Original article

Calibration under measurement errors

Gajendra K. Vishwakarma *, Abhishek Singh, Neha Singh

Department of Mathematics & Computing, Indian Institute of Technology Dhanbad, Dhanbad 826004, Jharkhand, India

ABSTRACT

The calibration technique is the process of adjusting weights for the better enhancement of the estimate of the population parameters using auxiliary information. In this manuscript, we have proposed ratio and regression type calibration estimators for the population mean under both correlated and uncorrelated measurement errors. The variances of the proposed calibrated estimators to the first order approximation are obtained and compared their efficiencies. A Monte Carlo simulation study has been carried out to show the effect of measurement errors on the proposed calibrated estimators.

© 2020 Published by Elsevier B.V. on behalf of King Saud University. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

The sample survey literature mainly comprised of estimating the population characteristics of the study variable using associated auxiliary information, the characteristics are closely related to the study variable and enhances our estimators for better efficiency. In this regard ratio, product and regression methods are good examples. This theory based on auxiliary information is further extended by using two or more than two auxiliary variables such as chain-type ratio cum regression, a linear combination of two auxiliary variables, etc. to estimate and enhance the estimator for better efficiency. These estimation procedures based on auxiliary information or merely on study variable presumed that all the data is free from error (without measurement errors). In the actual scenario, we never meet such ideality, that is the data is contaminated with or have hidden measurements error due to various kind of reasons, (see Cochran (1968) and Murthy (1967)). Measurement error is nothing, but a difference between the actual value and observed value. This can be revealed only when the measurement process is repeated or responses are compared to a gold standard (true value). There are many authors like Biemer and Stokes (1991), Shalabh (1997), Manisha and Singh (2001), Maneesha and Singh (2002), Carroll et al. (2006), Fuller (1987), Gregoire

and Salas (2009) and Kumar et al. (2011) have given the estimation procedures for the population characteristics under the measurement errors. There is a different concern that measurement errors have presented in both the study and auxiliary variable are uncorrelated. When the same surveyor or researchers are collecting information on both study and auxiliary variable with the measurement errors then there may be a correlation between measurement errors on both study and auxiliary variable. This correlated measurement errors may arise due to hidden biased tendencies of the surveyor or researchers and can affect the estimators in many ways (see Shalabh and Tsai, 2017). While dealing with sample surveys, you may face a heterogeneous population then it is convenient to use stratified sampling than to simple random sampling. In stratified sampling, we divide the heterogeneous population into homogeneous strata that enhances our estimator for better efficacy than to simple random sampling. Here, the auxiliary information is used at the estimation stages in stratified sampling to estimating the population means. As an illustration of stratified sampling, you may see Vishwakarma and Singh (2012), and Tailor et al. (2014) of estimation procedure for the means in stratified sampling.

The statisticians are always keen to find out or employ some new technique for better precision of their proposed estimators. Calibration approach is one of them for the stratified random sampling which provides better precision for the estimators using calibration weight. It is the technique of adjusting weights for estimating population characteristics associated with characteristics based on auxiliary information as constraints. The calibration technique firstly was proffered by Deville and Sarndal (1992) which minimize the distance function (chi-square) associated with constraints. The calibration approach for the combined generalized

* Corresponding author.
E-mail address: vishwagk@rediffmail.com (G.K. Vishwakarma).
Peer review under responsibility of King Saud University.

linear regression estimator (GREG) also was made known by Singh et al. (1998). Similarly, in the reference of this, many authors like Singh (2001), Tracy et al. (2003), Kim et al. (2007), Singh and Arnab (2011), Singh (2012), Singh (2013), Clement and Enang (2015), Sinha et al. (2017), Rao et al. (2017), Singh et al. (2019), Alka et al. (2019) and, Garg and Pachori (2019) have proposed different calibration approach for estimating the population characteristics associated to different calibration constraints based on auxiliary information.

Consider a finite population $U^* = (1, \ldots, N)$, for estimating the population characteristics of the variable of matter y (known to sample only) associated with auxiliary variable x, which is known to the whole population. Let u^* be the set of samples of size n drawn without replacement from a population set with some probabilistic sampling having inclusion probability $\pi_i = p(i \in u^*)$. In the support of the estimators and for the efficient estimation of the population characteristics $Y = \sum_i^N y_i$, Deville and Sarndal (1992) firstly have proposed calibration estimator of the population characteristics Y, which is established as $\hat{Y}_c = \sum_{i \in u^*} p_i y_i$, where p_i's are the calibration weights chosen to minimize their average distance from the basic design weights d_i utilized in Horvitz-Thompson estimator and equal to $1/\pi_i$. The Horvitz-Thompson estimator is as

$$\hat{Y}_{HT} = \sum_{i \in u^*} d_i y_i \tag{1}$$

associated with the constraint $\sum_{i \in u^*} p_i x_i = X$, where X is the population characteristics of the auxiliary variables. The distance measure is mostly in the chi-square form which is as

$$\varphi = \sum_{i \in u^*} \frac{(d_i - y_i)^2}{d_i q_i} \tag{2}$$

where q_i's is the positive value could be chosen suitably to estimators and also unrelated to d_i. Hence the resulting calibration estimator can be given as

$$\hat{Y}_c = \sum_{i \in u^*} p_i y_i = \hat{Y}_{HT} + \hat{B}\left(X - \hat{X}_{HT}\right) \tag{3}$$

where $\hat{B} = \left[\sum_{i \in u^*} d_i q_i \bar{x}_i^2\right]^{-1} \sum_{i \in u^*} d_i q_i \bar{x}_i \bar{y}_i$, $\hat{Y}_{HT}$ and $\hat{X}_{HT}$ are the Horvitz-Thompson estimators. The calibrated estimator $\hat{Y}_c$ is generalized by chosen different form of φ associated with the different type of constraints (See Singh et al. (1998), Singh (2003) and Tracy et al (2003)).

Further, the impact of measurement errors ridden data on the statistical properties of the estimators of parameters like population mean and population variance, cannot be avoided in the estimation until proper care is not taken. For example, in a satellite launching, the let study variable be the path reading of the satellite and speed of the satellite be the auxiliary characteristics. Second, in a vaccine testing, let study variable be the efficacy of the vaccine and time given to preparing a vaccine be the auxiliary characteristics. If there are any measurement errors then in a satellite launching mission, mission may fail or in a vaccine testing, vaccine could be pestilent. Similarly, in a making or testing of destructive bombs, in a dog fight of fighter plane or in any others exterminatory elements, measurements errors can't be ignored, because "a miss is as good as a mile".

We have used some existing basic calibration estimation and the major disadvantages of this existing research is that it doesn't count seriousness of the measurement errors. To tackle this seriousness or problems along with the variation of measurement errors. In this manuscript, we have explained the calibration technique under both correlated and uncorrelated measurement errors in stratified random sampling to examine the effect of measurement errors on the proposed estimators under different calibration weights. Here, we have described proposed estimators under two

different calibration weights, in the first case, it reduced to ratio form and while in the second case it reduced to regression form. A simulation study is conducted to examine the properties of proposed estimators under both cases in measurement errors (correlated and uncorrelated) and without measurements errors in stratified sampling. Also, the efficiency comparisons have been given to examine the precision of the proposed estimators under measurements errors in stratified sampling using different calibration weights over unbiased estimator in stratified sampling. The effect of the measurement errors has given as percentage contribution of measurement errors (PCME). Further, this manuscript followed by as: proposed calibration estimators are given in Section 2, Section 3 consists of efficiency comparisons, Section 4 describes the compilation of simulation study, Section 5 describes the discussion and Section 6 mention the conclusions and further study.

2. Proposed ratio and regression type calibrated estimators under measurement errors

Let us consider a finite heterogeneous population of size N, partitioned into L non-overlapping strata of sizes N_h, h = 1,2, ..., L, where $\sum_{h=1}^L N_h = N$. Let (y_{wehj}, x_{wehj}) be the pair of observed values instead of true pair values (y_{thj}, x_{thj}) of the study character y and the auxiliary character x respectively of the jth unit (j = 1, 2, ..., N_h) in the hth stratum. Also, let (y_{wehj}, x_{wehj}) be the pair of values on (y, x) drawn from the hth stratum (j = 1, 2, ..., n_h; h = 1, 2, ..., L). It is familiar that in stratified random sampling an unbiased estimator of the population means $\mu_{weY} = \sum_{h=1}^L W_h \mu_{weY_h}, \mu_{weX} = \sum_{h=1}^L W_h \mu_{weX_h}, \mu_{tY} = \sum_{h=1}^L W_h \mu_{tY_h}, \mu_{tX} = \sum_{h=1}^L W_h \mu_{tX_h}$; $W_h = \frac{N_h}{N}$ of the variable y and x are given by $\bar{y}_{west} = \sum_{h=1}^L W_h \bar{y}_{weh}, \bar{x}_{west} = \sum_{h=1}^L W_h \bar{x}_{weh}$ and $\bar{y}_{tst} = \sum_{h=1}^L W_h \bar{y}_{th}, \bar{x}_{tst} = \sum_{h=1}^L W_h \bar{x}_{th}$, where $\bar{y}_{weh} = \frac{1}{n_h} \sum_{j=1}^{n_h} y_{wehj}, \bar{x}_{weh} = \frac{1}{n_h} \sum_{j=1}^{n_h} x_{wehj}$, and $\bar{y}_{weh} = \frac{1}{n_h} \sum_{j=1}^{n_h} y_{wehj}, \bar{x}_{weh} = \frac{1}{n_h} \sum_{j=1}^{n_h} x_{wehj}$ are the sample mean of hth stratum of observed value and true value, and $\mu_{weY_h} = \frac{1}{N_h} \sum_{j=1}^{N_h} Y_{wehj}, \mu_{weX_h} = \frac{1}{N_h} \sum_{j=1}^{N_h} X_{wehj}$ and $\mu_{tY_h} = \frac{1}{N_h} \sum_{j=1}^{N_h} Y_{thj}, \mu_{tX_h} = \frac{1}{N_h} \sum_{j=1}^{N_h} X_{thj}$ are the population mean of hth stratum of the observed and true value. Let the observational errors be $u_{hj} = (x_{wehj} - x_{thj})$, $v_{hj} = (y_{wehj} - y_{thj})$, which are normally distributed with mean zero and variances σ_{uh}^2 and σ_{vh}^2 respectively. Also let $\rho_{tX_h Y_h}$ be the population correlation coefficient between Y_t and X_t in h^{th} stratum. For simplicity in exposition, assume that the variables u_{hj} and v_{hj} are correlated with correlation coefficient $\rho_{U_h V_h}$ as (y_{thj}, x_{thj}) are correlated. The usual mean estimator and its variance under measurement error are given by

