Journal of Modern Applied Statistical Methods

Volume 23 | Issue 1 Article 2

# Smoothing of Estimators of Population mean using Calibration Technique with Sample Errors

Matthew Iseh
*Department of Statistics, Akwa Ibom State University, Nigeria,*
matthewiseh@aksu.edu.ng

Mbuotidem Bassey
*Department of Statistics, Akwa Ibom State University, Nigeria,*
mbuoti66@gmail.com

Recommended Citation

Matthew Iseh, Mbuotidem Bassey (2024). Smoothing of Estimators of Population mean using Calibration Technique with Sample Errors. Journal of Modern Applied Statistical Methods, 23(1), https://doi.org/10.56801/Jmasm.V23.i1.2

Journal of Modern Applied Statistical Methods
2024, Vol. 23, No. 1,
Doi: 10.56801/Jmasm.V23.i1.2

ISSN 1538 – 9472

# Smoothing of Estimators of Population mean using Calibration Technique with Sample Errors

**Matthew Iseh**
Department of Statistics, Akwa Ibom State University, Nigeria

**Mbuotidem Bassey**
Department of Statistics, Akwa Ibom State University, Nigeria

The challenges bedeviling the performance of estimators of population parameters in survey samples as a result of measurement and nonresponse errors are of great concern to researchers and users of statistics. This study suggests new estimators and adopts the calibration approach in smoothing the existing and proposed estimators for optimal performance. We have proposed improved estimators for estimating the finite population mean under stratified random sampling in three different situations: first, the properties of the estimators are considered under nonresponse, then the study of the estimators for measurement errors and in the last case, the estimators are examined in the presence of both measurement and nonresponse errors simultaneously. Expressions for mean square errors are obtained for the suggested estimators. Empirical study has been carried out with two real datasets to validate the theoretical underpinnings of this study.

*Keywords:* Auxiliary variable; calibration; nonresponse; optimization; mean square error.

## 1. Introduction

While conducting sample survey, statistician often come across non-sampling errors like measurement errors, coverage errors and nonresponse errors. The measurements that we get on the units for estimating the characteristics under study are seldom correct. And in practical situations the observations on this units are not correctly measured and differ from true values of the observations. This difference between the observation values and true values on the characteristics under study are called measurement errors or observational errors and is quite frequent in survey sampling. It is a kind of non- sampling errors and may arise due to the following reasons, that:

- the respondent may not provide the required information. However, the question was meant for the proper respondent. Example many families in Africa do not record a birth in the family and hence no birth certificate is made as the birth was not registered. hence, in this case it may be possible that the respondent included in the sampling may be given an approximate figure for the age which may not be the actual age, as the birth was not registered

- sometime it may happen that the observation may be on the closely related substitutes called proxies, although the variable is well defined. As an example ; if we are interested to know the economic status of a person and supposed the person is not willing to answer this question, then we may pool out the desired information by modifying the question, for instance instead of asking his economic status directly; we can ask about his educational level. However this will be only a guess as it is not necessary that a highly educated man/woman is economically well established and vice-versa
- it may also be due to respondent has misunderstood a particular question and hence supplied the information accordingly, (Tabssum 2021).

Several authors like Singh and Karpe (2009), Shalabh (1997), Manisha (2001) and Sud and Srivastava (2000) have discussed the problem of measurement errors.

One more error that arises frequently during survey sampling are the nonresponse errors. This errors are also part of non-sampling errors and arise due to the following reason; the absence of the respondent at the time of survey/refusal to answer the question or inability to recall the answer. Authors such as Hansen and Hurwitz (1946), Rao, P.S.R.S. (1986), Khare and Srivastava (1993), Khara and Srivastava (1997), Tabasum, R. and Khan, I. A. (2006), Singh and Kumar (2008), Singh and Kumar (2010), Kumar, Singh and Gupta (2011), Singh, Kumar and Kozak (2012), Iseh and Bassey (2021a, 2021b), Iseh and Bassey (2022) and have studied the problem of nonresponse to a large extent, however, the challenges are still enormous and call for further research to enhance efficiency.

