AIMS Mathematics

AIMS Mathematics, 7(6): 10609–10624.
DOI: 10.3934/math.2022592
Received: 18 January 2022
Revised: 13 March 2022
Accepted: 24 March 2022
Published: 29 March 2022

http://www.aimspress.com/journal/Math

Research article

Improved generalized class of estimators in estimating the finite population mean using two auxiliary variables under two-stage sampling

Sohaib Ahmad[1], Sardar Hussain[2,*], Javid Shabbir[2], Muhammad Aamir[1], M. El-Morshedy[3,4], Zubair Ahmad[5] and Sharifah Alrajhi[6]

[1] Department of Statistics, Abdul Wali Khan University, Mardan, Pakistan
[2] Department of Statistics, Quaid-i-azam University, Islamabad, Pakistan
[3] Department of Mathematics, College of Science and Humanities in Al-Kharj, Prince Sattam bin Abdulaziz University, Al-Kharj 11942, Saudi Arabia
[4] Department of Mathematics, Faculty of Science, Mansoura University, Mansoura 35516, Egypt
[5] Department of Statistics, Yazd University, P.O. Box 89175-741, Yazd, Iran
[6] Statistics Department, Faculty of Science, King Abdul Aziz University, Jeddah, Kingdom of Saudi Arabia

* **Correspondence:** Email: shussain@stat.qau.edu.pk; Tel: +923439725799.

Abstract: This article addresses the problem of estimating the finite population mean using two auxiliary variables under two-stage sampling scheme. Further the proposed improved class of estimators are provided their genralized form. Mathematical properties of the existing and proposed improved generalized class of estimators are derived up to first order of approximation. We identified 11 members of the improved generalized class of estimators which are more efficient than existing estimators in terms of the percentage relative efficiency. We use two real data sets under two-stage sampling to compare the performances of all of considered estimators.

Keywords: bias; mean square error; auxiliary variables; two-stage sampling; efficiency
Mathematics Subject Classification: 62D05

1. Introduction

In survey sampling, it is well known fact that suitable use of the auxiliary information may improves the precision of an estimator for the unknown population parameters. The auxiliary information can be used either at the design stage or at estimation stage to increase the accuracy of the population parameter estimators. Several authors presented modified different type of estimators for estimating the finite population mean including [4,9,21–27].

The problem of estimation of finite population mean or total in two-stage sampling scheme using the auxiliary information has been well established. The two stage sampling scheme is an improvement over the cluster sampling, when it is not possible or easy to calculate all the units from the selected clusters. One of the main characteristic could be the budget, and it becomes too difficult to collect information from all the units within the selected clusters. To overcome this, one way is to select clusters, called first stage unit (fsus) and from the given population of interest, select a subsample from the selected clusters called the second stage units (ssu). This also benefits to increase the size of the first stage samples which consist of clusters, and assume to be heterogeneous groups. If there is no variation within clusters then might not be possible to collect information from all the units within selected clusters. In many situations, it is not possible to obtain the complete list of ultimate sampling units in large scale sample surveys, while a list of primary units of clusters may be available. In such situations, we select a random sample of first stage units or primary units using certain probability sampling schemes i.e simple random sampling (with or without replacement), systematic sampling and probability proportional to size (PPS), and then we can perform sub-sampling in selected clusters (first stage units). This approach is called two-stage sampling scheme.

Two-stage has a great varaity of applications, which go far beyond the immediate scop of sample survey. Whenever any process involves in chemical, physical, or biological tests that can be performed on a small amount of materail, it is likely to be drawn as a subsample from a larger amount that is itself a sample.

In large scale survey sampling, it is usual to adopt multistage sampling to estimate the population mean or total of the study variable y. [13] proposed a general class of estimators of a finite population mean using multi-auxiliary information under two stage sampling scheme. [1] proposed an alternative class of estimators in two stage sampling with two auxiliary variables. [10] proposed estimators for finite population mean under two-stage sampling using multivariate auxiliary information. [12] suggested a detailed note on ratio estimates in multi-stage sampling. [6] given some stratagies in two stage sampling using auxiliary information. [3] suggested a class of predictive estimaotrs in two stage sampling using auxiliary information. [8] gave a generalized method of estimation for two stage sampling using two auxiliary variables. [5] suggested chain ratio estimators in two stage sampling. For certain related work, we refer some latest articles, i.e., [14–20].

In this article, we propose an improved generalized class of estimators using two auxiliary variables under two-stage sampling scheme. The biases and mean sqaure errors of the proposed generalized class of estimators are derived up to first order of approximation. Based on the numerical results, the proposed class of estimators are more efficient than their existing counterparts.

2. Symbols and notation

Consider a finite population $U = \{U_1, U_2, \ldots, U_N\}$ is divided into N first-stage units (fsus) clusters in the population. Let *N* be the total number of first stage unit in population, *n* be the number of first stage units selected in the sample, M_i be the number of second stage units (ssus) belongs to the i^{th} first stage units (fsus), (*i*=1,2,…, *N*), and mi be the number of fsus selected from the i^{th} fsu in the sample of *n* fsus, (*i*=1,2,…,*n*).

Let y_{ij}, x_{ij} and z_{ij} be values of the study variable y and the auxiliary variables $(x \text{ and } z)$ respectively, for the j^{th} ssus $U_i = (j = 1,2,\ldots,M_i)$, in the i^{th} fsus. The population mean of the study variable y and the auxiliary variables $(x,\ z)$ are given by:

$$\bar{Y} = \frac{1}{N}\sum_{i=1}^{N} u_i\bar{Y}_i, \qquad \bar{X} = \frac{1}{N}\sum_{i=1}^{N} u_i\bar{X}_i, \qquad \bar{Z} = \frac{1}{N}\sum_{i=1}^{N} u_i\bar{Z}_i,$$

where

$$\bar{Y}_i = \frac{1}{M_i}\sum_{j=1}^{M_i} y_{ij},\ \ \bar{X}_i = \frac{1}{M_i}\sum_{j=1}^{M_i} x_{ij},\ \ \bar{Z}_i = \frac{1}{M_i}\sum_{j=1}^{M_i} z_{ij},\ \ (i = 1,2,\ldots,N).$$

$u_i = \frac{M_i}{\bar{M}}$, and $\bar{M} = \frac{M}{N}$, $M = \sum_{i=1}^{N} M_i$,

$R = \frac{\bar{Y}}{\bar{X}}$, and $R_i = \frac{\bar{Y}_i}{\bar{X}_i}$,

$$S_{by}^2 = \frac{1}{N-1}\sum_{i=1}^{N} (u_i\bar{Y}_i - \bar{Y})^2,$$

$$S_{bx}^2 = \frac{1}{N-1}\sum_{i=1}^{N} (u_i\bar{X}_i - \bar{X})^2,$$

$$S_{bz}^2 = \frac{1}{N-1}\sum_{i=1}^{N} (u_i\bar{Z}_i - \bar{Z})^2,$$

$$S_{byx} = \frac{1}{N-1}\sum_{i=1}^{N} (u_i\bar{Y}_i - \bar{Y})(u_i\bar{X}_i - \bar{X}), \qquad S_{byz} = \frac{1}{N-1}\sum_{i=1}^{N} (u_i\bar{Z}_i - \bar{Z})(u_i\bar{Y}_i - \bar{Y}),$$

$$S_{bxz} = \frac{1}{N-1}\sum_{i=1}^{N} (u_i\bar{Z}_i - \bar{Z})(u_i\bar{X}_i - \bar{X}), \qquad S_{iy}^2 = \frac{1}{M_i-1}\sum_{j=1}^{M_i} (y_{ij} - \bar{Y}_i)^2,$$

$$S_{ix}^2 = \frac{1}{M_i-1}\sum_{j=1}^{M_i} (x_{ij} - \bar{X}_i)^2, \qquad S_{iz}^2 = \frac{1}{M_i-1}\sum_{j=1}^{M_i} (z_{ij} - \bar{Z}_i)^2,$$

$$S_{iyx} = \frac{1}{M_i-1}(y_{ij} - \bar{Y}_i)(x_{ij} - \bar{X}_i), \qquad S_{iyz} = \frac{1}{M_i-1}(y_{ij} - \bar{Y}_i)(z_{ij} - \bar{Z}_i),$$

$$S_{ixz} = \frac{1}{M_i-1}\sum_{j=1}^{M_i} (x_{ij} - \bar{X}_i)(z_{ij} - \bar{Z}_i), \qquad (i = 1,2,\ldots N).$$