$$\widehat{\bar{y}}_{west} = \sum_{h=1}^L W_h \bar{y}_{weh} \tag{4}$$

$$V\left(\widehat{\bar{y}}_{west}\right) = \sum_{h=1}^L W_h^2 \frac{\sigma_{weYh}^2}{n_h} = \sigma_{west}^2 (\text{say}), \tag{5}$$

where $\sigma_{weYh}^2 = \frac{1}{N_h - 1} \sum_{j=1}^{N_h} \left(Y_{wehj} - \bar{\mu}_{weY_h}\right)^2 = \sigma_{tY_h}^2 + \sigma_{V_h}^2 = \frac{1}{N_h - 1} \sum_{j=1}^{N_h} \left(Y_{thj} - \bar{\mu}_{tY_h}\right)^2 + \frac{1}{N_h - 1} \sum_{j=1}^{N_h} \left(V_{hj} - \bar{V}_h\right)^2$ is the hth stratum population

variance. the unbiased estimate of this variance σ^2_{west} can be written as

$$\hat{V}\left(\widehat{\bar{y}}_{west}\right) = \sum_{h=1}^{L} W_h^2 \frac{\hat{\sigma}^2_{weYh}}{n_h} = \hat{\sigma}^2_{west} \tag{6}$$

where $\hat{\sigma}^2_{weyh} = \frac{1}{n_h-1}\sum_{j=1}^{n_h}\left(y_{wehj} - \bar{y}_{weh}\right)^2 = \sigma^2_{tyh} + \sigma^2_{vh} = \frac{1}{n_h-1}\sum_{j=1}^{n_h}\left(y_{thj} - \bar{y}_{th}\right)^2 + \frac{1}{n_h-1}\sum_{j=1}^{n_h}\left(v_{hj} - \bar{v}_h\right)^2$ is the hth stratum sample variance, $\bar{V}_h \approx 0 \approx \bar{U}_h$, $\bar{\mu}_{tYh}, \bar{y}_{th}, \bar{u}_h$ and $\bar{v}_h$ are population stratum mean, sample stratum mean and sample mean of the errors.

To obtaining the expression up to first degree of approximation, we have defined some error terms are given by as:

$e_{h0} = \frac{\left(\bar{y}_{weh} - \mu_{tY_h}\right)}{\mu_{tY_h}}$, $e_{h1} = \frac{\left(\bar{x}_{weh} - \mu_{tX_h}\right)}{\mu_{tX_h}}$ are such that $E(e_{h0}) = 0 = E(e_{h1})$ with $E(e_{h0}^2) = \frac{1}{n_h}\left(C_{tY_h}^2 + C_{V_h}^2\right)$,

$E(e_{h1}^2) = \frac{1}{n_h}\left(C_{tX_h}^2 + C_{U_h}^2\right)$ and $E(e_{h0}e_{h1}) = \frac{1}{n_h}\left(\rho_{tXY_h}C_{tX_h}C_{tY_h} + \rho_{UV_h}C_{U_h}C_{V_h}\right)$. Where $C_{tY_h} = \frac{\sigma_{tY_h}}{\mu_{tY_h}}$, $C_{tX_h} = \frac{\sigma_{tX_h}}{\mu_{tX_h}}$ and $C_{V_h} = \frac{\sigma_{V_h}}{\mu_{tY_h}}$, $C_{U_h} = \frac{\sigma_{U_h}}{\mu_{tX_h}}$.

Thus, variance of the estimator up to first order of approximation is given by:

$$V\left(\widehat{\bar{y}}_{west}\right) = \sum_{h=1}^{L} W_h^2 \mu_{tY_h}^2 \frac{\left(C_{tY_h}^2 + C_{V_h}^2\right)}{n_h} \tag{7}$$

which is same as in Eq. (5). Now, we are going to define new proposed calibration estimators instead of $\left(\widehat{\bar{y}}_{west}\right)$ under measurement errors, here calibration estimators are as

Case 1: Ratio form

The estimator based on Kim et al. (2007) for the population mean in stratified random sampling under measurement errors with new calibration weight W'_h is given as

$$\widehat{\bar{y}}'_{west} = \sum_{h=1}^{L} W'_h \bar{y}_{weh} \tag{8}$$

where W'_h are chosen such that the chi-square-type distance function

$$\sum_{h=1}^{L} \frac{\left(W'_h - W_h\right)^2}{Q_h W_h} \tag{9}$$

is minimized associated to constraint

$$\sum_{h=1}^{L} W'_h \bar{x}_{weh} = \sum_{h=1}^{L} W_h \mu_{weX_h} = \mu_{weX} \tag{10}$$

where Q_h provide form of calibration estimator and can be arbitrary chosen.

Then we have the following theorem

Theorem 1. *A calibrated estimator for the population mean under measurement errors along with calibrated weight W'_h is given by:*

$$\widehat{\bar{y}}'_{west} = \sum_{h=1}^{L} W_h \bar{y}_h + \hat{\beta}'_{we} \sum_{h=1}^{L} W_h \left(\mu_{weX_h} - \bar{x}_h\right) \tag{11}$$

where $\hat{\beta}'_{we} = \sum_{h=1}^{L} Q_h W_h \bar{x}_{weh} \bar{y}_{weh} / \sum_{h=1}^{L} Q_h W_h \bar{x}_{weh}^2$ and $\sum_{h=1}^{L} W'_h \bar{x}_h = \sum_{h=1}^{L} W_h \mu_{weX_h}$ with

$\bar{x}_{weh} = \bar{x}_{th} + \bar{u}_h, \bar{y}_{weh} = \bar{y}_{th} + \bar{v}_h, \mu_{weX_h} = \mu_{tX_h}, \mu_{weY_h} = \mu_{tY_h}$

Proof. By defining λ_a as Lagrange multiplier, and from Eqs. (8)–(10), we have Lagrange function as

$$\phi_1 = \sum_{h=1}^{L} \frac{\left(W'_h - W_h\right)^2}{Q_h W_h} - 2\lambda_a \left(\sum_{h=1}^{L} W'_h \bar{x}_{weh} - \sum_{h=1}^{L} W_h \bar{x}_{weh}\right) \tag{12}$$

Differentiating ϕ_1 in Eq. (12) with respect to calibration weight W'_h and equating to equal to 0, we have

$$W'_h = W_h + \lambda_a Q_h W_h \bar{x}_{weh} \tag{13}$$

Multiplying above Eq. (13) by $\bar{x}_{weh}$ on both side and taking summation we have

$$\sum_{h=1}^{L} W'_h \bar{x}_{weh} = \sum_{h=1}^{L} W_h \bar{x}_{weh} + \lambda_a \sum_{h=1}^{L} Q_h W_h \bar{x}_{weh}^2 \tag{14}$$

solving above Eq. (14), we have Lagrange multiplier as

$$\lambda_a = \frac{\left(\sum_{h=1}^{L} W'_h \bar{x}_{weh} - \sum_{h=1}^{L} W_h \bar{x}_{weh}\right)}{\sum_{h=1}^{L} Q_h W_h \bar{x}_{weh}^2} \tag{15}$$

and by putting the value of λ_a in Eq. (13), we have calibrated weight as

$$W'_h = W_h + \frac{Q_h W_h \bar{x}_{weh}}{\sum_{h=1}^{L} Q_h W_h \bar{x}_{weh}^2}\left[\sum_{h=1}^{L} W'_h \bar{x}_{weh} - \sum_{h=1}^{L} W_h \bar{x}_{weh}\right] \tag{16}$$

and hence from Eqs. (8) and (16), we have calibrated estimator as

$$\widehat{\bar{y}}'_{west} = \sum_{h=1}^{L} W_h \bar{y}_{weh}$$
$$+ \frac{\sum_{h=1}^{L} Q_h W_h \bar{x}_{weh} \bar{y}_{weh}}{\sum_{h=1}^{L} Q_h W_h \bar{x}_{weh}^2}\left[\sum_{h=1}^{L} W_h \mu_{weX_h} - \sum_{h=1}^{L} W_h \bar{x}_{weh}\right] \tag{17}$$

finally, new estimator under calibration weight W'_h can be written as

$$\widehat{\bar{y}}'_{west} = \sum_{h=1}^{L} W_h \bar{y}_h + \hat{\beta}'_{we} \sum_{h=1}^{L} W_h \left(\mu_{weX_h} - \bar{x}_h\right) \tag{18}$$

Also, setting $Q_h = \left(\bar{x}_{weh}\right)^{-1}$, we get combined ratio estimator in stratified random sampling under measurement errors as

$\widehat{\bar{y}}'_{west} = \left(\sum_{h=1}^{L} W_h \bar{y}_{weh} / \sum_{h=1}^{L} W_h \bar{x}_{weh}\right)\mu_{weX}$, hence ratio form.