## 2. Methodology/Existing Estimators

- The Hansen and Hurvitz (1946) estimator in stratified random sampling under measurement errors and nonresponse for estimating population mean is given by.

$$\bar{y}^*_{s(HH)} = \sum_{h=1}^{L} P_h \, \bar{y}^*_h$$

The expression for the variance of $\bar{y}^*_{s(HH)}$ is given by:

$$V\left(\bar{y}^*_{s\,(HH)}\right) = \sum_{h=1}^{L} P_h^2 A_h \tag{1}$$

where $\bar{y}^*_h = \left(\frac{n_{1h}}{n_h}\right)\bar{y}_{1h} + \left(\frac{n_{2h}}{n_h}\right)\bar{y}'_{2h}$, $and\ P_h = \frac{N_h}{N}$

$$A_h = \lambda_{2h}\left(S^2_{hY.} + S^2_{hU}\right) + \theta_h\left(S^2_{hY(2)} + S^2_{hU(2)}\right);\ \lambda_{2h} = \frac{1}{n_h} - \frac{1}{N_h};\ \theta_h = \frac{P_{2h}\left(g_h - 1\right)}{n_h}$$

where, $\bar{y}^*_h$ *and* $\bar{y}'_{2h}$ are the sampling means based on nonresponse and $K_h$ units of sub- sample from $n_{2h}$ nonresponding groups respectively.

- The separate ratio estimator for stratified random sampling under measurement error and nonresponse is given by:

$$\bar{y}^{*}_{s(R)} = \sum_{h=1}^{L} P_h \frac{\bar{y}^{*}_{h}}{\bar{x}^{*}_{h}} \bar{X}_h$$

The expressions for the Bias and MSE of $\bar{y}^{*}_{h(R)}$ are given by

$$Bias(\bar{y}^{*}_{s(R)}) \cong \sum_{h=1}^{L} \frac{P_h}{\bar{X}_h}(R_h B_h - C_h)$$

$$\text{MSE}\left(\bar{y}^{*}_{s(R)}\right) \cong \sum_{h=1}^{L} P_h^2 \left(A_h + R_h^2 B_h - 2R_h C_h\right) \quad (2)$$

where $R_h \frac{\bar{Y}}{\bar{X}}, B_h = \lambda_{2h}\left(S^2_{hX} + S^2_{hV}\right) + \theta_h\left(S^2_{hX(2)} + S^2_{hV(2)}\right)$ and

$C_h = \lambda_{2h}\, \rho_{hXY}\, S_{hY} S_{nX} + \theta_h \rho_{hXY(2)} S_{hY(2)}\, S_{nX(2)}$

- The separate difference estimator in stratified random sampling under measurement error and nonresponse is given by:

$$\bar{y}^{*}_{s(D)} = \sum_{h=1}^{L} P_h[\bar{y}^{*}_{h} + d_h(\bar{X}_h - \bar{x}^{*}_{h})]$$

where $\bar{x}^{*}_{h} = \frac{N_h \bar{X}_h - n_h \bar{x}^{*}_{h}}{N_h - n_h}$ and $d_h$ is a constant.

The expression for minimum variance of $\bar{y}^{*}_{s(D)}$ is given by:

$$V\left(\bar{y}^{*}_{s(D)}\right) min = \sum_{h=1}^{L} P_h^2 \left[A_h - \frac{C_h^2}{B_h}\right] \quad (3)$$

where the optimum value of $d_h$ is $d_{h(opt)} = -\frac{C_h}{t_h\, B_h}$ and $t_h = \frac{n_h}{N_h - n_h}$

- Azeem and Hanif (2016) estimator under stratified random sampling is given by

$$\bar{y}^{*}_{s(AH)} = \sum_{h=1}^{L} P_h \bar{y}^{*}_{h} \left(\frac{\bar{x}^{*}_{h}}{\bar{X}_h}\right) \left(\frac{\bar{x}^{*}_{h}}{\bar{x}_h}\right)$$

The expressions for both Bias and MSE of the estimator of $\bar{y}^{*}_{s(AH)}$ are as follows

$$Bias\left(\bar{y}^{*}_{s(AH)}\right) \cong \sum_{h=1}^{L} \frac{p_h}{\bar{X}_h}\left(t_h^2\, R_h B_h - q_h C_h\right)$$

$$MSE\left(\bar{y}^{*}_{s(AH)}\right) \cong \sum_{h=1}^{L} P_h \left(A_h + q_h^2\, R_n^2\, B_h - 2q_h\, R_h\, C_h\right) \quad (4)$$

where $q_h = \frac{N_h + n_h}{N_h - n_h}$

- Zahid and Shabbir (2018) proposed an estimator for population mean in stratified random sampling as:

$$\bar{y}^{*}_{s(p)} = \sum_{h=1}^{L} P_h \left[m_h\, \bar{y}^{*}_{h} + m_{2h}(\bar{X}_h - \bar{x}'^{*}_{h}) \left(\frac{\bar{x}_h}{\bar{x}'^{*}_{h}}\right) \exp\, (1 - \propto_h) \left(\frac{\bar{x}_h - \bar{x}'^{*}_{h}}{\bar{x}_h + \bar{x}'^{*}_{h}}\right)\right]$$

where $m_{1h}$ and $m_{2h}$ are constants whose values are to be determine and $\propto_h$ is the scalar chosen arbitrarily.