Similarly for sample data:

$$\bar{y} = \frac{1}{n}\sum_{i=1}^{n} u_i\bar{y}_i = \bar{y}^*(say),\ \bar{x} = \frac{1}{n}\sum_{i=1}^{n} u_i\bar{x}_i = \bar{x}^*(say),\ \bar{z} = \frac{1}{n}\sum_{i=1}^{n} u_i\bar{z}_i = \bar{z}^*(say),$$

where

$$\bar{y}_i = \frac{1}{m_i}\sum_{j=1}^{m_i} y_{ij}, \qquad \bar{x}_i = \frac{1}{m_i}\sum_{j=1}^{m_i} x_{ij}, \qquad \bar{z}_i = \frac{1}{m_i}\sum_{j=1}^{m_i} z_{ij},$$

$$s_{by}^2 = \frac{1}{n-1}\sum_{i=1}^{n} (u_i\bar{y}_i - \bar{y})^2, \qquad s_{bx}^2 = \frac{1}{n-1}\sum_{i=1}^{n} (u_i\bar{x}_i - \bar{x})^2,$$

$$s_{bz}^2 = \frac{1}{n-1}\sum_{i=1}^{n} (u_i\bar{z}_i - \bar{z})^2, \qquad s_{byx} = \frac{1}{n-1}\sum_{i=1}^{n} (u_i\bar{y}_i - \bar{y})(u_i\bar{x}_i - \bar{x}),$$

$$s_{byz} = \frac{1}{n-1}\sum_{i=1}^{n} (u_i\bar{y}_i - \bar{y})(u_i\bar{z}_i - \bar{z}), \qquad s_{bxz} = \frac{1}{n-1}\sum_{i=1}^{n} (u_i\bar{x}_i - \bar{x})(u_i\bar{z}_i - \bar{z}),$$

$$s_{iy}^2 = \frac{1}{m_i-1}(y_{ij} - \bar{y}_i)^2, \qquad s_{ix}^2 = \frac{1}{m_i-1}(x_{ij} - \bar{x}_i)^2,$$

$s_{iz}^2 = \frac{1}{m_i-1}(z_{ij} - \bar{z}_i)^2,$ $s_{iyx} = \frac{1}{m_i-1}\sum_{j=1}^{m_i} (y_{ij} - \bar{y}_i)(x_{ij} - \bar{x}_i),$

$s_{iyz} = \frac{1}{m_i-1}\sum_{j=1}^{m_i} (y_{ij} - \bar{y}_i)(z_{ij} - \bar{z}_i),$ $s_{ixz} = \frac{1}{m_i-1}\sum_{j=1}^{m_i} (x_{ij} - \bar{x}_i)(z_{ij} - \bar{z}_i).$

In order to obtain the biases and mean sqaured errors, we consider the following relative error terms:

$e_0 = \frac{\bar{y}^* - \bar{Y}}{\bar{Y}},\ e_1 = \frac{\bar{x}^* - \bar{X}}{\bar{X}},\ e_2 = \frac{\bar{z}^* - \bar{Z}}{\bar{Z}},$

$E(e_0^2) = \lambda C_{by}^2 + \frac{1}{nN}\sum_{i=1}^{n} u_i^2\theta_i C_{iy}^2 = V_y,$

$E(e_1^2) = \lambda C_{bx}^2 + \frac{1}{nN}\sum_{i=1}^{n} u_i^2\theta_i C_{ix}^2 = V_x,$

$E(e_2^2) = \lambda C_{bz}^2 + \frac{1}{nN}\sum_{i=1}^{n} u_i^2\theta_i C_{iz}^2 = V_z,$

$E(e_0e_1) = \lambda C_{byx} + \frac{1}{nN}\sum_{i=1}^{n} u_i^2\theta_i C_{iyx} = V_{yx},$

$E(e_0e_2) = \lambda C_{byz} + \frac{1}{nN}\sum_{i=1}^{n} u_i^2\theta_i C_{iyz} = V_{yz},$

$E(e_1e_2) = \lambda C_{bxz} + \frac{1}{nN}\sum_{i=1}^{n} u_i^2\theta_i C_{ixz} = V_{xz},$

$C_{by} = \frac{S_{by}}{\bar{Y}},$ $C_{bx} = \frac{S_{bx}}{\bar{X}},$ $C_{bz} = \frac{S_{bz}}{\bar{Z}},$

$C_{byx} = \frac{S_{byx}}{\bar{Y}\bar{X}},$ $C_{byz} = \frac{S_{byz}}{\bar{Y}\bar{Z}},$ $C_{bxz} = \frac{S_{bxz}}{\bar{X}\bar{Z}},$

$C_{iyx} = \frac{S_{iyx}}{\bar{Y}\bar{X}},$ $C_{iyz} = \frac{S_{byz}}{\bar{Y}\bar{Z}},$ $C_{ixz} = \frac{S_{ixz}}{\bar{X}\bar{Z}},$

$C_{iy} = \frac{S_{iy}}{\bar{Y}},$ $C_{ix} = \frac{S_{ix}}{\bar{X}},$ $C_{iz} = \frac{S_{iz}}{\bar{Z}},$

where,

$\theta_i = \left(\frac{1}{m_i} - \frac{1}{M_i}\right),\ \lambda = (\frac{1}{n} - \frac{1}{N}).$

3. Existing estimators

In this section, we consider several estimators of the finite population mean under two-stage sampling that are available in the sampling literature, the properties of all estimators considered here are obtained up-to the first order of approximation.