Theorem 2. *The variance of the estimator $\widehat{\bar{y}}'_{west}$ under measurement errors, to the first order approximation is given by:*

$$V\left(\widehat{\bar{y}}'_{west}\right) = \left(\frac{\mu_{tX_h}}{\bar{x}_{tst}}\right)^2 \sum_{h=1}^{L} W_h^2 \mu_{tY_h}^2 \frac{\left(C_{tY_h}^2 + C_{V_h}^2\right)}{n_h} \tag{19}$$

Proof. From the Eq. (8), we have

$$\widehat{\bar{y}}'_{west} = \sum_{h=1}^{L} W'_h \bar{y}_{weh} \text{ and}$$

$$V\left(\widehat{\overline{y}}'_{west}\right) = V\left(\sum_{h=1}^{L} W'_h \bar{y}_{weh}\right) = \sum_{h=1}^{L} W_h'^2 V\left(\bar{y}_{weh}\right)$$

$$= \sum_{h=1}^{L} W_h'^2 V(1 + e_{h0}) = \sum_{h=1}^{L} W_h'^2 E(e_{h0}^2) \qquad (20)$$

$$= \sum_{h=1}^{L} W_h'^2 \mu_{tY_h}^2 \frac{\left(C_{tY_h}^2 + C_{V_h}^2\right)}{n_h}$$

From Eq. (16) we have

$$W'_h = W_h + \frac{Q_h W_h \bar{x}_{weh}}{\sum_{h=1}^{L} Q_h W_h \bar{x}_{weh}^2} \left[\sum_{h=1}^{L} W'_h \bar{x}_{weh} - \sum_{h=1}^{L} W_h \bar{x}_{weh}\right] \qquad (21)$$

and also, from Eq. (10) we have

$$\sum_{h=1}^{L} W'_h \bar{x}_{weh} = \sum_{h=1}^{L} W_h \mu_{weX_h} = \mu_{weX} = \mu_{tX}, \text{ so, we have Eq. (21) as}$$
following

$$W'_h = W_h + \frac{Q_h W_h \bar{x}_{weh}}{\sum_{h=1}^{L} Q_h W_h \bar{x}_{weh}^2} \left[\mu_{tX} - \sum_{h=1}^{L} W_h \bar{x}_{weh}\right] \qquad (22)$$

On setting $Q_h = \left(\bar{x}_{weh}\right)^{-1}$ in Eq. (22), we have new weight as

$$W'_h = \frac{W_h \mu_{tX}}{\sum_{h=1}^{L} W_h \bar{x}_{weh}} = \frac{W_h \mu_{tX}}{\bar{x}_{west}} = \frac{W_h \mu_{tX}}{\bar{x}_{tst}}, \text{ since, now substituting this } W'$$
in Eq. (20),

We have variance of the calibrated estimator $\widehat{\overline{y}}'_{west}$ as following

$$V\left(\widehat{\overline{y}}'_{west}\right) = \left(\frac{\mu_{tX_h}}{\bar{x}_{tst}}\right)^2 \sum_{h=1}^{L} W_h^2 \mu_{tY_h}^2 \frac{\left(C_{tY_h}^2 + C_{V_h}^2\right)}{n_h}, \text{ which proves the the-}$$
orem 2.

Lemma 1. The variance of the estimator $\widehat{\overline{y}}'_{west}$ without measurement errors, to the first order approximation is given by:

$$V\left(\widehat{\overline{y}}'_{west}\right) = V\left(\widehat{\overline{y}}'_{west}\right) = \left(\frac{\mu_{tX_h}}{\bar{x}_{tst}}\right)^2 \sum_{h=1}^{L} W_h^2 \mu_{tY_h}^2 \frac{\left(C_{tY_h}^2\right)}{n_h} \qquad (23)$$

Proof. From the theorem 2, we have variance of the estimator $\widehat{\overline{y}}'_{west}$ under measurement errors as following

$$V\left(\widehat{\overline{y}}'_{west}\right) = \left(\frac{\mu_{tX_h}}{\bar{x}_{tst}}\right)^2 \sum_{h=1}^{L} W_h^2 \mu_{tY_h}^2 \frac{\left(C_{tY_h}^2 + C_{V_h}^2\right)}{n_h}$$

If there are no measurement errors then in the expression of the variance of the estimator $\widehat{\overline{y}}'_{west}$, the coefficient of variation C_{V_h} will be zero and the variance of the estimator $\widehat{\overline{y}}'_{west}$ will be now as following

$$V\left(\widehat{\overline{y}}'_{west}\right) = V\left(\widehat{\overline{y}}'_{west}\right) = \left(\frac{\mu_{tX_h}}{\bar{x}_{tst}}\right)^2 \sum_{h=1}^{L} W_h^2 \mu_{tY_h}^2 \frac{\left(C_{tY_h}^2\right)}{n_h} \text{ this proves}$$
lemma1.

Lemma 2. The variance of the estimator $\widehat{\overline{y}}'_{west}$ under measurement errors, to the first order approximation has greater variance than to variance of the estimator $\widehat{\overline{y}}'_{west}$ without measurement errors:

$$V\left(\widehat{\overline{y}}'_{west}\right) > V\left(\widehat{\overline{y}}'_{tst1}\right) \qquad (24)$$

Proof. From the theorem 2 and lemma 1, the variance of the estimator $\widehat{\overline{y}}'_{west}$ under measurement errors in Eq. (19) can be expressed as following

$$V\left(\widehat{\overline{y}}'_{west}\right) = V\left(\widehat{\overline{y}}'_{tst}\right) + \left(\frac{\mu_{tX_h}}{\bar{x}_{tst}}\right)^2 \sum_{h=1}^{L} W_h^2 \mu_{tY_h}^2 \frac{\left(C_{V_h}^2\right)}{n_h} \qquad (25)$$

That is $V\left(\widehat{\overline{y}}'_{west}\right) > V\left(\widehat{\overline{y}}'_{tst}\right)$, hence the proof of the lemma 2.

Case 2: Regression form

Again, we are considering here new estimator, and this estimator based on Sinha et al. (2017) for the population mean in stratified random sampling under measurement errors along with weight W_h'' and is given by:

$$\widehat{\overline{y}}''_{west} = \sum_{h=1}^{L} W''_h \bar{y}_{weh} \qquad (26)$$

where W_h'' are chosen such that the chi-square-type distance

$$\sum_{h=1}^{L} \frac{\left(W_h'' - W_h\right)^2}{Q_h W_h} \qquad (27)$$

is minimized s. to c.

$$\sum_{h=1}^{L} W''_h \bar{x}_{weh} = \sum_{h=1}^{L} W_h \mu_{weX_h} = \mu_{weX} \text{ and } \sum_{h=1}^{L} W''_h = \sum_{h=1}^{L} W_h = 1 \qquad (28)$$

Theorem 3. A calibrated estimator for the population mean under measurement errors along with calibrated weight W_h'' is given by:

$$\widehat{\overline{y}}''_{west} = \sum_{h=1}^{L} W_h \bar{y}_{weh} + \hat{\beta}''_{we} \sum_{h=1}^{L} W_h \left(\mu_{tX_h} - \bar{x}_{weh}\right) \qquad (29)$$

where $\hat{\beta}''_{we} = \dfrac{\left[\left(\sum_{h=1}^{L} Q_h W_h \bar{x}_{weh} \bar{y}_{weh}\right)\left(\sum_{h=1}^{L} Q_h W_h\right) - \left(\sum_{h=1}^{L} Q_h W_h \bar{y}_{weh}\right)\left(\sum_{h=1}^{L} Q_h W_h \bar{x}_{weh}\right)\right]}{\left[\left(\sum_{h=1}^{L} Q_h W_h\right)\left(\sum_{h=1}^{L} Q_h W_h \bar{x}_{weh}^2\right) - \left(\sum_{h=1}^{L} Q_h W_h \bar{x}_{weh}\right)^2\right]}$

Proof. By defining λ_b and λ_c as Lagrange multiplier, and from Eqs. (26) to (28), we have Lagrange function as

$$\phi_2 = \sum_{h=1}^{L} \frac{\left(W_h'' - W_h\right)^2}{Q_h W_h} - 2\lambda_b \left(\sum_{h=1}^{L} W''_h \bar{x}_{weh} - \sum_{h=1}^{L} W_h \bar{x}_{weh}\right)$$

$$- 2\lambda_c \left(\sum_{h=1}^{L} W''_h - \sum_{h=1}^{L} W_h\right) \qquad (30)$$

Differentiating ϕ_2 in Eq. (30) with respect to calibration weight W_h'' and equating to equal to 0, we have

$$W''_h = W_h \left[1 + \lambda_b Q_h \bar{x}_h + \lambda_c Q_h\right] \qquad (31)$$

the following equations are obtained solving above Eq. (31) by multiplying $\bar{x}_{weh}$ and over summation

$$\sum_{h=1}^{L} W''_h = \sum_{h=1}^{L} W_h + \lambda_b \sum_{h=1}^{L} Q_h W_h \bar{x}_{weh} + \lambda_c \sum_{h=1}^{L} Q_h W_h \qquad (32)$$

$$\sum_{h=1}^{L} W''_h \bar{x}_{weh} = \sum_{h=1}^{L} W_h \bar{x}_{weh} + \lambda_b \sum_{h=1}^{L} Q_h W_h \bar{x}^2_{weh} + \lambda_c$$
$$\times \sum_{h=1}^{L} Q_h W_h \bar{x}_{weh} \qquad (33)$$

by substituting the $\sum_{h=1}^{L} W''_h \bar{x}_{weh} = \sum_{h=1}^{L} W_h \mu_{weX_h}$, $\sum_{h=1}^{L} W''_h = \sum_{h=1}^{L} W_h = 1$ and solving Eqs. (32) and (33), we have

$$\lambda_b = \sum_{h=1}^{L} Q_h W_h \left[\sum_{h=1}^{L} W_h \mu_{weX_h} - \sum_{h=1}^{L} W_h \bar{x}_{weh} \right] /$$
$$\left[\left(\sum_{h=1}^{L} Q_h W_h \right) \left(\sum_{h=1}^{L} Q_h W_h \bar{x}^2_{weh} \right) - \left(\sum_{h=1}^{L} Q_h W_h \bar{x}_{weh} \right)^2 \right]$$