The expression for the bias and MSE of the estimator $\bar{y}^*_{s(p)}$ are

$Bias\left(\bar{y}^*_{s(P)}\right) \cong \sum_{h=1}^{L} P_h \left[(m_{1h}-1)\bar{Y}_h + m_{1h}\left(\frac{C_h t_h R_h C_h}{\bar{X}_h} + \frac{f_h R_h t_h^2 B_h}{\bar{X}_h}\right) + m_{2h}\left(\frac{C_h t_h R_h C_h}{\bar{X}_h}\right)\right]$

$$\text{MSE}\left(\bar{y}^*_{s(P)}\right) \cong \sum_{h=1}^{L} P_h^2 \left[\bar{Y}_h^2 - \frac{A_{h1} E_{h1}^2 + B_{h1} D_{h1}^2 - 2C_{h1} D_{h1} E_{h1}}{A_{h1} B_{h1} - C_{h1}^2}\right] \quad (5)$$

where $A_{h1} = \bar{Y}_h^2 + A_h + e_h^2 t_h^2 R_h^2 B_h + 4e_h t_h R_h C_h + 2f_h t_h^2 R_h^2 B_h$, $B_{h1} = t_h^2 B_h$

$C_{h1} = t_h C_h + 2e_h t_h^2 R_h B_h$,

$D_{h1} = \bar{Y}_h^2 + e_h t_h R_h C_h + f_h t_h^2 R_h^2 B_h$,

$E_{h1} = e_h t_h^2 R_h B_h$ ,

$e_h = \frac{1+\alpha_h}{2}$, $f_h = \alpha_h^2 + 4\alpha_h + 3$

- ❖ Rajest et al (2020) proposed an estimator for population mean in stratified random sampling when nonresponse is observed for both study and auxiliary variables as:

$t^* = \sum_{h=1}^{L} P_h \left[\bar{y}_h^* + \alpha \log\left(\frac{\bar{x}_h^*}{\bar{X}_h}\right)\right]$

with bias and mse given as

$\text{Bias}(t^*) = -\frac{1}{2}\sum_{h=1}^{L} \frac{P_h}{\bar{X}_h^2} \alpha_h B_{hQ}$

$$\text{MSE}(t^*) = \sum_{h=1}^{L} P_h^2 \left(A_{hQ} + \frac{\alpha_h^2}{\bar{X}_h^2} B_{hQ} + 2\frac{\alpha_h}{\bar{X}_h} C_{hQ}\right) \quad (6)$$

and

$$\text{MSE}(t^*)_{min} = \sum_{h=1}^{L} P_h^2 \left(A_{hQ} - \frac{C_{hQ}^2}{B_{hQ}}\right) \quad (7)$$

## 2.1 Proposed Estimators

By adopting the Rajest et al (2020) estimator, we proposed the following estimator in the presence of nonresponse

$$\bar{y}^*_{pr} = \sum_{h=1}^{L} P_h \left[\bar{y}_h^* + (1-\alpha)\log\left(\frac{\bar{x}_h^*}{\bar{X}_h}\right)\right] \quad (8)$$

In order to obtain the expression for bias and MSE of the proposed estimator, we assume that:

$\eta^*_{hY} = \sum_{i=1}^{n_h} (Y^*_{hi} - \bar{Y}_h)$

$\frac{\eta^*_{hY}}{n_h} = \frac{\sum_{i=1}^{n_h} Y^*_{hi} - n_h \bar{Y}_h}{n_h}$

$$\bar{y}_h^* = \bar{Y}_h + \frac{\eta^*_{hY}}{n_h} \quad (9)$$

Similarly,

$\eta_{hX}^{*} = \sum_{i=1}^{n_h}(X_{hi}^{*} - \bar{X}_h)$

$\frac{\eta_{hX}^{*}}{n_h} = \frac{\sum_{i=1}^{n_h} X_{hi}^{*} - n_h\bar{X}_h}{n_h}$

