, S. (2003). *Advanced sampling theory with applications*. Dordrecht: Kluwer Academic Publisher.

Singh S (2004). Golden and silver jubilee year-2003 of the linear regression estimators. *Proceedings of the American Statistical Association*, *Survey Method Section*, Toronto: American Statistical Association, 4382-4389.

Singh, S., and Arnab, R. (2011). On calibration of design weights. *Metron*, **LXIX**, 185-205.

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

Sukhatme, P.V., Sukhatme, B.V., Sukhatme, S. and Asok, C. (1984). *Sampling Theory of Surveys with Applications*. Iowa State University Press. (USA).

Tracy, D.S., Singh, S. and Arnab, R. (2003). Note on calibration in stratified and double sampling. *Survey Methodology*, **29**, 99-104.

Wu, C. and Sitter, R.R. (2001). A model calibration approach to using complete auxiliary information from survey data. *Journal of the American Statistical Association*, **96**, 185-193.