$$\lambda_c = - \left[\sum_{h=1}^{L} Q_h W_h \bar{x}_{weh} \left(\sum_{h=1}^{L} W_h \mu_{weX_h} - \sum_{h=1}^{L} W_h \bar{x}_{weh} \right) \right] /$$
$$\left[\left(\sum_{h=1}^{L} Q_h W_h \right) \left(\sum_{h=1}^{L} Q_h W_h \bar{x}^2_{weh} \right) - \left(\sum_{h=1}^{L} Q_h W_h \bar{x}_{weh} \right)^2 \right]$$

putting both the value of λ_b and λ_c in Eq. (31), we have new calibration weight as

$$W''_h = W_h$$
$$+ \frac{\left[\left\{ Q_h W_h \bar{x}_{weh} \left(\sum_{h=1}^{L} Q_h W_h \right) - Q_h W_h \left(\sum_{h=1}^{L} Q_h W_h \bar{x}_{weh} \right) \right\} \left\{ \sum_{h=1}^{L} W_h \left(\mu_{weX_h} - \bar{x}_{weh} \right) \right\} \right]}{\left[\left(\sum_{h=1}^{L} Q_h W_h \right) \left(\sum_{h=1}^{L} Q_h W_h \bar{x}^2_{weh} \right) - \left(\sum_{h=1}^{L} Q_h W_h \bar{x}_{weh} \right)^2 \right]}$$
$$(34)$$

and the new estimator under calibration weight W''_h can be written as

$$\widehat{\bar{y}}''_{west} = \sum_{h=1}^{L} W_h \bar{y}_{weh} + \hat{\beta}''_{we} \sum_{h=1}^{L} W_h \left(\mu_{weX_h} - \bar{x}_{weh} \right) \qquad (35)$$

$$\hat{\beta}''_{we} = \frac{\left[\left(\sum_{h=1}^{L} Q_h W_h \bar{x}_{weh} \bar{y}_{weh} \right) \left(\sum_{h=1}^{L} Q_h W_h \right) - \left(\sum_{h=1}^{L} Q_h W_h \bar{y}_{weh} \right) \left(\sum_{h=1}^{L} Q_h W_h \bar{x}_{weh} \right) \right]}{\left[\left(\sum_{h=1}^{L} Q_h W_h \right) \left(\sum_{h=1}^{L} Q_h W_h \bar{x}^2_{weh} \right) - \left(\sum_{h=1}^{L} Q_h W_h \bar{x}_{weh} \right)^2 \right]}$$

hence, proves the theorem 3. Also, when we put the value $Q_h = 1$ in Eq. (35) then we have calibrated estimator in the regression form as

$$\widehat{\bar{y}}''_{west} = \sum_{h=1}^{L} W_h \bar{y}_{weh}$$
$$+ \frac{\left[\left(\sum_{h=1}^{L} W_h \bar{x}_{weh} \bar{y}_{weh} \right) - \left(\sum_{h=1}^{L} W_h \bar{y}_{weh} \right) \left(\sum_{h=1}^{L} W_h \bar{x}_{weh} \right) \right]}{\left[\left(\sum_{h=1}^{L} W_h \bar{x}^2_{weh} \right) - \left(\sum_{h=1}^{L} W_h \bar{x}_{weh} \right)^2 \right]}$$
$$\times \sum_{h=1}^{L} W_h \left(\mu_{weX_h} - \bar{x}_{weh} \right)$$
$$(36)$$

The estimator in Eq. (35) has another form is given by:

$$\widehat{\bar{y}}''_{west1} = \sum_{h=1}^{L} W_h \bar{y}_{weh} + \beta''_{we} \sum_{h=1}^{L} W_h \left(\mu_{weX_h} - \bar{x}_{weh} \right) \qquad (37)$$

where is a constant to be determined such that the variance of the estimator $\widehat{\bar{y}}''_{west}$ is minimum. The regression coefficient $\hat{\beta}''_{we}$ is considered as consistent estimator of β''_{we}. Therefore, the variance of

the calibrated estimator $\widehat{\bar{y}}''_{west}$ under measurement errors, up to first order approximation, is asymptotically equivalent to the variance of the estimator $\widehat{\bar{y}}''_{west}$ (see Salinas et al., 2019).

Theorem 4. *The variance of the calibrated estimator $\widehat{\bar{y}}''_{west}$ under correlated measurement errors, to the first order approximation is given by:*

$$V\left(\widehat{\bar{y}}''_{west1} \right)_{\min} = \sum_{h=1}^{L}$$
$$\times \frac{W_h^2 \mu_{tY_h}^2}{n_h} \left\{ \left(C_{tY_h}^2 + C_{V_h}^2 \right) - \frac{\left(\rho_{tXY_h} C_{tX_h} C_{tY_h} + \rho_{UV_h} C_{U_h} C_{V_h} \right)^2}{\left(C_{tX_h}^2 + C_{U_h}^2 \right)} \right\}$$
$$(38)$$

Proof. From the Eq. (37) and following the approach of Koyuncu and Kadilar (2014), and Salinas et al. (2019), we have variance of the estimator $\widehat{\bar{y}}''_{west}$ as

$$V\left(\widehat{\bar{y}}''_{west1} \right) = V\left[\sum_{h=1}^{L} W_h \bar{y}_{weh} + \beta''_{we} \sum_{h=1}^{L} W_h \left(\mu_{weX_h} - \bar{x}_{weh} \right) \right]$$
$$= \left[\sum_{h=1}^{L} W_h^2 \left\{ \mu_{tY_h}^2 E(e_{h0}^2) + \beta''_{we} 2 \mu_{tX_h}^2 E(e_{h1}^2) - 2\beta''_{we} \mu_{tY_h} \mu_{tX_h} E(e_{h0} e_{h1}) \right\} \right]$$
$$= \left[\sum_{h=1}^{L} \frac{W_h^2}{n_h} \left\{ \mu_{tY_h}^2 \left(C_{tY_h}^2 + C_{V_h}^2 \right) + \beta''_{we} 2 \mu_{tX_h}^2 \left(C_{tX_h}^2 + C_{U_h}^2 \right) \right. \right.$$
$$\left. \left. - 2\beta''_{we} \mu_{tY_h} \mu_{tX_h} \left(\rho_{tXY_h} C_{tX_h} C_{tY_h} + \rho_{UV_h} C_{U_h} C_{V_h} \right) \right\} \right]$$
$$(39)$$

To obtaining the optimum value of for the minimum variance of the estimator $\widehat{\bar{y}}''_{west}$, we are differentiating $V\widehat{\bar{y}}''_{west}$ along with Eq. (39) with respect to equating it equal to zero, that is $\frac{\partial \left(V\left(\widehat{\bar{y}}''_{west} \right) \right)}{\partial \beta''_{we}} = 0$, we have

$$\beta''_{we} = \frac{\mu_{tY_h}}{\mu_{tX_h}} \frac{\left(\rho_{tXY_h} C_{tX_h} C_{tY_h} + \rho_{UV_h} C_{U_h} C_{V_h} \right)}{\left(C_{tX_h}^2 + C_{U_h}^2 \right)} \qquad (40)$$

Putting the value of from the Eq. (40) in $V\left(\widehat{\bar{y}}''_{west} \right)$ along with Eq. (39), we have minimum variance of the calibrated estimator $\widehat{\bar{y}}''_{west1}$ as

$$V\left(\widehat{\bar{y}}''_{west1} \right)_{\min} = \sum_{h=1}^{L}$$
$$\times \frac{W_h^2 \mu_{tY_h}^2}{n_h} \left\{ \left(C_{tY_h}^2 + C_{V_h}^2 \right) - \frac{\left(\rho_{tXY_h} C_{tX_h} C_{tY_h} + \rho_{UV_h} C_{U_h} C_{V_h} \right)^2}{\left(C_{tX_h}^2 + C_{U_h}^2 \right)} \right\}$$

hence the proof of the theorem 4.

Lemma 3. *The variance of the estimator $\widehat{\bar{y}}'_{west1}$ under uncorrelated measurement errors, to the first order approximation is given by:*

$$V\left(\widehat{\bar{y}}''_{west1} \right)_{\min} = \sum_{h=1}^{L}$$
$$\times \frac{W_h^2 \mu_{tY_h}^2}{n_h} \left\{ \left(C_{tY_h}^2 + C_{V_h}^2 \right) - \frac{\left(\rho_{tXY_h} C_{tX_h} C_{tY_h} \right)^2}{\left(C_{tX_h}^2 + C_{U_h}^2 \right)} \right\}$$
$$(41)$$

Proof. From the theorem 4, we have variance of the estimator $\widehat{\bar{y}}''_{west1}$ under correlated measurement errors as following

$$V\left(\widehat{\bar{y}}''_{west}\right)_{\min} = \sum_{h=1}^{L} \frac{W_h^2 \mu_{tY_h}^2}{n_h} \left\{ \left(C_{tY_h}^2 + C_{V_h}^2\right) - \frac{\left(\rho_{tXY_h} C_{tX_h} C_{tY_h} + \rho_{UV_h} C_{U_h} C_{V_h}\right)^2}{\left(C_{tX_h}^2 + C_{U_h}^2\right)} \right\}$$

If there are uncorrelated measurement errors then in the expression of the variance of the estimator $\widehat{\bar{y}}''_{west1}$, the correlation coefficient ρ_{UV_h} will be zero and the variance of the estimator $\widehat{\bar{y}}''_{west1}$ will be now as following

$$V\left(\widehat{\bar{y}}''_{west1}\right)_{\min} = \sum_{h=1}^{L} \frac{W_h^2 \mu_{tY_h}^2}{n_h} \left\{ \left(C_{tY_h}^2 + C_{V_h}^2\right) - \frac{\left(\rho_{tXY_h} C_{tX_h} C_{tY_h}\right)^2}{\left(C_{tX_h}^2 + C_{U_h}^2\right)} \right\} \quad \text{this}$$

proves lemma3.