$$\bar{x}_h^{*} = \bar{X}_h + \frac{\eta_{hX}^{*}}{n_h} \quad (10)$$

Substuting Eqs. 9 and 10 in Eq. 8, gives

$$\bar{y}_{pr}^{*} = \sum_{h=1}^{L} P_h \left[\left(\bar{Y}_h + \frac{\eta_{hY}^{*}}{n_h}\right) + (1-\alpha)\log\left(\frac{\bar{X}_h + \frac{\eta_{hX}^{*}}{n_h}}{\bar{X}_h}\right)\right]$$

$$= \sum_{h=1}^{L} P_h \left[\left(\bar{Y}_h + \frac{\eta_{hY}^{*}}{n_h}\right) + \log\left(\frac{\bar{X}_h + \frac{\eta_{hX}^{*}}{n_h}}{\bar{X}_h}\right) - \alpha\log\left(\frac{\bar{X}_h + \frac{\eta_{hX}^{*}}{n_h}}{\bar{X}_h}\right)\right] \quad (11)$$

recall that

$\log(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \cdots for \ \mid x \mid < 1$ we have for

$\mid \frac{\eta_{hX}^{*}}{\bar{X}_h n_h} \mid < 1$

$$\log\left(1 + \frac{\eta_{hX}^{*}}{\bar{X}_h n_h}\right) = \frac{\eta_{hX}^{*}}{\bar{X}_h n_h} - \frac{1}{2}\left(\frac{\eta_{hX}^{*}}{\bar{X}_h n_h}\right)^2 + \frac{1}{3}\left(\frac{\eta_{hX}^{*}}{\bar{X}_h n_h}\right)^3 - \cdots$$

Since $\frac{\bar{X}_h + \frac{\eta_{hX}^{*}}{n_h}}{\bar{X}_h} \Rightarrow \frac{\bar{X}_h n_h + \eta_{hX}^{*}}{\bar{X}_h n_h} = 1 + \frac{\eta_{hX}^{*}}{\bar{X}_h n_h}$

Therefore, Eq. 11 becomes

$$= \sum_{h=1}^{L} P_h \left[\left(\bar{Y}_h + \frac{\eta_{hY}^{*}}{n_h}\right) + \frac{\eta_{hX}^{*}}{\bar{X}_h n_h} - \frac{1}{2}\left(\frac{\eta_{hX}^{*}}{\bar{X}_h n_h}\right)^2 + \frac{1}{3}\left(\frac{\eta_{hX}^{*}}{\bar{X}_h n_h}\right)^3 - \alpha\left\{\frac{\eta_{hX}^{*}}{\bar{X}_h n_h} - \frac{1}{2}\left(\frac{\eta_{hX}^{*}}{\bar{X}_h n_h}\right)^2 + \frac{1}{3}\left(\frac{\eta_{hX}^{*}}{\bar{X}_h n_h}\right)^3\right\}\right]$$

By ignoring powers of $\frac{\eta_{hX}^{*}}{\bar{X}_h n_h}$ greater than 2, we have

$$= \sum_{h=1}^{L} P_h \left[\bar{Y}_h + \frac{\eta_{hY}^{*}}{n_h} + \frac{\eta_{hX}^{*}}{\bar{X}_h n_h} - \frac{1}{2}\left(\frac{\eta_{hX}^{*}}{\bar{X}_h n_h}\right)^2 - \alpha\left\{\frac{\eta_{hX}^{*}}{\bar{X}_h n_h} - \frac{1}{2}\left(\frac{\eta_{hX}^{*}}{\bar{X}_h n_h}\right)^2\right\}\right]$$

$$\bar{y}_{pr}^{*} - \bar{Y}_h = \sum_{h=1}^{L} P_h \left[\bar{Y}_h + \frac{\eta_{hY}^{*}}{n_h} + \frac{\eta_{hX}^{*}}{\bar{X}_h n_h} - \frac{1}{2}\left(\frac{\eta_{hX}^{*}}{\bar{X}_h n_h}\right)^2 - \alpha\left\{\frac{\eta_{hX}^{*}}{\bar{X}_h n_h} - \frac{1}{2}\left(\frac{\eta_{hX}^{*}}{\bar{X}_h n_h}\right)^2\right\} - \bar{Y}_h\right]$$

$$E\left(\bar{y}_{pr}^{*} - \bar{Y}_h\right) = \sum_{h=1}^{L} P_h \left[E\left(\frac{\eta_{hY}^{*}}{n_h}\right) + E\left(\frac{\eta_{hX}^{*}}{\bar{X}_h n_h}\right) - \frac{1}{2}E\left(\frac{\eta_{hX}^{*}}{\bar{X}_h n_h}\right)^2 - \alpha\left\{E\left(\frac{\eta_{hX}^{*}}{\bar{X}_h n_h}\right) - \frac{1}{2}E\left(\frac{\eta_{hX}^{*}}{\bar{X}_h n_h}\right)^2\right\}\right] \quad (12)$$