(i) The usual mean estimator $\bar{y}^* = \bar{y}_0^*$ and its variance under two-stage sampling are given by:

$$\bar{y}_0^* = \frac{1}{n}\sum_{i=1}^{n} u_i\bar{y}_i, \tag{1}$$

and

$$V(\bar{y}_0^*) = \bar{Y}^2V_y = MSE(\bar{y}_0^*). \tag{2}$$

(ii) The usual ratio estimator under two-stage sampling, is given by:

$$\bar{y}_R^* = \bar{y}^* \left(\frac{\bar{X}}{\bar{x}^*}\right), \tag{3}$$

where $\bar{X}$ is the known population mean of x.
The bias and MSE of $\bar{y}_R^*$ to first order of approximation, are given by:

$$Bias(\bar{y}_R^*) = \bar{Y}\left[V_x - V_{yx}\right], \tag{4}$$

and

$$MSE(\bar{y}_R^*) = \bar{Y}^2\left[V_y + V_x - 2V_{yx}\right]. \tag{5}$$

(iii) [2] Exponential ratio type estimator under two-stage sampling, is given by:

$$\bar{y}_E^* = \bar{y}^* exp\left(\frac{\bar{X}-\bar{x}^*}{\bar{X}+\bar{x}^*}\right). \tag{6}$$

The bias and MSE of $\bar{y}_E^*$ to first order of approximation, are given by:

$$Bias(\bar{y}_E^*) = \bar{Y}\left[\frac{3}{8}V_x - \frac{1}{2}V_{yx}\right], \tag{7}$$

and

$$MSE(\bar{y}_E^*) = \bar{Y}^2\left[V_y + \frac{1}{4}V_x - V_{yx}\right]. \tag{8}$$

(iv) The traditional difference estimator under two-stage sampling is given by:

$$\bar{y}_D^* = \bar{y}^* + d(\bar{X} - \bar{x}^*), \tag{9}$$

where d is the constant.
The minimum variance of $\bar{y}_D^*$, is given by:

$$V(\bar{y}_{D\,min}^*) = \bar{Y}^2 V_y(1 - \rho^{*2}) = MSE(\bar{y}_D^*), \tag{10}$$

where $\rho^* = \frac{V_{yx}}{\sqrt{V_y}\sqrt{V_x}}$.
The optimum value of d is $d_{opt} = \frac{\bar{Y}V_{yx}}{\bar{X}V_x}$.
(v) [7] Difference type estimator under two-stage sampling, is given by:

$$\bar{y}_{Rao}^* = d_0\bar{y}^* + d_1(\bar{X} - \bar{x}^*), \tag{11}$$

where d_0 and d_1 are constants.
The bias and minimum MSE of $\bar{y}_{Rao}^*$ to first order of approximation, is given by:

$$Bias(\bar{y}_{Rao}^*) = (d_0 - 1)\bar{Y}, \tag{12}$$

and

$$MSE(\bar{y}_{Rao}^*)_{min} \cong \frac{\bar{Y}^2(V_xV_y - V_{yx}^2)}{V_xV_y - V_{yx}^2 + V_x} = \frac{\bar{Y}^2 V_y(1-\rho^{*2})}{1+V_y(1-\rho^{*2})}. \tag{13}$$

The optimum values of d_0 and d_1 are:

$d_{0opt}=\frac{V_x}{V_xV_y-V_{yx}^2+V_x}$ and $d_{1opt}=\frac{\bar{Y}V_{yx}}{\bar{X}(V_xV_y-V_{yx}^2+V_x)}$.

(vi) The difference-in-ratio type estimator under two-stage sampling, is given by:

$$\bar{y}_{DR}^{*}=[d_2\bar{y}^{*}+d_3(\bar{X}-\bar{x}^{*})]\left(\frac{\bar{X}}{\bar{x}^{*}}\right), \tag{14}$$

where d_2 and d_3 are constants.

The bias and minimum MSE of $\bar{y}_{DR}^{*}$ to first order of approximation, are given by:

$$Bias(\bar{y}_{DR}^{*})\cong\bar{Y}(d_2-1)-d_2\bar{Y}V_{yx}+d_3\bar{X}V_x+d_2V_x\bar{Y}, \tag{15}$$

and

$$MSE(\bar{y}_{DR}^{*})_{min}\cong\frac{\bar{Y}^2(V_x^2V_y-V_xV_{yx}^2-V_xV_y+V_{yx}^2)}{(V_x^2-V_xV_y+V_{yx}^2-V_x)}. \tag{16}$$

The optimum values of d_2 and d_3 are:

$$d_{2opt}=\frac{V_x(V_x-1)}{V_x^2-V_xV_y+V_{yx}^2-V_x},$$

$$d_{3opt}=\frac{-\bar{Y}^2(V_x^2+V_xV_y-V_xV_{yx}-V_{yx}^2-V_x+V_{yx})}{\bar{X}(V_x^2-V_xV_y+V_{yx}^2-V_x)}.$$

(vii) The difference-in-exponential ratio type estimator under two-stage sampling, is given by:

$$\bar{y}_{DE}^{*}=[d_4\bar{y}^{*}+d_5(\bar{X}-\bar{x}^{*})]\exp\left(\frac{\bar{X}-\bar{x}^{*}}{\bar{X}-\bar{x}^{*}}\right), \tag{17}$$

where d_4 and d_5 are constants.

The bias and minimum MSE of $\bar{y}_{DE}^{*}$ to first order of approximation, are given by:

$$Bias(\bar{y}_{DE}^{*})=(d_4-1)\bar{Y}-\frac{1}{2}d_4\bar{Y}V_{yx}+\frac{3}{8}d_4\bar{Y}V_x+\frac{1}{2}d_5V_x, \tag{18}$$

and

$$MSE(\bar{y}_{DE}^{*})_{min}\cong\frac{-\bar{Y}^2\left(V_x^3+16V_x^2V_y-16V_xV_{yx}^2-64V_xV_y+64V_{yx}^2\right)}{64V_xV_y-64V_{yx}^2+64V_x}. \tag{19}$$

The optimum values of d_4 and d_5 are:

$$d_{4opt}=-\frac{1}{8}\frac{V_x(V_x-8)}{(V_xV_y-V_{yx}^2+V_x)},$$

$$d_{5opt}=\frac{\bar{Y}\left(V_x^2+4V_xV_y-V_xV_{yx}-4V_{yx}^2-4V_x+8V_{yx}\right)}{\frac{1}{8}\bar{X}(V_xV_y-V_{yx}^2+V_x)}.$$

(viii) The difference-difference type estimator under two stage sampling, is given by:

$$\bar{y}_{DD}^{*}=\bar{y}^{*}+d_6(\bar{X}-\bar{x}^{*})+d_7(\bar{Z}-\bar{z}^{*}), \tag{20}$$

where d_6 and d_7 are constants.

The minimum variance or MSE of $\bar{y}_{DD}^{*}$ to first order of approximation, is given by:

$$MSE(\bar{y}_{DD}^{*})_{min} \cong \frac{\bar{Y}^2(V_xV_yV_z - V_xV_{yz}^2 - V_{xz}^2V_y + 2V_{xz}V_{yx}V_{yz} - V_{yx}^2V_z)}{V_xV_z - V_{xz}^2}. \quad (21)$$

The optimum values of d_6 and d_7 are:

$$d_6 = \frac{-\bar{Y}(V_{xz}V_{yz} - V_{yx}V_z)}{\bar{X}(V_xV_z - V_{xz}^2)},$$

$$d_7 = \frac{\bar{Y}(V_xV_{yz} - V_{xz}V_{yx})}{\bar{Z}(V_xV_z - V_{xz}^2)}.$$

(ix) The difference-difference type estimator under two stage sampling, is given by:

$$\bar{y}_{DD(R)}^{*} = d_8\bar{y}^* + d_9(\bar{X} - \bar{x}^*) + d_{10}(\bar{Z} - \bar{z}^*), \quad (22)$$

where d_8, d_9 and d_{10} are constants.