Lemma 4. The variance of the estimator $\widehat{\bar{y}}'_{west1}$ without measurement errors, to the first order approximation is given by:

$$V\left(\widehat{\bar{y}}''_{west1}\right)_{\min} = V\left(\widehat{\bar{y}}''_{tst1}\right)_{\min} = \sum_{h=1}^{L} \frac{W_h^2 \mu_{tY_h}^2 C_{tY_h}^2}{n_h} \left(1 - \rho_{tXY_h}^2\right) \quad (42)$$

Proof. From the theorem 4, we have variance of the estimator $\widehat{\bar{y}}''_{west1}$ under correlated measurement errors as following

$$V\left(\widehat{\bar{y}}''_{west1}\right)_{\min} = \sum_{h=1}^{L} \frac{W_h^2 \mu_{tY_h}^2}{n_h} \left\{ \left(C_{tY_h}^2 + C_{V_h}^2\right) - \frac{\left(\rho_{tXY_h} C_{tX_h} C_{tY_h} + \rho_{UV_h} C_{U_h} C_{V_h}\right)^2}{\left(C_{tX_h}^2 + C_{U_h}^2\right)} \right\}$$

If there are no measurement errors then in the expression of the variance of the estimator $\widehat{\bar{y}}''_{west1}$, the correlation coefficient ρ_{UV_h} and coefficient of variations C_{U_h}, C_{V_h} will be zero, and the variance of the estimator $\widehat{\bar{y}}''_{west1}$ will be now as following

$$V\left(\widehat{\bar{y}}''_{west1}\right)_{\min} = V\left(\widehat{\bar{y}}''_{tst1}\right)_{\min} = \sum_{h=1}^{L} \frac{W_h^2 \mu_{tY_h}^2 C_{tY_h}^2}{n_h} \left(1 - \rho_{tXY_h}^2\right) \quad \text{this}$$

proves lemma3.

The expression for variance in lemma 4 of Eq. (42), is a variance of the classical linear regression under stratified sampling when there are no measurement errors. Similarly, the calibrated estimator $\widehat{\bar{y}}''_{west1}$ from the Eq. (36) under measurement errors will be classical linear regression estimator under stratified sampling if there are no measurement errors. Therefore, this also shows that the variance of the estimator $\widehat{\bar{y}}''_{west1}$ from the Eq. (37) will be equivalent to the variance of the estimator $\widehat{\bar{y}}''_{west1}$ from the Eq. (36) under both correlated and uncorrelated measurements errors. That is

$$V\left(\widehat{\bar{y}}''_{west1}\right)_{\min} \approx V\left(\widehat{\bar{y}}''_{west}\right)_{\min} = \sum_{h=1}^{L} \frac{W_h^2 \mu_{tY_h}^2}{n_h}$$
$$\left\{ \left(C_{tY_h}^2 + C_{V_h}^2\right) - \frac{\left(\rho_{tXY_h} C_{tX_h} C_{tY_h} + \rho_{UV_h} C_{U_h} C_{V_h}\right)^2}{\left(C_{tX_h}^2 + C_{U_h}^2\right)} \right\} \quad (43)$$

Lemma 5. The variance of the estimator $\widehat{\bar{y}}''_{west1}$ under measurement errors, to the first order approximation has greater variance than to variance of the estimator $\widehat{\bar{y}}''_{west1}$ without measurement errors

$$V\left(\widehat{\bar{y}}''_{west1}\right) > V\left(\widehat{\bar{y}}''_{tst1}\right) \quad (44)$$

Proof. From the lemma 3 and lemma 4, the variance of the estimator $\widehat{\bar{y}}''_{west1}$ under measurement errors in Eq. (41) can be expressed as following

$$V\left(\widehat{\bar{y}}''_{west1}\right) = V\left(\widehat{\bar{y}}''_{tst1}\right) + \sum_{h=1}^{L} \frac{W_h^2 \mu_{tY_h}^2}{n_h} \left(C_{V_h}^2 - \rho_{tXY_h}^2 C_{tY_h}^2\right) \quad (45)$$

That is $V\left(\widehat{\bar{y}}''_{west1}\right) > V\left(\widehat{\bar{y}}''_{tst1}\right)$, hence the proof of the lemma 5.

3. Efficiency comparisons

In this section, we are giving efficiency comparisons of the calibrated estimators $\widehat{\bar{y}}'_{west}$ and $\widehat{\bar{y}}''_{west1}$ over unbiased estimator $\widehat{\bar{y}}_{west}$ under measurement errors.

(i) $V\left(\widehat{\bar{y}}'_{west}\right) < V\left(\widehat{\bar{y}}_{west}\right)$ iff

$$\left\{ V\left(\widehat{\bar{y}}_{west}\right) - V\left(\widehat{\bar{y}}'_{west}\right) \right\} > 0 \quad (46)$$

from the Eqs. (19) and (7), the variance of the estimator $\widehat{\bar{y}}'_{west}$ can also be expressed as following

$$V\left(\widehat{\bar{y}}'_{west}\right) = \left(\frac{\mu_{tX_h}}{\bar{x}_{tst}}\right)^2 V\left(\widehat{\bar{y}}_{west}\right) \text{ i.e.}$$
$$V\left(\widehat{\bar{y}}'_{west}\right) = V\left(\widehat{\bar{y}}_{west}\right) - V\left(\widehat{\bar{y}}_{west}\right)$$
$$\left\{ 2\mu_{tX_h}^{-1}\left(\bar{x}_{tst} - \mu_{tX_h}\right) - 3\left(\mu_{tX_h}^{-1}\left(\bar{x}_{tst} - \mu_{tX_h}\right)\right)^2 - O_P\left(n_h^{-3/2}\right) \right\} \quad (47)$$

Since $\left(\frac{\mu_{tX_h}}{\bar{x}_{tst}}\right)^2 = 1 - \left\{ 2\mu_{tX_h}^{-1}\left(\bar{x}_{tst} - \mu_{tX_h}\right) - 3\left(\mu_{tX_h}^{-1}\right.\right.$ $\left.\left.\left(\bar{x}_{tst} - \mu_{tX_h}\right)\right)^2 - O_P\left(n_h^{-3/2}\right) \right\}$ see Kim et al. (2007). It is clear from Eq. (46) $V\left(\widehat{\bar{y}}'_{west}\right) < V\left(\widehat{\bar{y}}_{west}\right)$.

(ii) $V\left(\widehat{\bar{y}}''_{west}\right)_{\min} < V\left(\widehat{\bar{y}}_{west}\right)$ iff

$$\left\{ V\left(\widehat{\bar{y}}'_{west}\right) - V\left(\widehat{\bar{y}}'_{west}\right)_{\min} \right\} > 0 \quad (48)$$

from the Eqs. (38), (7) and (48), we have following expression

$$\left\{ V\left(\widehat{\bar{y}}_{west}\right) - \left(\widehat{\bar{y}}''_{west1}\right)_{\min} \right\} =$$
$$\sum_{h=1}^{L} \frac{W_h^2 \mu_{tY_h}^2}{n_h} \left\{ \frac{\left(\rho_{tXY_h} C_{tX_h} C_{tY_h} + \rho_{UV_h} C_{U_h} C_{V_h}\right)^2}{\left(C_{tX_h}^2 + C_{U_h}^2\right)} \right\} > 0 \quad (49)$$

which is always true.

Hence, the calibrated estimators $\widehat{\bar{y}}'_{west}, \widehat{\bar{y}}'_{west1}$ and $\widehat{\bar{y}}''_{west1}$ theoretically have shown supremacy over unbiased estimator $\widehat{\bar{y}}_{west}$ under measurement errors.

4. Simulation study

To examine the merit of the proposed calibrated estimators under measurement errors, we have conducted Monte-Carlo simulation study (see Shalabh and Tsai 2017,). The simulation study is

conducted through using R studio. We have drawn heterogeneous strata of the sizes $N_1 = 150, N_2 = 90$ and $N_3 = 120$ with the total population $N = 450$. The different sizes of strata are drawn from a 4-variate multivariate normal distribution with means $\left(\mu_{tX_h}, \ \mu_{tY_h}, 0, \ 0 \right)$ and covariance matrix

$$\begin{pmatrix} \sigma^2_{tX_h} & \rho_{tXY_h}\sigma_{tX_h}\sigma_{tY_h} & 0 & 0 \\ \rho_{tXY_h}\sigma_{tX_h}\sigma_{tY_h} & \sigma^2_{tY_h} & 0 & 0 \\ 0 & 0 & \sigma^2_{U_h} & \rho_{UV_h}\sigma_{U_h}\sigma_{V_h} \\ 0 & 0 & \rho_{UV_h}\sigma_{U_h}\sigma_{V_h} & \sigma^2_{V_h} \end{pmatrix}$$

where h = 1, 2, and 3 for the different strata with the details of the combination of the parameters. Which are given below as:

$\left(\mu_{tX_1}, \mu_{tX_2}, \mu_{tX_3} = 94, 88, 91 \right)$,

$\left(\mu_{tY_1}, \mu_{tY_2}, \mu_{tY_3} = 99, 92, 97 \right)$, $\left(\mu_{tY_1}, \right.$

$\left. \mu_{tY_2}, \mu_{tY_3} = 99, 92, 97 \right)$, $\left(\sigma^2_{tX_1}, \sigma^2_{tX_2}, \sigma^2_{tX_3} = 37.07, 35.76, 31.23 \right)$,

$\left(\sigma^2_{tY_1}, \sigma^2_{tY_2}, \sigma^2_{tY_3} = 33.59, 36.47, 34.90 \right)$,

$\left(\rho_{tXY_1}, \rho_{tXY_2}, \rho_{tXY_3} = 0.9, 0.5, 0.7 \right)$, for all possible combination we have

$\left\{ \sigma^2_{U_1} = \sigma^2_{U_2} = \sigma^2_{U_3} = (0, \ 1, \ 2, \ 3), \sigma^2_{V_1} = \sigma^2_{V_2} = \sigma^2_{V_3} = (0, \ 1, \ 2, \ 3) \right\}$

and the correlation coefficient for the correlated and uncorrelated measurement errors are as given following $\left\{ \rho_{UV_1} = \rho_{UV_2} = \rho_{UV_3} = (0.9, \ 0.7, \ 0.5, \ 0.0, -0.5, -0.7, -0.9) \right\}$. Similarly, for heterogeneous case are given by as $\left\{ \rho_{UV_1}, \ \rho_{UV_2}, \ \rho_{UV_3} = 0.5, \ 0.7, \ 0.9 \ \& -0.5, -0.7, -0.9 \right\}$.