Since,

$$E\left(\frac{\eta_{hX}^{*}}{\bar{X}_h n_h}\right) = E\left(\frac{\eta_{hY}^{*}}{\bar{Y}_h n_h}\right) = 0$$

$$E\left(\frac{\eta_{hY}^{*}}{n_h}\right)^2 = \gamma_{h2} S_{hy}^2 + \theta_h S_{hy(2)}^2 = A_{hQ}$$

$$E\left(\frac{\eta_{hX}^{*}}{n_h}\right)^2 = \gamma_{h2} S_{hx}^2 + \theta_h S_{hx(2)}^2 = B_{hQ}$$

$$E\left(\frac{\eta_{hX}^{*}}{n_h}\right) E\left(\frac{\eta_{hY}^{*}}{n_h}\right) = \gamma_h \rho_{hxy} S_{hx} S_{hy} + Q_h \rho_{hxy(2)} S_{hx(2)} S_{hy(2)} = C_{hQ}$$

Thus Eq. 12 becomes

$$B\left(\bar{y}_{pr}^{*}\right) = \sum_{h=1}^{L} P_h \left[-\frac{1}{2}E\left(\frac{\eta_{hX}^{*}}{\bar{X}_h n_h}\right)^2 + \alpha\left\{\frac{1}{2}E\left(\frac{\eta_{hX}^{*}}{\bar{X}_h n_h}\right)^2\right\}\right]$$

$$= -\frac{1}{2}\sum_{h=1}^{L} \frac{P_h}{\bar{X}_h^2}\left[\left\{\gamma_{h2} S_{hx}^2 + \theta_h S_{hx(2)}^2\right\} - \alpha\left\{\gamma_{h2} S_{hx}^2 + \theta_h S_{hx(2)}^2\right\}\right]$$

$$= -\frac{1}{2}\sum_{h=1}^{L} \frac{P_h}{\bar{X}_h^2}\left[\gamma_{h2} S_{hx}^2 (1-\alpha) + \theta_h S_{hx(2)}^2 (1-\alpha)\right]$$

$$= -\frac{1}{2}\sum_{h=1}^{L} \frac{P_h}{\bar{X}_h^2}\left[(1-\alpha)\gamma_{h2} S_{hx}^2 + \theta_h S_{hx(2)}^2\right]$$

$$= -\frac{1}{2}\sum_{h=1}^{L} \frac{P_h}{\bar{X}_h^2}\left[(1-\alpha) B_{hQ}\right] \quad (13)$$

The expression for MSE is derived as follows

$$\left(\bar{y}_{pr}^{*} - \bar{Y}_h\right)^2 = \sum_{h=1}^{L} P_h^2 \left[\frac{\eta_{hY}^{*}}{n_h} + \frac{\eta_{hX}^{*}}{\bar{X}_h n_h} - \frac{1}{2}\left(\frac{\eta_{hX}^{*}}{\bar{X}_h n_h}\right)^2 - \alpha\left\{\frac{\eta_{hX}^{*}}{\bar{X}_h n_h} - \frac{1}{2}\left(\frac{\eta_{hX}^{*}}{\bar{X}_h n_h}\right)^2\right\}\right]^2$$

$$= \sum_{h=1}^{L} P_h^2 \left\{\left[\frac{\eta_{hY}^{*}}{n_h} + \frac{\eta_{hX}^{*}}{\bar{X}_h n_h} - \frac{1}{2}\left(\frac{\eta_{hX}^{*}}{\bar{X}_h n_h}\right)^2 - \alpha\left\{\frac{\eta_{hX}^{*}}{\bar{X}_h n_h} - \frac{1}{2}\left(\frac{\eta_{hX}^{*}}{\bar{X}_h n_h}\right)^2\right\}\right]\left[\frac{\eta_{hY}^{*}}{n_h} + \frac{\eta_{hX}^{*}}{\bar{X}_h n_h} - \frac{1}{2}\left(\frac{\eta_{hX}^{*}}{\bar{X}_h n_h}\right)^2 - \alpha\left\{\frac{\eta_{hX}^{*}}{\bar{X}_h n_h} - \frac{1}{2}\left(\frac{\eta_{hX}^{*}}{\bar{X}_h n_h}\right)^2\right\}\right]\right\}$$