The bias and MSE of $\bar{y}_{DD(R)}^{*}$ to first order of approximation is given by:

$$Bias(\bar{y}_{DD(R)}^{*}) = \bar{Y}(d_8 - 1), \quad (23)$$

and

$$MSE(\bar{y}_{DD(R)}^{*}) \cong \frac{\bar{Y}^2(V_xV_yV_z - V_xV_{yz}^2 - V_{xz}^2V_y + 2V_{xz}V_{yx}V_{yz} - V_{yx}^2V_z)}{V_xV_yV_z - V_xV_{yz}^2 - V_{xz}^2V_y + 2V_{xz}V_{yx}V_{yz} - V_{yx}^2V_z + V_xV_z - V_{xz}^2}. \quad (24)$$

The optimum values of d_8, d_9 and d_{10} are given by:

$$d_8 = \frac{V_xV_z - V_{xz}^2}{V_xV_yV_z - V_xV_{yz}^2 - V_{xz}^2V_y + 2V_{xz}V_{yx}V_{yz} - V_{yx}^2V_z + V_xV_z - V_{xz}^2},$$

$$d_9 = \frac{-\bar{Y}(V_{xz}V_{yz} - V_{yx}V_z)}{\bar{X}(V_xV_yV_z - V_xV_{yz}^2 - V_{xz}^2V_y + 2V_{xz}V_{yx}V_{yz} - V_{yx}^2V_z + V_xV_z - V_{xz}^2)},$$

$$d_{10} = \frac{\bar{Y}(V_xV_{yz} - V_{xz}V_{yx})}{\bar{Z}(V_xV_yV_z - V_xV_{yz}^2 - V_{xz}^2V_y + 2V_{xz}V_{yx}V_{yz} - V_{yx}^2V_z + V_xV_z - V_{xz}^2)}.$$

4. Proposed estimator

The principal advantage of our proposed improved generalized class of estimators under two-stage sampling is that it is more flexible, efficient, than the existing estimators. The mean square errors based on two data sets are minimum and percentage relative efficiency is more than hundred as compared to the existing estimators considered here. We identified 11 estimators as members of the proposed class of estimators by substituting the different values of $w_i (i = 1,2,3)$, δ and γ. On the lines of [2,7], we propose the following generalized improved class of estimators under two stage sampling for estimation of finite population mean using two auxiliary varaible as given by:

$$\bar{y}_G^* = [w_1\bar{y}^* + w_2(\bar{X} - \bar{x}^*) + w_3(\bar{Z} - \bar{z}^*)]\left[\left\{exp\ \ \delta\left(\frac{\bar{X}-\bar{x}^*}{\bar{X}+\bar{x}^*}\right)\right\}\left(\frac{\bar{X}}{\bar{x}^*}\right)^{\gamma}\right], \tag{25}$$

where $w_i(i = 1,2,3)$ are constants, whose values are to be determined; δ and γ are constants i.e., $(0 \le \delta,\gamma \le 1)$ and can be used to construct the different estimators.

Using (25), solving $\bar{y}_G^*$ in terms of errors, we have

$$\bar{y}_G^* - \bar{Y} = (w_1 - 1)\bar{Y} + w_1\bar{Y}\left\{e_0 - \frac{1}{2}\alpha_1 e_1 + \frac{1}{8}\alpha_2 e_1^2 - \frac{1}{2}\alpha_1 e_0 e_1\right\}$$
$$-w_2\bar{X}\left\{e_1 - \frac{1}{2}\alpha_1 e_1^2\right\} - w_3\bar{Z}\left\{e_2 - \frac{1}{2}\alpha_1 e_1 e_2\right\},$$

where

$\alpha_1 = \delta + 2\gamma$ and $\alpha_2 = \delta(\delta + 2) + 4\gamma(\delta + \gamma + 1)$.

The bias and MSE of $\bar{y}_G^*$ are given by:

$$Bias(\bar{y}_G^*) \cong (w_1 - 1)\bar{Y} + w_1\bar{Y}\left\{\frac{1}{8}\alpha_2 V_x - \frac{1}{2}\alpha_1 V_{yx}\right\} + w_2\bar{X}\alpha_1\frac{V_x}{2} + w_3\bar{Z}\alpha_1\frac{V_{xz}}{2}, \tag{26}$$

and

$$MSE(\bar{y}_G^*) \cong (w_1 - 1)^2 + w_1^2\bar{Y}^2 A + w_2^2\bar{X}^2 B + w_3^2\bar{Z}^2 C - w_1\bar{Y}^2 D - w_2\bar{Y}\bar{X}E$$
$$-w_3\bar{Y}\bar{Z}F + 2w_1w_2\bar{Y}\bar{X}G + 2w_1w_3\bar{Y}\bar{Z}H + 2w_2w_3\bar{X}\bar{Z}I, \tag{27}$$

where

$$A = V_y + \frac{1}{4}V_x(\alpha_1^2 + \alpha_2) - 2\alpha_1 V_{yx}, \qquad B = V_x, \qquad C = V_z\,,$$

$$D = \frac{1}{4}\alpha_2 V_x - \alpha_1 V_{yx}, \qquad E = \alpha_1 V_x, \qquad F = \alpha_1 V_{xz},$$

$$G = \alpha_1 V_x - V_{yx}, \qquad H = \alpha_1 V_{xz} - V_{xz},\ \ I = V_{xz}.$$

Solving (27), the minimum MSE of $\bar{y}_G^*$ to first order of approximation are given by:

$$MSE(\bar{y}_G^*)_{min} = \bar{Y}^2[1 - \frac{\Omega_2}{4\Omega_1}], \tag{28}$$

where

$$\Omega_1 = ABC - AI^2 - BH^2 - CG^2 + 2GHI + BC - I^2,$$

and

$$\Omega_2 = ABF^2 + ACE^2 - 2AEFI + BCD^2 - 2BDFH - 2CDEG - D^2I^2 + 2DEHI$$
$$+2DFGI - E^2H^2 + 2EFGH - F^2G^2 + 4BCD + BF^2 - 4BFH + CE^2$$
$$-4CEG - 4DI^2 - 2EFI + 4EHI + 4FGI + 4BC + 4I^2.$$

The optimum values of $w_i (i = 1, 2, 3)$ are given by:

$w_{1opt} = \frac{\Omega_3}{2\Omega_1}, w_{2opt} = \frac{\bar{Y}\Omega_4}{2\bar{X}\Omega_1}$, and $w_{3opt} = \frac{\bar{Y}\Omega_5}{2\bar{Z}\Omega_1}$,

where

$$\Omega_3 = BCD - BFH - CEG - DI^2 + EHI + FGI + 2GI + 2BC - 2I^2,$$

$$\Omega_4 = ACE - AFI - CDG + DHI - EH^2 + FGH + CE - 2CG - FI + 2HI,$$

$$\Omega_5 = ABF - AEI - BDH + DGI + EGH - FG^2 + BF - 2BH - EI + 2GI.$$

From (28), we produce the following two estimators called $\bar{y}^*_{G1}$ and $\bar{y}^*_{G2}$. Put $(\delta = 0, \gamma = 1)$ and $(\delta = 1, \gamma = 0)$ in (25), we get the following two estimators respectively:

$$(i) \quad \bar{y}^*_{G1} = [w_4\bar{y}^* + w_5(\bar{X} - \bar{x}^*) + w_6(\bar{Z} - \bar{z}^*)]\left(\frac{\bar{X}}{\bar{x}^*}\right),$$

$$(ii) \quad \bar{y}^*_{G2} = [w_7\bar{y}^* + w_8(\bar{X} - \bar{x}^*) + w_9(\bar{Z} - \bar{z}^*)]\exp\left(\frac{\bar{X}-\bar{x}^*}{\bar{X}+\bar{x}^*}\right),$$

where $w_i (i = 4, 5, 6, 7, 8, 9)$ are constants. Solving $\bar{y}^*_{G1}$, in terms of errors, we have:

$$(\bar{y}^*_{G1} - \bar{Y}) = [-\bar{Y} + w_4\bar{Y} + w_4\bar{Y}e_o - w_5\bar{X}e_1 - w_6\bar{Z}e_2].\left[1 - \frac{1}{2}e_1 + \frac{3}{8}e_1^2\right],$$

$$(\bar{y}^*_{G1} - \bar{Y}) = \begin{bmatrix} w_4\bar{Y} - w_4\bar{Y}e_1 + \frac{3}{8}w_4\bar{Y}e_1^2 + w_4\bar{Y}e_0 - \frac{1}{2}w_4\bar{Y}e_0e_1 - \bar{Y} + \frac{1}{2}\bar{Y}e_1 - \frac{3}{8}\bar{Y}e_1^2 \\ -w_5\bar{X}e_1 + \frac{1}{2}w_5\bar{X}e_1^2 - w_6\bar{Z}e_2 + \frac{1}{2}w_6\bar{Z}e_1e_2 \end{bmatrix}. \tag{29}$$