Further, we have drawn samples $n_1 = 50$, $n_2 = 30$ and $n_3 = 40$ from each stratum with the total sample $n = 120$. The values of the means, estimated variances, relative efficiencies and percentage contribution of the measurement error are computed based on 5000 replications and are represented in the Tables 1–3 or by graphs in Figs. 1 and 2, see Appendix A. The relative efficiencies (RE) and percentage contribution of the measurement error (PCME) are defined respectively as

$$RE(*) = \frac{V\left(\widehat{\overline{y}}_{west} \right)}{V(*)}, \tag{50}$$

$$PCME = \frac{V(*)_m - V(*)_o}{V(*)_o} \times 100, \tag{51}$$

where $V(*)_0$ are the variances when there are no measurement errors, $V(*)_m$ are the variances when there are measurement errors and $* = \widehat{\overline{y}}_{west}, \widehat{\overline{y}}'_{west}$ and $\widehat{\overline{y}}''_{west}$.

Note: For the convenient of the graphical representation, the estimators $\widehat{\overline{y}}_{west}, \widehat{\overline{y}}'_{west}$ and $\widehat{\overline{y}}''_{west}$ are represented by y, y1 and y2 respectively in all Figs. 1 and 2.

5. Discussion

To analyze the contribution of measurement errors, we have obtained two proposed calibration estimators $\widehat{\overline{y}}'_{west}$ and $\widehat{\overline{y}}''_{west}$ with their respective variances to the first order approximation under measurement errors. The two calibrated estimators $\widehat{\overline{y}}'_{west}$ and $\widehat{\overline{y}}''_{west}$ are stated in theorem 1 and 3, and their respective variances are stated in theorem 3 and 4. Lemma 1 and lemma 4 state the

variances of the calibrated estimators $\widehat{\overline{y}}'_{west}$ and $\widehat{\overline{y}}''_{west}$ without measurement errors and in lemma 3, the variance of the calibrated estimator $\widehat{\overline{y}}''_{west}$ under uncorrelated measurement errors have shown. Lemma 2 and lemma 5 have shown that the variances of the calibrated estimators $\widehat{\overline{y}}'_{west}$ and $\widehat{\overline{y}}''_{west}$ have greater variances than to the respective same estimators under without measurement errors. Theoretically, lemma 2 and 5 have shown that there is contribution of measurement errors (correlated and uncorrelated). Also, in section 3, we have shown the supremacy of the calibrated estimators $\widehat{\overline{y}}'_{west}$ and $\widehat{\overline{y}}''_{west}$ over mean per unit estimator $\widehat{\overline{y}}_{west}$ under measurement errors through the efficiency comparisons.

Thus, theoretically we have shown the aim of the study but for more analytical study, we have conducted a simulation study in Section 4. The results are shown in Tables 1–3 as a numerical and as a graphical in Figs. 1 and 2. We can see in the Tables 1–3, the relative efficiency of the proposed calibrated estimators $\widehat{\overline{y}}'_{west}$ and $\widehat{\overline{y}}''_{west}$ are greater than mean per unit estimator $\widehat{\overline{y}}_{west}$ under measurement errors. Also, we have represented combination of data in Figs. 1 and 2 along with data level. The Figs. 1 and 2 represent the results with changing values of the correlation coefficient of the measurement errors. We can see from Figs. 1 and 2, the relative efficiency of the proposed calibrated estimators over mean per unit estimator have increased with the increasing values of the positive homogeneous correlation coefficient and have decreased with the increasing values of the negative homogeneous correlation coefficient of the measurement errors. Alike homogeneous correlation coefficient of the measurement errors, heterogeneous case has also shown the same pattern for the relative efficiency of the proposed calibrated estimators.

Theoretically, lemma 2 and 5 have shown that there is contribution of measurement errors (correlated and uncorrelated) and numerically, we have calculated as percentage contribution of measurement errors (PCME) through simulation study. We can see from the Tables 1 and 3 or from the Figs. 1 and 2, the PCME of the proposed calibrated estimators under measurement errors are increasing or decreasing with same pattern. In all the Tables 1–3 or in Figs. 1 and 2, we can see the PCME of the calibrated estimators are increasing with the increased variability of the measurement errors. In Tables 1, the proposed calibrated estimators have high values of PCME for all the cases of the uncorrelated measurement errors, that is the proposed estimators are highly affected by the measurement errors. Also, in heterogeneous correlated measurement errors case, the proposed estimators are highly affected by the measurement errors for all the variability cases of the measurement errors, But, for the positive correlated measurement errors cases not much affected. In Table 2, the proposed estimators are not much highly affected by measurement errors except in the case $\sigma^2_{U_h} = 2$, $\sigma^2_{V_h} = 2$ and $\sigma^2_{U_h} = 3$, $\sigma^2_{V_h} = 3$ for which the calibrated estimators have high values of PCME. While, in Table 3, the proposed estimators are highly affected by measurement errors for all possible correlation coefficient and variability of the measurement errors due to high values of the PCME.

In Figs. 1 and 2, the effect of measurement errors is represented with the changing values of the correlation coefficient of the measurement errors. We can see from Figs. 1 and 2, the PCME of the proposed estimators are decreasing with increased values of the positive correlation coefficient of the measurement errors and are increasing with increased values of the negative correlation coefficient of the measurement errors. Even, the PCME of the pro-

posed estimators are high in the case of negative correlation coefficient of the measurement errors. This shows the proposed calibrated estimators are highly affected by the measurement errors than to all cases when the negative correlation coefficient of the measurement errors is observed.

6. Conclusions and further study

In this manuscript, we have proposed calibrated estimators in stratified sampling under both the correlated and uncorrelated measurement errors. The proposed calibrated estimators have shown superiority over usual mean estimator under all the cases of measurement errors and also when there are no measurement errors by theoretically as well as numerically. To examine the effect of measurement errors, we have calculated percentage contribution of measurement error ($PCME$) and we found that the $PCME$ of the calibrated estimators are increasing with increased variability of the measurement errors presented in both the study and auxiliary variables. Also, the $PCME$ of the calibrated estimators decreases with the positive increment of correlation coefficient of the measurement errors concerning to when there is no correlation in the measurement errors of the both variables. While, the $PCME$ of the calibrated estimators increases with the negative increment of the correlation coefficient of the measurement errors concerning to when there is no correlation in both the variables of the measurement errors. The calibrated estimators are highly affected by the measurement errors when the correlation coefficient in the measurement errors of the both variables are negative. A proper care should be taken in those cases where proposed estimators are highly affected by the measurement errors and specially for negative correlation coefficient of the measurement errors.

A further study can be done by applying different estimators on a calibration estimator having a calibration weight subject to fix single constraint (mean or variance), as in Eqs. (9) and (10) (see Kim et al. (2007), Clement (2017)). Second, by applying a calibration weight subject to many constraints (see Koyuncu and Kadilar (2014), Clement (2018) and Salinas et al. (2019)).

7. Disclosure of any funding to the study

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

Appendix A

Table 1
Shows variances, *PCME* and relative efficiencies of the calibrated estimators over usual mean estimator for uncorrelated and heterogeneous correlations cases.

Estimators	$\rho_{uv1} = \rho_{uv2} = \rho_{uv3} = 0.0,\ \sigma^2_{U_h} = 2,\ \sigma^2_{V_h} = 2$			
	$\widehat{\bar{y}}_{west}$	$V(*)$	RE	$PCME$
$\bar{y}_{west}$	79.3	0.201	1.00	0
	79.2	0.111	1.82	0
$\bar{y}''_{west}$	79.2	0.110	1.83	0

	$\rho_{uv1} = \rho_{uv2} = \rho_{uv3} = 0.0$											
	$\sigma^2_{U_h} = 1,\ \sigma^2_{V_h} = 1$				$\sigma^2_{U_h} = 2,\ \sigma^2_{V_h} = 2$				$\sigma^2_{U_h} = 3,\ \sigma^2_{V_h} = 3$			
	$\widehat{\bar{y}}_{west}$	$V(*)$	RE	$PCME$	$\widehat{\bar{y}}_{west}$	$V(*)$	RE	$PCME$	$\widehat{\bar{y}}_{west}$	$V(*)$	RE	$PCME$
$\bar{y}_{west}$	79.3	0.209	1.00	4.125	79.4	0.217	1.00	7.826	79.4	0.224	1.00	11.46
$\bar{y}'_{west}$	79.4	0.123	1.71	10.90	79.5	0.135	1.61	21.60	79.6	0.146	1.53	32.29
$\bar{y}''_{west}$	79.4	0.122	1.72	10.75	79.5	0.133	1.63	21.39	79.6	0.145	1.54	32.02
	$\rho_{uv1} = 0.5, \rho_{uv2} = 0.7, \rho_{uv3} = 0.9$											
$\bar{y}_{west}$	79.3	0.209	1.00	4.183	79.4	0.216	1.00	7.502	79.3	0.236	1.00	17.30
$\bar{y}'_{west}$	79.3	0.116	1.81	4.558	79.3	0.120	1.81	8.173	79.2	0.131	1.80	18.19
$\bar{y}''_{west}$	79.3	0.115	1.82	4.514	79.3	0.119	1.82	8.100	79.2	0.130	1.81	18.38
	$\rho_{uv1} = -0.5, \rho_{uv2} = -0.7, \rho_{uv3} = -0.9,$											
$\bar{y}_{west}$	79.3	0.205	1.00	1.983	79.3	0.221	1.00	9.909	79.2	0.216	1.00	7.586
$\bar{y}'_{west}$	79.2	0.130	1.58	17.55	79.3	0.164	1.35	48.11	79.3	0.173	1.25	56.09
$\bar{y}''_{west}$	79.2	0.129	1.59	17.48	79.3	0.162	1.36	48.04	79.3	0.171	1.26	55.95

Table 2
Shows variances, PCME and relative efficiencies of the calibrated estimators over usual mean estimator for positive correlations.