$$= \sum_{h=1}^{L} P_h^2 \left\{\left(\frac{\eta_{hY}^{*}}{n_h}\right)^2 + \left(\frac{\eta_{hY}^{*}}{n_h}\right)\left(\frac{\eta_{hX}^{*}}{\bar{X}_h n_h}\right) - \alpha\left(\frac{\eta_{hX}^{*}}{\bar{X}_h n_h}\right)\left(\frac{\eta_{hY}^{*}}{n_h}\right) + \left(\frac{\eta_{hY}^{*}}{n_h}\right)\left(\frac{\eta_{hX}^{*}}{\bar{X}_h n_h}\right) + \left(\frac{\eta_{hX}^{*}}{\bar{X}_h n_h}\right)^2 - \alpha\left(\frac{\eta_{hX}^{*}}{\bar{X}_h n_h}\right)^2 - \alpha\left(\frac{\eta_{hY}^{*}}{n_h}\right)\left(\frac{\eta_{hX}^{*}}{\bar{X}_h n_h}\right) - \alpha\left(\frac{\eta_{hX}^{*}}{\bar{X}_h n_h}\right)^2 + \alpha^2\left(\frac{\eta_{hX}^{*}}{\bar{X}_h n_h}\right)^2\right\}$$

$$E\left(\bar{y}_{pr}^{*}-\bar{Y}_{h}\right)^{2}=\sum_{h=1}^{L}P_{h}^{2}\left\{E\left(\frac{\eta_{hY}^{*}}{n_{h}}\right)^{2}+2E\left(\frac{\eta_{hY}^{*}}{n_{h}}\right)\left(\frac{\eta_{hX}^{*}}{\bar{X}_{h}n_{h}}\right)-2\alpha E\left(\frac{\eta_{hX}^{*}}{\bar{X}_{h}n_{h}}\right)\left(\frac{\eta_{hY}^{*}}{n_{h}}\right)+E\left(\frac{\eta_{hX}^{*}}{\bar{X}_{h}n_{h}}\right)^{2}-2\alpha E\left(\frac{\eta_{hX}^{*}}{\bar{X}_{h}n_{h}}\right)^{2}+\alpha^{2}E\left(\frac{\eta_{hX}^{*}}{\bar{X}_{h}n_{h}}\right)^{2}\right\} \quad (14)$$

Substituting for the individual expectations in Eq.14 gives

$$=\sum_{h=1}^{L}P_{h}^{2}\left\{\gamma_{h2}S_{hy}^{2}+\theta_{h}S_{hy(2)}^{2}+\frac{2}{\bar{X}_{i}}\left(\gamma_{h}\rho_{hxy}S_{hx}S_{hy}+Q_{h}\rho_{hxy(2)}S_{hx(2)}S_{hy(2)}\right)-\frac{2}{\bar{X}_{h}}\alpha\left(\gamma_{h}\rho_{hxy}S_{hx}S_{hy}+Q_{h}\rho_{hxy(2)}S_{hx(2)}S_{hy(2)}\right)+\frac{1}{\bar{X}_{h}{}^{2}}\left(\gamma_{h2}S_{hx}^{2}+\theta_{h}S_{hx(2)}^{2}\right)-\frac{2}{\bar{X}_{h}{}^{2}}\alpha\left(\gamma_{h2}S_{hx}^{2}+\theta_{h}S_{hx(2)}^{2}\right)+\alpha^{2}\frac{1}{\bar{X}_{h}{}^{2}}\left(\gamma_{h2}S_{hx}^{2}+\theta_{h}S_{hx(2)}^{2}\right)\right\}$$

$$=\sum_{h=1}^{L}P_{h}^{2}\left\{A_{hQ}+\frac{1}{\bar{X}_{h}{}^{2}}B_{hQ}(1-2\alpha)+\frac{1}{\bar{X}_{h}{}^{2}}\alpha^{2}B_{hQ}+\frac{2}{\bar{X}_{h}}C_{hQ}(1-\alpha)\right\}$$

$$MSE\left(\bar{y}_{pr}^{*}\right)=\sum_{h=1}^{L}P_{h}^{2}\left\{A_{hQ}+\frac{1}{\bar{X}_{h}{}^{2}}B_{hQ}(\alpha^{2}-2\alpha+1)+\frac{2}{\bar{X}_{h}}C_{hQ}(1-\alpha)\right\} \quad (15)$$