The bias and MSE of $\bar{y}^*_{G1}$, to first order of approximation is given by:

$$\text{Bias}(\bar{y}^*_{G1}) = \frac{3}{8}w_4\bar{Y}V_x^2 \text{-} \frac{1}{2}w_4\bar{Y}V_{yx} - \frac{3}{8}\bar{Y}V_x^2 + \frac{1}{2}w_5\bar{X}V_x^2 + \frac{1}{2}w_6\bar{Z}V_{xz},$$

By squaring and taking expectation of (29), we get the mean square error:

$$MSE(\bar{y}^*_{G1}) = \begin{bmatrix} w_6^2V_z^2 + w_4^2V_y^2 + w_5^2V_x^2 - 2w_4^2RV_{yx} + 2w_4RV_{yx} - 2Rw_5V_x^2 \\ -2w_4R^2V_x^2 + w_4^2RV_x^2 + 2w_4Rw_5V_x^2 - 2Rw_6V_{xz} + w_4^2\bar{Y}^2 \\ -2w_4\bar{Y}^2 + \bar{Y}^2 + R^2V_x^2 - 2w_4w_5V_{yx} - 2w_4w_6V_{yz} + 2w_5w_6V_{xz} \\ +2w_4Rw_6V_{xz} \end{bmatrix}. \tag{30}$$

Differentiate (30) with respect to w_4, w_5 and w_6, we get the optimum values of w_4, w_5 and w_6 i.e.,

$$w_{4opt} = \frac{\bar{Y}^2(V_x^2V_z^2 - V_{xz}^2)}{\begin{bmatrix} -R^2V_x^4V_z^2 + RV_x^4V_z^2 + \bar{Y}^2V_x^2V_z^2 + R^2V_x^2V_{xz}^2 + V_x^2V_y^2V_z^2 \\ -RV_x^2V_{xz}^2 - \bar{Y}^2V_{xz}^2 - V_x^2V_{yz}^2 - V_{xz}^2V_y^2 - V_{yx}^2V_z^2 + 2V_{xz}V_{yx}V_{yz} \end{bmatrix}},$$

$$w_{5opt} = \frac{\begin{bmatrix} -R^3V_x^4V_z^2 - R^2V_x^4V_z^2 - R^3V_x^2V_{xz}^2 - RV_x^2V_y^2V_z^2 + R^2V_x^2V_{xz}^2 \\ -\bar{Y}^2V_{yx}V_z^2 + RV_x^2V_{yz}^2 + RV_{xz}^2V_y^2 + RV_{yx}^2V_z^2 + \bar{Y}^2V_{xz}V_{yz} - 2R_{Vxz}V_{yx}V_{yz} \end{bmatrix}}{\begin{bmatrix} -R^2V_x^4V_z^2 + RV_x^2V_z^2 + \bar{Y}^2V_x^2V_z^2 + R^2V_x^2V_{xz}^2 + V_x^2V_y^2V_z^2 \\ -RV_x^2V_{xz}^2 - \bar{Y}^2V_{xz}^2 - V_x^2V_{yz}^2 - V_{xz}^2V_y^2 - V_{yx}^2V_z + 2V_{xz}V_{yx}V_{yz} \end{bmatrix}},$$

$$w_{6opt} = \frac{\bar{Y}^2(V_x^2V_{yz}-V_{xz}V_{yx})}{\begin{bmatrix} -R^2V_x^4V_z^2+RV_x^4V_z^2+\bar{Y}^2V_x^2V_z^2+R^2V_x^2V_{xz}^2+V_x^2V_y^2V_z^2 \\ -RV_x^2V_{xz}^2-\bar{Y}^2V_{xz}^2-V_x^2V_{yz}^2-V_{xz}^2V_y^2-V_{yx}^2V_z^2+2V_{xz}V_{yx}V_{yz} \end{bmatrix}}.$$

Substituting the optimum values of w_4, w_5 and w_6 in (30), we get the minimum mean square error of $\bar{y}_{G1}^*$, given by:

$$MSE(\bar{y}_{G1}^*)_{min} = \frac{\bar{Y}^2\begin{bmatrix} R^2V_x^4V_z^2-RV_x^4V_z^2-R^2V_x^2V_{xz}^2-V_x^2V_y^2V_z^2 \\ +RV_x^2V_{xz}^2+V_x^2V_{yz}^2+V_{xz}^2V_y^2+V_{yx}^2V_z^2-2V_{xz}V_{yx}V_{yz} \end{bmatrix}}{\begin{bmatrix} R^2V_x^4V_z^2-RV_x^4V_z^2-\bar{Y}^2V_x^2V_z^2-R^2V_x^2V_{xz}^2-V_x^2V_y^2V_z^2 \\ +RV_x^2V_{xz}^2+\bar{Y}^2V_{xz}^2+V_x^2V_{yz}^2+V_{xz}^2V_y^2+V_{yx}^2V_z^2-2V_{xz}V_{yx}V_{yz} \end{bmatrix}}. \tag{31}$$

Solving $\bar{y}_{G2}^*$, in terms of errors, we have

$$(\bar{y}_{G2}^* - \bar{Y}) = [w_7\bar{Y} + w_7\bar{Y}e_0 - \bar{Y} - w_8\bar{X}e_1 - w_9\bar{Z}e_2](1 - e_1 + e_1^2),$$

or

$$(\bar{y}_{G2}^* - \bar{Y}) = \begin{bmatrix} w_7\bar{Y} + w_7\bar{Y}e_0 - \bar{Y} - w_8\bar{X}e_1 - w_9\bar{z}e_2 - w_7\bar{Y}e_1 - w_7\bar{Y}e_1 + \bar{Y}e_1 \\ +w_8\bar{X}e_1^2 - w_9\bar{Z}e_1e_2 + w_7\bar{Y}e_1^2 - \bar{Y}e_1^2 \end{bmatrix}. \tag{32}$$

The Bias and MSE $\bar{y}_{G2}^*$, to first order of approximation, is given by:

Bias($\bar{y}_{G2}^*$) = $w_8\bar{X}V_x^2 - w_9\bar{Z}V_{xz} + w_7\bar{Y}V_x^2 - \bar{Y}V_x^2$.