Estimators	$\rho_{uv1}=\rho_{uv2}=\rho_{uv3}=0.5$,											
	$\sigma^2_{U_h}=1,\ \sigma^2_{V_h}=1$				$\sigma^2_{U_h}=2,\ \sigma^2_{V_h}=2$				$\sigma^2_{U_h}=3,\ \sigma^2_{V_h}=3$			
	$\widehat{\bar{y}}_{west}$	$V(*)$	RE	PCME	$\widehat{\bar{y}}_{west}$	$V(*)$	RE	PCME	$\widehat{\bar{y}}_{west}$	$V(*)$	RE	PCME
$\bar{y}_{west}$	79.3	0.205	1.00	2.174	79.3	0.211	1.00	4.872	79.4	0.216	1.00	7.757
$\bar{y}'_{west}$	79.3	0.116	1.76	5.312	79.3	0.122	1.73	10.35	79.4	0.131	1.65	18.37
$\bar{y}''_{west}$	79.3	0.116	1.78	5.216	79.3	0.121	1.74	10.20	79.4	0.130	1.67	18.15
$\rho_{uv1}=\rho_{uv2}=\rho_{uv3}=0.7$,												
$\bar{y}_{west}$	79.3	0.206	1.00	2.659	79.4	0.212	1.00	5.617	79.4	0.228	1.00	13.49
$\bar{y}'_{west}$	79.3	0.114	1.81	3.224	79.3	0.118	1.81	6.262	79.2	0.126	1.81	13.88
$\bar{y}''_{west}$	79.3	0.113	1.82	3.161	79.3	0.117	1.82	6.166	79.2	0.125	1.82	14.04
$\rho_{uv1}=\rho_{uv2}=\rho_{uv3}=0.9$,												
$\bar{y}_{west}$	79.3	0.208	1.00	3.317	79.3	0.216	1.00	7.351	79.5	0.227	1.00	13.24
$\bar{y}'_{west}$	79.3	0.112	1.86	1.075	79.2	0.117	1.85	5.408	79.4	0.121	1.88	9.213
$\bar{y}''_{west}$	79.3	0.111	1.87	1.055	79.2	0.116	1.86	5.465	79.4	0.120	1.90	9.177

Table 3
Shows variances, PCME and relative efficiencies of the calibrated estimators over usual mean estimator for negative correlations.

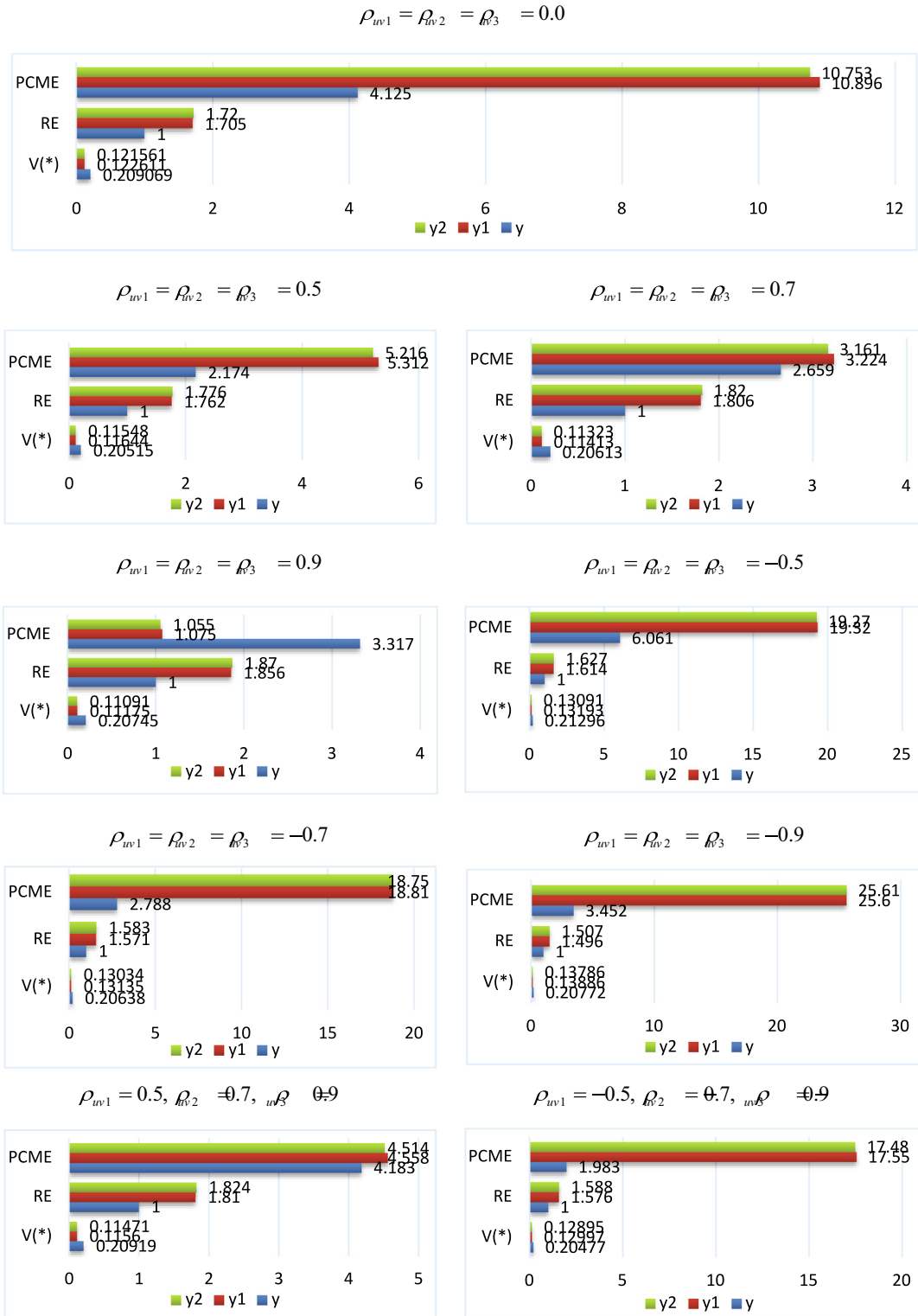| Estimators	$\rho_{uv1}=\rho_{uv2}=\rho_{uv3}=-0.5$,											
	$\sigma^2_{U_h}=1,\ \sigma^2_{V_h}=1$				$\sigma^2_{U_h}=2,\ \sigma^2_{V_h}=2$				$\sigma^2_{U_h}=3,\ \sigma^2_{V_h}=3$			
	$\widehat{\bar{y}}_{west}$	$V(*)$	RE	PCME	$\widehat{\bar{y}}_{west}$	$V(*)$	RE	PCME	$\widehat{\bar{y}}_{west}$	$V(*)$	RE	PCME
$\bar{y}_{west}$	79.3	0.213	1.00	6.061	79.3	0.224	1.00	11.61	79.3	0.220	1.00	9.366
$\bar{y}'_{west}$	79.4	0.132	1.61	19.32	79.3	0.159	1.41	43.66	79.3	0.166	1.32	50.47
$\bar{y}''_{west}$	79.4	0.131	1.63	19.27	79.3	0.158	1.42	43.62	79.3	0.165	1.33	50.35
$\rho_{uv1}=\rho_{uv2}=\rho_{uv3}=-0.7$,												
$\bar{y}_{west}$	79.3	0.206	1.00	2.788	79.2	0.215	1.00	7.265	79.3	0.219	1.00	8.984
$\bar{y}'_{west}$	79.2	0.131	1.57	18.81	79.1	0.160	1.35	44.52	79.3	0.174	1.26	56.88
$\bar{y}''_{west}$	79.2	0.130	1.58	18.75	79.1	0.159	1.36	44.50	79.3	0.172	1.27	56.77
$\rho_{uv1}=\rho_{uv2}=\rho_{uv3}=-0.9$,												
$\bar{y}_{west}$	79.2	0.208	1.00	3.452	79.3	0.211	1.00	5.273	79.3	0.222	1.00	10.60
$\bar{y}'_{west}$	79.1	0.139	1.50	25.60	79.2	0.158	1.34	43.07	79.4	0.198	1.12	78.95
$\bar{y}''_{west}$	79.1	0.138	1.51	25.61	79.2	0.157	1.35	43.00	79.4	0.196	1.13	78.86