Minimizing $MSE\left(\bar{y}_{pr}^{*}\right)$ with respect to $\alpha$ and solving gives

$$\alpha=\left(1+\frac{\bar{X}_{h}C_{hQ}}{B_{hQ}}\right)$$

Thus, Eq. 15 could be written in terms of minimum MSE as

$$MSE(\bar{y}_{pr}^{*})_{min}=\sum_{h=1}^{L}P_{h}^{2}\left\{A_{hQ}+\frac{1}{\bar{X}_{h}{}^{2}}B_{hQ}\left[\left(1+\frac{\bar{X}_{h}C_{hQ}}{B_{hQ}}\right)^{2}-2\left(1+\frac{\bar{X}_{h}C_{hQ}}{B_{hQ}}\right)+1\right]+\frac{2}{\bar{X}_{h}}C_{hQ}\left[1-\left(1+\frac{\bar{X}_{h}C_{hQ}}{B_{hQ}}\right)\right]\right\} \quad (16)$$

**2.2 Estimation Using Calibration Technique**

Emphasis of this study is on using calibration method to smoothen the performance of both the existing and suggested estimators. The procedure is as follows follows;

Calibration of the Proposed Estimator

Given the proposed estimator in Eq. 8

$$\bar{y}_{pr}^{*}=\sum_{h=1}^{L}P_{h}\,\bar{y}_{h}^{**}$$

where $\bar{y}_{h}^{**}=\bar{y}_{h}^{*}+(1-\alpha)\log\left(\frac{\bar{x}_{h}^{*}}{\bar{X}^{*}}\right)$

Then the calibration estimator given as

$$\bar{y}_{prc}^{*}=\sum_{h=1}^{L}w_{h}\,\bar{y}_{h}^{**} \quad (17)$$

is minimized using the distance measure

$$\frac{\sum_{h=1}^{L}(w_{h}-P_{h})^{2}}{Q_{h}P_{h}}$$

subject to the constraint

$$\sum_{h=1}^{L} w_h \bar{x}_h^* = \bar{X} \tag{18}$$

Now, the optimization problem given as

$$L = \frac{\sum_{h=1}^{L}(w_h - P_h)^2}{Q_h P_h} - 2\lambda\left(\sum_{h=1}^{L} w_h \bar{x}_h^* - \bar{X}\right) \tag{19}$$

is solved to minimize the distance between the existing and calibration weights as;

$$\frac{\delta L}{\delta w_h} = \frac{2(w_h - P_h)}{Q_h P_h} - 2\lambda \bar{x}_h^* = 0$$

$$w_h = P_h[1 + \lambda \bar{x}_h^* Q_h]$$

Substituting for $w_h$ in Eq. 18 and solving for $\lambda$ gives

$$\lambda = \frac{\bar{X} - \sum_{h=1}^{L} P_h \bar{x}_h^*}{\sum_{h=1}^{L} P_h Q_h \bar{x}_h^{*2}}$$

Thus, the calibration weights become

$$w_h = P_h + \left[\frac{\bar{X} - \sum_{h=1}^{L} P_h \bar{x}_h^*}{\sum_{h=1}^{L} P_h Q_h \bar{x}_h^{*2}}\right] \bar{x}_h^* Q_h P_h$$

Eq. 17 then becomes

$$= \sum_{h=1}^{L} P_h \bar{y}_h^{**} + \sum_{h=1}^{L} P_h Q_h \bar{x}_h^* \bar{y}_h^{**} \left[\frac{\bar{X} - \sum_{h=1}^{L} P_h \bar{x}_h^*}{\sum_{h=1}^{L} P_h Q_h \bar{x}_h^{*2}}\right] \tag{20}$$

Substituting for $\bar{y}_h^{**}$ yields

$$= \sum_{h=1}^{L} P_h \bar{y}_h^* + \sum_{h=1}^{L} P_h (1-\alpha) \log\left(\frac{\bar{x}_h^*}{\bar{X}^*}\right) + \sum_{h=1}^{L} P_h Q_h \bar{x}_h^* \left[\bar{y}_h^* + (1-\alpha) \log\left(\frac{\bar{x}_h^*}{\bar{X}^*}\right)\right] \left[\frac{\bar{X} - \sum_{h=1}^{L} P_h \bar{x}_h^*}{\sum_{h=1}^{L} P_h Q_h \bar{x}_h^{*2}}\right] \tag{21}$$