By squaring and taking expectation of (32), we get the mean square error:

$$MSE(\bar{y}_{G2}^*) = 4Rw_7V_{yx} - 4RV_x^2w_8 + \bar{Y}^2 - 2w_7\bar{Y}^2 + 3R^2V_x^2 - 6w_7R^2V_x^2 + 4w_7w_9RV_{xz}$$

$$-2w_7w_8V_{yx} - 2w_7\bar{Y}w_9V_{yz} + 2w_8w_9V_{xz} - 4Rw_9V_{xz} - 4w_7^2RV_{yx} + 3w_7^3R^2V_x^2$$

$$+w_9^2V_z^2 + w_8^2V_x^2 + 4w_7w_8RV_x^2 + w_7^2\bar{Y}^2. \tag{33}$$

Differentiate (33) with respect to w_7, w_8 and w_9, we get the optimum values of w_7, w_8 and w_9 i.e.,

$$w_{7opt} = \frac{(V_x^2V_z^2-V_{xz}^2)(-R^2V_x^2+\bar{Y}^2)}{R^2V_x^4V_z^2+\bar{Y}^2V_x^2V_{yz}^2-\bar{Y}^2V_x^2V_z^2+\bar{Y}^2V_{xz}^2-2\bar{Y}V_{xz}V_{yx}V_{yz}+V_{yx}^2V_z^2},$$

$$w_{8opt} = \frac{\begin{bmatrix} 2\bar{Y}^2RV_x^2V_{yz}^2-\bar{Y}^2R^2V_x^2V_{xz}V_{yz}+R^2V_x^2V_{yx}V_z^2 \\ +\bar{Y}^3V_{xz}V_{yz}-\bar{Y}^2V_{yz}V_z^2-4\bar{Y}RV_{xz}V_{yx}V_{yz}+2RV_{yx}^2V_z \end{bmatrix}}{\begin{bmatrix} R^2V_x^4V_z^2+\bar{Y}^2V_x^2V_{yz}^2-\bar{Y}^2V_x^2V_z^2-R^2V_x^2V_{xz}^2 \\ +\bar{Y}^2V_{xz}^2-2\bar{Y}V_{xz}V_{yx}V_{yz}+V_{yx}^2V_z^2 \end{bmatrix}},$$

$$w_{9opt} = -\frac{-\bar{Y}R^2V_x^4V_{yz}+\bar{Y}^3V_x^2V_{yz}+R^2V_x^2V_{xz}V_{yx}-\bar{Y}^2V_{xz}V_{yx}}{R^2V_x^4V_z^2+\bar{Y}^2V_x^2V_{yz}^2-\bar{Y}^2V_x^2V_z^2-R^2V_x^2V_{xz}^2+\bar{Y}^2V_{xz}^2-2\bar{Y}V_{xz}V_{yx}V_{yz}+V_{yx}^2V_z^2}.$$

Substituting the optimum values of w_7, w_8 and w_9 in (33), we get the minimum mean square error of $\bar{y}_{G2}^*$, given by:

$$MSE(\bar{y}_{G2}^*)_{min} = \frac{(R^2V_x^2-\bar{Y}^2)(\bar{Y}^2V_x^2V_{yz}^2+V_{yx}^2V_z^2-2\bar{Y}V_{xz}V_{yx}V_{yz})}{\begin{bmatrix} -\bar{Y}^2(V_x^2V_{yz}^2-V_x^2V_z^2+V_{xz}^2)+2\bar{Y}^2V_{xz}V_{yx}V_{yz} \\ -R^2V_x^4V_z^2+R^2V_x^2V_{xz}^2-V_{yx}^2V_z^2 \end{bmatrix}}. \tag{34}$$

We can generate the considered and many more estimators from (25), by substituting the different values of $w_i (i = 1,2,3)$, δ and γ, given in Table 1.

Table 1. Members of the proposed generalized family of estimators.

w_1	w_2	w_3	σ	γ	Estimator
1	0	0	0	0	$\bar{y}^*$
1	0	0	0	1	$\bar{y}_R^*$
1	0	0	1	0	$\bar{y}_E^*$
1	d	0	0	0	$\bar{y}_D^*$
d_0	d_1	0	0	0	$\bar{y}_{Rao}^*$
d_2	d_3	0	0	1	$\bar{y}_{DR}^*$
d_4	d_5	0	1	0	$\bar{y}_{DE}^*$
0	d_6	d_7	0	0	$\bar{y}_{DD}^*$
d_8	d_9	d_{10}	0	0	$\bar{y}_{DD(R)}^*$
w_4	w_5	w_6	0	1	$\bar{y}_{G1}^*$
w_7	w_8	w_9	1	0	$\bar{y}_{G2}^*$

5. Numerical study

Population 1. [Source: [11], Model Assisted Survey Sampling]
There are 124 countries (second stage units) divided into 7 continents (first stage units) according to locations. Continent 7^{th} consists of only one country therefore, we placed 7^{th} continent in 6^{th} continent.
We considered:

y=1983 import (in millions U.S dollars),
x=1983 export (in millions U.S dollars),
z=1982 gross national product (in tens of millions of U.S dollars).

The data are divided into 6 clusters, having N =6, and n =3. Also $\sum_{i=1}^{N} M_i$ =124, $\bar{M}$ =20.67. In Table 2, we show cluster sizes, and population means of the study variable (y) and the auxiliary variables (x, z). Tables 3 and 4 give some results.

Table 2. Cluster sizes with population means.

No. of clusters	M_i	m_i	u_i	$\bar{Y}_i$	$\bar{X}_i$	$\bar{Z}_i$
1	38	15	1.8387	2254.6	1901.1	1029.158
2	14	6	0.6774	25533.14	22083.21	25671.57
3	11	4	0.5323	3602.82	5835.455	5028.818
4	33	13	1.5968	12156.79	12438.85	7533.939
5	24	10	1.1613	34226.79	33198	16314.42
6	4	2	0.1936	26392.5	29360.5	43967.75

Table 3. Statistical computation.

$(u_i\bar{Y}_i-\bar{Y})^2$	$(u_i\bar{X}_i-\bar{X})^2$	$(u_i\bar{Z}_i-\bar{Z})^2$	$(u_i\bar{Y}_i-\bar{Y})$ $(u_i\bar{X}_i-\bar{X})$	$(u_i\bar{X}_i-\bar{X})$ $(u_i\bar{Z}_i-\bar{Z})$	$(u_i\bar{Y}_i-\bar{Y})$ $(u_i\bar{Z}_i-\bar{Z})$
109395354.7	116233397.3	69704363.4	112762554.6	90010971.3	87323155.9
7243544.5	465731.6	51103837.3	1836722	4878593.3	19239878.4
160959577.1	124780258.7	57219949.1	141720068	84498047.6	95969259.7
23109139.4	31199313.33	3200403.1	26851243.6	9992516.3	8599916.4
632160672.8	589329821.2	75771962.7	610369671.8	211316533.3	218860811.8
90158398.4	73831543.1	2989684.1	81587582.8	14857085.7	16417829.8

Table 4. Statistical computations of variances and covariances.