$$\rho_{uv1} = \rho_{uv2} = \rho_{uv3} = 0.0$$

$$\rho_{uv1} = \rho_{uv2} = \rho_{uv3} = 0.5$$

$$\rho_{uv1} = \rho_{uv2} = \rho_{uv3} = 0.7$$

$$\rho_{uv1} = \rho_{uv2} = \rho_{uv3} = 0.9$$

$$\rho_{uv1} = \rho_{uv2} = \rho_{uv3} = -0.5$$

$$\rho_{uv1} = \rho_{uv2} = \rho_{uv3} = -0.7$$

$$\rho_{uv1} = \rho_{uv2} = \rho_{uv3} = -0.9$$

$$\rho_{uv1} = 0.5, \ \rho_{uv2} = 0.7, \ \rho_{uv3} = 0.9$$

$$\rho_{uv1} = -0.5, \ \rho_{uv2} = 0.7, \ \rho_{uv3} = 0.9$$

Fig. 1. Shows variances, *PCME* and relative efficiencies of the calibrated estimators over usual mean estimator when $\sigma^2_{U_h} = 1$, $\sigma^2_{V_h} = 1$.
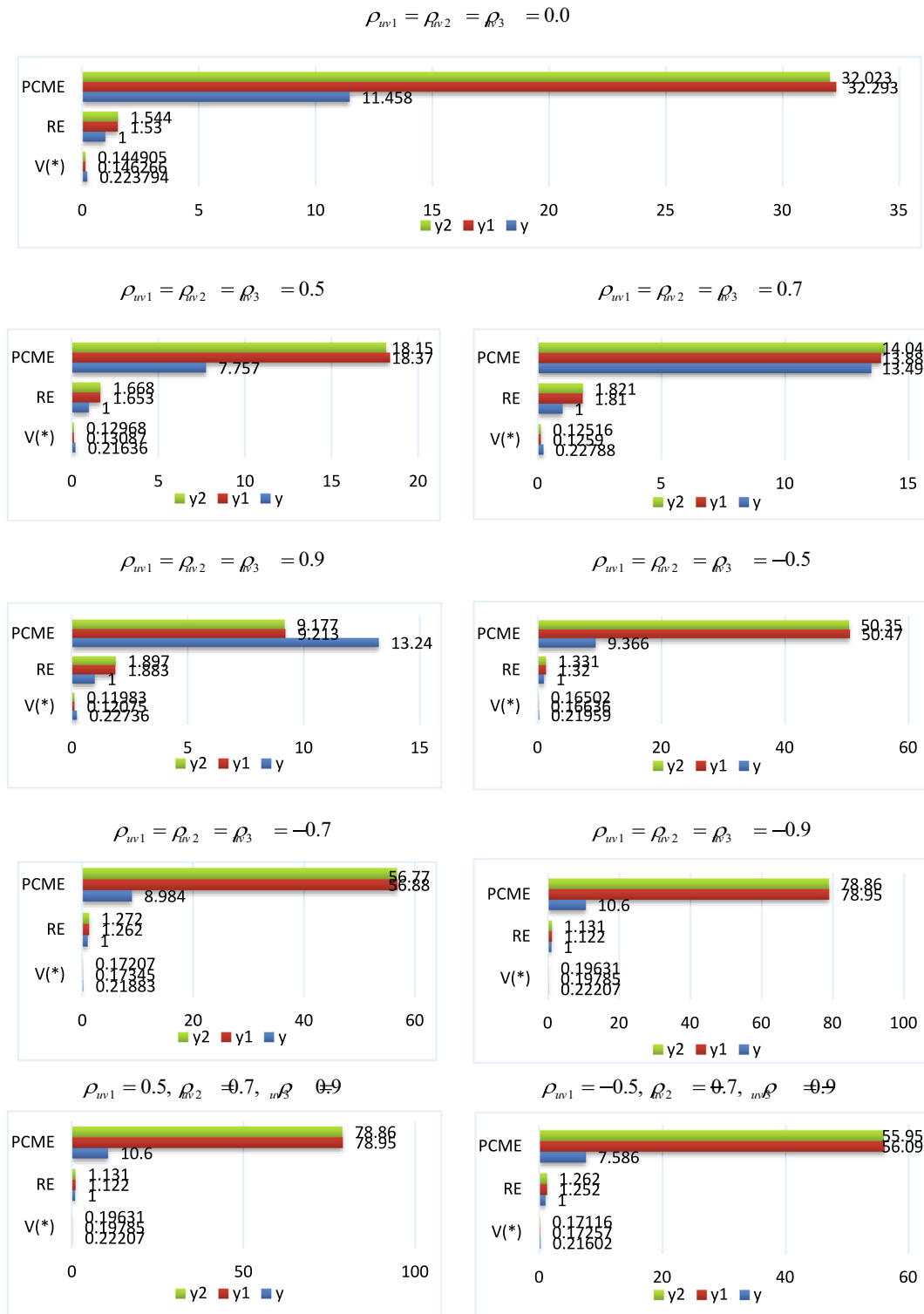
Fig. 2. Shows variances, *PCME* and relative efficiencies of the calibrated estimators over usual mean estimator when $\sigma^2_{U_h} = 3$, $\sigma^2_{V_h} = 3$.

References

Alka, Rai, P.K., Qasim, M., 2019. Two-step calibration of design weights under two auxiliary variables in sample survey. J. Stat. Comput. Simul.. https://doi.org/10.1080/00949655. 2019.1616205.

Biemer, P., and S.L. Stokes (1991). Approaches to the modelling of measurement error; Measurement error in Surveys. Hoboken, NJ: Wiley, 485–516.

Carroll, R.J., Ruppert, D., Stefanski, L.A., Crainiceanu, C.M., 2006. Measurement Error in Nonlinear Models: A Modern Perspective. CRC Press, Boca Raton, FL.

Clement, E.P., Enang, E.I., 2015. Calibration approach alternative ratio estimator for population mean in stratified sampling. Int. J. Statist. Econ. 16 (1), 83–93.

Clement, E., 2017. A new ratio estimator of mean in survey sampling by calibration estimation. Elixir Int. J. Statist. 106, 46461–46465.

Clement, E.P., 2018. A note on calibration weightings for stratified double sampling with equal probability. Commun. Statist. Theory Methods 47 (12), 2835–2847.

Cochran, W.G., 1968. Errors of measurement in statistics. Technometrics 10 (4), 637–666.

Deville, J.C., Sarndal, C.E., 1992. Calibration estimators in survey sampling. J. Am. Stat. Assoc. 87, 376–382.

Fuller, W.A., 1987. Measurement Error Models. John Wiley and Sons Inc, Hoboken, NJ.

Garg, N., Pachori, M., 2019. Use of coefficient of variation in calibration estimation of population mean in stratified sampling. Commun. Statist. Theory Methods. https://doi.org/10.1080/03610926.2019.1622729.

Gregoire, T.G., Salas, C., 2009. Ratio estimation with measurement error in the auxiliary variate. Biometrics 65 (2), 590–598.

Kim, J.M., Sungur, E.A., Heo, T.Y., 2007. Calibration approach estimators in stratified sampling. Statist. Probabil. Lett. 77, 99–103.

Koyuncu, N., Kadilar, C., 2014. A new calibration estimator in stratified double sampling. Hacettepe J. Mathemat. Statist. 43 (2), 1–9.

Kumar, M., Singh, R., Singh, A.K., Smarandache, F., 2011. Some ratio type estimators under measurement error. World Appl. Sci. J. 14 (2), 272–276.

Maneesha, Singh, R.K., 2002. Role of regression estimator involving measurement Errors. Braz. J. Probabil. Statist. 16, 39–46.

Manisha, Singh, R.K., 2001. An estimation of population mean in the presence of measurement error. J. Indian Soc. Agric. Statist. 54 (1), 13–18.

Murthy, M.N., 1967. Sampling Theory and Methods. Statistical Publishing Society, Calcutta.

Rao, D.K., Khan, M., Singh, G., 2017. On calibrated weights in stratified sampling. Austr. N. Z. Ind. Appl. Mathemat. J. 59, 190–204.

Salinas, V.I., Sedory, S.A., Singh, S., 2019. Calibrated estimators in two-stage sampling. Commun. Statist. Theory Methods 48 (6), 1449–1469.

Shalabh, 1997. Ratio method of estimation in the presence of measurement errors. J. Indian Soc. Agric. Statist. 50, 150–155.

Shalabh, Tsai, J.R., 2017. Ratio and product methods of estimation of population mean in the presence of correlated measurement error. Commun. Statist. Simul. Comput. 46 (7), 5566–5593.

Singh, S., Horn, H., Yu, F., 1998. Estimation of variance of the general regression estimator: Higher level calibration approach. Survey methodology 24 (1), 41–50.

Singh, D., Sisodia, B.V.S., Nidhi, Pundir, S., 2019. Some calibration estimators for finite population mean in two-stage stratified random sampling. Commun. Statist. Theory Methods. https://doi.org/10.1080/03610926.2019.1596283.

Singh, S. (2001). Generalized calibration approach for estimating variance in survey sampling. Annals of the Institute of Statistical Mathematics, 53, 404–417.

Singh, S, 2003. Advanced sampling theory with applications, II. Kluwer Academic Publishers.

Singh, S., 2012. On the calibration of design weights using a displacement function. Metrika 75 (1), 85–107.

Singh, S., 2013. A dual problem of calibration of design weights. Statistics 47 (3), 566–574.

Singh, S., Arnab, R., 2011. On the calibration of design weights. Metron. LXIX (2), 185–205.

Sinha, N., Sisodia, B.V.S., Singh, S., Singh, S.K., 2017. Calibration approach estimation of mean in stratified sampling and stratified double sampling. Commun. Statist. Theory Method 46 (10), 4932–4942.

Tailor, R., Chouhan, S., Kim, J.M., 2014. Ratio and product type exponential estimators of population mean in double sampling for stratification. Commun. Statist. Appl. Methods 21 (1), 1–9.

Tracy, D.S., Singh, S., Arnab, R., 2003. Note on calibration in stratified and double sampling. Surv. Methodol. 29, 99–104.

Vishwakarma, G.K., Singh, H.P., 2012. A general procedure for estimating the mean using double sampling for stratification and multi-auxiliary information. J. Statist. Plann. Inference 142 (5), 1252–1261.