S_{iy}^2	S_{ix}^2	S_{iz}^2	S_{ixy}	S_{ixz}	S_{iyz}
14634002.89	13229390.42	3667896.461	12035361.66	5676138.848	7031654.09
5199331742	3024354709	6568461403	3920918987	42963119811	5785526585
17474303.56	67544530.07	63348742.76	3322301379	62246450.49	32019714.86
510689624	689903319	440717912.5	586829812.3	522773788	447429378.1
1530618991	1588803380	408376223	1544450491	757258674.4	7559765056
1361248223	1782024492	5663081987	1557362451	3157897870	2755900798

$$S_{by}^2 = 204605337.7, \quad S_{bx}^2 = 187168013, \quad S_{bz}^2 = 51998039.99,$$

$$S_{byx} = 195025568.6, \quad S_{byz} = 89282170.45, \quad S_{bxz} = 83110749.52$$

$$V_y = 0.27028, \quad V_x = 0.25137, \quad V_z = 0.30933,$$

$$V_{yx} = 0.25723, \quad V_{yz} = 0.24573, \quad V_{xz} = 0.22493.$$

$$\bar{Y} = 14604.76564, \quad \bar{X} = 14276.72113, \quad \bar{Z} = 10241.22672.$$

Population 2. [Source: [11], Model Assisted Survey Sampling]

Similarly we considered the data as mentioned in Population 1,
y=1983 import (in millions U.S dollars),
x=1981 military expenditure (in tens of millions U.S dollars),
z=1980 population (in millions).
The data are divided into 6 clusters having $N = 6$, $n = 3$, $\sum_{i=1}^{N} M_i = 124$, $\bar{M} = 20.67$.
In Table 5, we show cluster sizes, and means of the study variable (y) and the auxiliary variables (x, z). Tables 6 and 7 give some computation results.

Table 5. Cluster sizes with population means.

No. of clusters	M_i	m_i	u_i	$\bar{Y}_i$	$\bar{X}_i$	$\bar{Z}_i$
1	38	15	1.8387	13.03684	418.3421	11.88421
2	14	6	0.6774	27.35	10065.21	26.1857
3	11	4	0.5323	23.13636	484.45	21.8818
4	33	13	1.5968	79.65455	3377.75	75.2424
5	24	10	1.1613	20.28333	4929.41	20.9583
6	4	2	0.1936	74.15	30676.25	70.975

Table 6. Statistical computation.

$(u_i\bar{Y}_i-\bar{Y})^2$	$(u_i\bar{X}_i-\bar{X})^2$	$(u_i\bar{Z}_i-\bar{Z})^2$	$(u_i\bar{Y}_i-\bar{Y})(u_i\bar{X}_i-\bar{X})$	$(u_i\bar{X}_i-\bar{X})(u_i\bar{Z}_i-\bar{Z})$	$(u_i\bar{Y}_i-\bar{Y})(u_i\bar{Z}_i-\bar{Z})$
109395354.7	11504653.07	171.5489	35430628.04	44434.42674	136842.964
7243544.5	7165707.84	296.6622	−461113.24893	7355674.750	−47336.072
160959577.1	15247937.12	544.5471	49517362.59	91132.5364	295949.600
23109139.4	1556060.53	7312.6397	6074534.22	106667.8685	416415.347
632160672.8	2494694.19	112.4143	39914472.69	−16750.3943	−268005.273
90158398.4	3214861.18	451.071	−16997583.03	−38085.13867	201365.558

Table 7. Statistical computations of variances and covariances.

S_{iy}^2	S_{ix}^2	S_{iz}^2	S_{iyx}	S_{ixz}	S_{iyz}
270.9083357	594166.8257	222.4889331	6380.484353	5806.286629	245.4143812
3906.928847	1281691972	3683.070549	2135135.281	2082979.098	3792.853077
1339.404545	461472.2727	1174.031636	13075.32182	12298.74909	1253.851727
45082.17318	53848774.81	83850.37836	1082424.717	1476243.493	43109.42511
368.9423188	52672480.78	364.7838949	117010.6551	116939.2322	366.7860145
18401.07	3453923758	16855.5025	7970505.317	7628400.308	17611.33833

$S_{by}^2 = 2002.428957$, $S_{bx}^2 = 8236782.79$, $S_{bz}^2 = 1782.2076$,

$S_{byx} = 28451.30273$, $S_{byz} = 1888.920758$, $S_{xz} = 27939.79$,

$V_y = 0.48633$, $V_x = 0.39654$, $V_z = 0.72000$,

$V_{yx} = 0.14250$, $V_{yz} = 0.48726$, $V_{xz} = 0.16552$,

$\bar{Y} = 36.7702$, $\bar{X} = 4163.56$, $\bar{Z} = 34.8552$.

The results based on Tables 2–7 are given in Tables 8 and 9 having biasses, mean square errors, and percentage relative efficiencies of the poposed and exisitng estimators w.r.t $\bar{y}_0^*$. Tables 8 and 9 show that the proposed estimators perform well as compared to the existing estimators considered here.

Table 8. Biases of different estimators in both data sets.

Estimator	Population 1	Population 2
$\bar{y}_0^*$, $\bar{y}_D^*$, $\bar{y}_{DD}^*$	0	0
$\bar{y}_R^*$	−85.58393	9.341102
$\bar{y}_E^*$	−501.692	2.847944
$\bar{y}_{Rao}^*$	−102.2916	−11.14854
$\bar{y}_{DR}^*$	62517.58	−5.729291
$\bar{y}_{DE}^*$	−1040.227	−4.254352
$\bar{y}_{DD(R)}^*$	−8687.674	0.8665627
$\bar{y}_{G1}^*$	−911.2082	42.56601
$\bar{y}_{G2}^*$	4097419	−660.0231

Table 9. MSE and PRE of different estimators w.r.t $\bar{y}_0^*$.

Estimators	Population 1 MSE	Population 1 PRE	Population 2 MSE	Population 2 PRE
$\bar{y}_0^*$	57649726.19	100	657.541	100
$\bar{y}_R^*$	1532261.42	3762.39	808.353	81.3433
$\bar{y}_E^*$	16186983.44	356.149	598.912	109.789
$\bar{y}_D^*$	1503085.83	3835.42	588.307	111.768
$\bar{y}_{Rao}^*$	1492567.93	3862.45	409.935	160.401
$\bar{y}_{DR}^*$	1489069.32	3871.53	341.831	192.358
$\bar{y}_{DE}^*$	1189664.93	4845.88	366.982	179.176
$\bar{y}_{DD}^*$	1025752.55	5620.24	208.189	315.839
$\bar{y}_{DD(R)}^*$	1020843.33	5647.26	180.409	364.472
$\bar{y}_{G1}^*$	1019205.50	5656.34	165.866	396.429
$\bar{y}_{G2}^*$	747118.42	7716.28	159.646	411.875

The following expression is used to obtain the Percent Relative Efficiency (*PRE*), i.e.,

$$PRE = \frac{MSE(\bar{y}_0^*)}{MSE(\bar{y}_i^*)} \times 100,$$

where $i = 0, R, E, D, Rao, DR, DE, DD, DD(R), G_1, G_2$.

6. Discussion

As mentioned above, we used two real data sets to obtain the biases, MSEs or variances and PREs of all estimators under two-stage sampling scheme when using two auxiliary variables. In Tables 2–4 and Tables 5–7, we present the summary statistic of both population. From Tables 8 and 9, we observed that the proposed class of estimators $\bar{y}_{G1}^*$ and $\bar{y}_{G2}^*$ are more precise than the existing estimators $\bar{y}_0^*$, $\bar{y}_R^*$, $\bar{y}_E^*$, $\bar{y}_D^*$, $\bar{y}_{Rao}^*$, $\bar{y}_{DR}^*$, $\bar{y}_{DE}^*$, $\bar{y}_{DD}^*$, $\bar{y}_{DD(R)}^*$ in terms of MSEs and PREs. It is clear that the proposed improved generalized class of estimators, i.e., performs better than the estimators. As we increase the sample size the mean square error values decreases, and percentage relative efficiency give best results, which are the expected results.

7. Conclusions

In this manuscript, we proposed a generalized class of estimators using two auxiliary variables under two-stage sampling for estimating the finite population mean. In addition, some well-known estimators of population mean like traditional unbiased estimator, usual ratio, exponential ratio type, traditional difference type, Rao difference type, difference-in- ratio type, difference-in-exponential ratio type, difference-in-difference, difference-difference ratio type estimator are created to be members of our suggested improved generalized class of estimators. Expression for the biases and mean squared error have been generated up to the first order of approximation. We identified 11 estimators as members of the proposed class of estimators by substituting the different values of $w_i (i = 1,2,3)$, δ and γ. Both generalized class of estimators $\bar{y}_{G1}^*$ and $\bar{y}_{G2}^*$ perform better as compared to all other considered estimators, although $\bar{y}_{G2}^*$ is the best. In Population 2, the performance of ratio estimator $(\bar{y}_R^*)$ is weak. The gain in Population 1 is more as compared to Population 2.

Acknowledgments

The authors are thankful to the Editor-in-Chief and two anonymous referees for their careful reading of the paper and valuable comments which leads to a significant improvement in article.

Conflict of interest

The authors declare no conflict of interest.

References

1. M. S. Ahmed, Some estimators for a finite population mean under two-stage sampling using multivariate auxiliary information, *Appl. Math. Comput.*, **153** (2004), 505–511. https://doi.org/10.1016/S0096-3003(03)00651-9
2. S. Bahl, R. Tuteja, Ratio and product type exponential estimators, *J. Inform. Optim. Sci.*, **12** (1991), 159–164. https://doi.org/10.1080/02522667.1991.10699058
3. M. Hossain, M. S. Ahmed, A class of predictive estimators in two-stage sampling using auxiliary information, *Int. J. Inform. Manage. Sci.*, 12 (2001), 49–56.
4. S. Hussain, S. Ahmad, M. Saleem, S. Akhtar, Finite population distribution function estimation with dual use of auxiliary information under simple and stratified random sampling, *Plos One*, **15 (**2020), e0239098. https://doi.org/10.1371/journal.pone.0239098
5. S. S. Mishra, B. Pradhan, Chain ratio estimator in two stage sampling, *J. Reliab. Stat. Stud.*, **8** (2015).
6. P. Panda, *Some strategies in two stage sampling using auxiliary information*, Utkal University, 1998.
7. T. Rao, On certail methods of improving ration and regression estimators, *Commun. Stat.-Theor. M.*, **20** (1991), 3325–3340. https://doi.org/10.1080/03610929108830705
8. L. N. Sahoo, N. Mahapatra, S. C. Senapati, A generalized method of estimation for two-stage sampling using two auxiliary variables, *J. Stat. Theor. Pract.*, **3** (2009), 831–839. https://doi.org/10.1080/15598608.2009.10411963
9. S. Hussain, M. Zichuan, S. Hussain, A. Iftikhar, M. Asif, S. Akhtar, et al., On estimation of distribution function using dual auxiliary information under nonresponse using simple random sampling, *J. Probab. Stat.*, 2020, 1693612. https://doi.org/10.1155/2020/1693612
10. L. Sahoo, C. Senapati, G. N. Singh, An alternative class of estimators in two-stage sampling with two auxiliary variables, *Qual. Control Appl. Stat.*, **52** (2007), 489–490.
11. C. E. Särdnal, B. Swensson, J. Wretman, *Model assisted survey sampling*, Springer, 1992. https://doi.org/10.1007/978-1-4612-4378-6
12. T. Smith, A note on ratio estimates in multi-stage sampling, *J. R. Stat. Soc. Series A*, 1969, 426–430. https://doi.org/10.2307/2344121
13. M. Srivastava, N. Garg, A general class of estimators of a finite population mean using multi-auxiliary information under two stage sampling scheme, *J. Reliab. Stat. Stud.*, **1** (2009), 103–118.
14. H. Nawaz, M. Rabia, A. Amir, M. Aslam, Two-stage sampling exponential and chain ratio type estimators in the presence of measurement error, *J. Stat. Manag. Syst.*, 2021, 1–16. https://doi.org/10.1080/09720510.2021.1933708

15. A. Haq, M. Abbas, M. Khan, Estimation of finite population distribution function in a complex survey sampling *Commun. Stat.-Theor. M.*, 2021, 1–23. https://doi.org/10.1080/03610926.2021.1955386
16. G. K. Vishwakarma, S. M. Zeeshan, Generalized ratio-cum-product estimator for finite population mean under two-phase sampling scheme, *J. Mod. Appl. Stat. Meth.*, **19** (2021), 7. https://doi.org/10.22237/jmasm/1608553320
17. D. Singh, B. V. Sisodia, S. Nidhi, S. Pundir, Some calibration estimators for finite population mean in two-stage stratified random sampling, *Commun. Stat.-Theor. M.*, **49** (2020), 4234–4247. https://doi.org/10.1080/03610926.2019.1596283
18. A. Biswas, K. Aditya, U. C. Sud, P. Basak, Product type Calibration estimation of finite population total under two stage sampling, 2020.
19. V. I. Salinas, S. A. Sedory, S. Singh, Calibrated estimators in two-stage sampling, *Commun. Stat.-Theor. M.*, **48** (2019), 1449–1469. https://doi.org/10.1080/03610926.2018.1433850
20. N. K. Bii, C. O. Onyango, J. Odhiambo, Estimating a finite population mean under random non response in two stage cluster sampling with replacement, *J. Phys. Conf. Ser.*, **1132** (2018), 012071. https://doi.org/10.1088/1742-6596/1132/1/012071
21. S. Ahmad, S. Hussain, S. Ahmad, Finite population distribution function estimation using auxiliary information under simple random sampling, *Stat. Comput. Interdisc. Res.*, **3** (2021), 29–38. https://doi.org/10.52700/scir.v3i1.25
22. J. Shabbir, S. Gupta, Estimation of finite population mean in simple and stratified random sampling using two auxiliary variables, *Commun. Stat.-Theor. M.*, **46** (2017), 10135–10148. https://doi.org/10.1080/03610926.2016.1231822
23. H. P. Singh, S. Singh, M. Kozak, A family of estimators of finite-population distribution function using auxiliary information, *Acta Appl. Math.*, **104** (2008), 115–130. https://doi.org/10.1007/s10440-008-9243-1
24. S. K. Yadav, S. Misra, S. S. Mishra, N. Chutiman, Improved ratio estimators of population mean in adaptive cluster sampling, *J. Stat. Appl. Probab. Lett.*, **3** (2016), 1–6.
25. S. Al-Marzouki, C. Christophe, S. Akhtar, J. A. Nasir, S. Ahmad, S. Hussain, et al., Estimation of finite population mean under pps in presence of maximum and minimum values, *AIMS Math.*, **6** (2021), 5397–5409. https://doi.org/10.3934/math.2021318
26. S. Hussain, S. Ahmad, S. Akhtar, A. Javed, U. Yasmeen, Estimation of finite population distribution function with dual use of auxiliary information under non-response, *Plos One*, **15** (2020), e0243584. https://doi.org/10.1371/journal.pone.0243584
27. C. Kadilar, H. Cingi, Improvement in estimating the population mean in simple random sampling, *Appl. Math. Lett.*, **19** (2006), 75–79. https://doi.org/10.1016/j.aml.2005.02.039