Assuming $Q_h = 1$, we have

$$= \sum_{h=1}^{L} P_h \bar{y}_h^* + \sum_{h=1}^{L} P_h (1-\alpha) \log\left(\frac{\bar{x}_h^*}{\bar{X}^*}\right) + \left[\sum_{h=1}^{L} P_h \bar{x}_h^* \bar{y}_h^* + (1-\alpha) \sum_{h=1}^{L} P_h \bar{x}_h^* \log\left(\frac{\bar{x}_h^*}{\bar{X}^*}\right)\right] \left[\frac{\bar{X} - \sum_{h=1}^{L} P_h \bar{x}_h^*}{\sum_{h=1}^{L} P_h \bar{x}_h^{*2}}\right]$$

$$= \sum_{h=1}^{L} P_h \bar{y}_h^* + (1-\alpha) \sum_{h=1}^{L} P_h \log\left(\frac{\bar{x}_h^*}{\bar{X}^*}\right) + \beta\left(\bar{X} - \sum_{h=1}^{L} P_h \bar{x}_h^*\right) + (1-\alpha) \sum_{h=1}^{L} P_h \bar{x}_h^* \log\left(\frac{\bar{x}_h^*}{\bar{X}^*}\right) \left[\frac{\bar{X} - \sum_{h=1}^{L} P_h \bar{x}_h^*}{\sum_{h=1}^{L} P_h \bar{x}_h^{*2}}\right] \tag{22}$$

Where $\beta = \frac{\sum_{h=1}^{L} P_h \bar{x}_h^* \bar{y}_h^*}{\sum_{h=1}^{L} P_h \bar{x}_h^{*2}}$

Let $\bar{y}_h^* = \bar{Y}\left(1 + e_y\right)$, $\bar{x}_h^* = \bar{X}\left(1 + e_x\right)$, Such that Eq. 22 can be written as

$$\bar{y}_{prc}^{*} = \sum_{h=1}^{L} P_h \bar{Y}\left(1+e_y\right) + (1-\alpha)\sum_{h=1}^{L} P_h \log\left(\frac{\bar{X}\,(1+e_x)}{\bar{X}^*}\right) + \beta\left(\bar{X} - \sum_{h=1}^{L} P_h \bar{X}\,(1+e_x)\right) + (1-\alpha)\sum_{h=1}^{L} P_h \bar{X}\,(1+e_x)\log\left(\frac{\bar{X}\,(1+e_x)}{\bar{X}^*}\right)\left[\frac{\bar{X}-\sum_{h=1}^{L} P_h \bar{X}\,(1+e_x)}{\sum_{h=1}^{L} P_h \bar{X}^2(1+e_x)^2}\right]$$

Assume that $\bar{X}^* = \bar{X}$,

$$= \sum_{h=1}^{L} P_h \bar{Y} + \sum_{h=1}^{L} P_h \bar{Y} e_y + (1-\alpha)\sum_{h=1}^{L} P_h \log(1+e_x) + \beta\left(\bar{X} - \sum_{h=1}^{L} P_h \bar{X} - \sum_{h=1}^{L} P_h \bar{X} e_x\right) + (1-\alpha)\left[\sum_{h=1}^{L} P_h \bar{X} + \sum_{h=1}^{L} P_h \bar{X} e_x\right]\log(1+e_x)\left[\frac{\bar{X}-\sum_{h=1}^{L} P_h \bar{X} - \sum_{h=1}^{L} P_h \bar{X} e_x}{\sum_{h=1}^{L} P_h \bar{X}^2(1+e_x)^2}\right]$$

$$= \sum_{h=1}^{L} P_h \bar{Y} + \sum_{h=1}^{L} P_h \bar{Y} e_y + (1-\alpha)\sum_{h=1}^{L} P_h \log(1+e_x) + \beta\left(\bar{X} - \sum_{h=1}^{L} P_h \bar{X} - \sum_{h=1}^{L} P_h \bar{X} e_x\right) + (1-\alpha)\left[\sum_{h=1}^{L} P_h \bar{X} + \sum_{h=1}^{L} P_h \bar{X} e_x\right]\log(1+e_x)\left[\frac{\bar{X}-\sum_{h=1}^{L} P_h \bar{X} - \sum_{h=1}^{L} P_h \bar{X} e_x}{\sum_{h=1}^{L} P_h \bar{X}^2(1+e_x)^2}\right]$$

But $\log(1+e_x) = e_x - \frac{e_x^2}{2} + \frac{e_x^3}{3} - \cdots$

$$= \sum_{h=1}^{L} P_h \bar{Y} + \sum_{h=1}^{L} P_h \bar{Y} e_y + (1-\alpha)\sum_{h=1}^{L} P_h \left(e_x - \frac{e_x^2}{2} + \frac{e_x^3}{3} - \cdots\right) + \beta\left(\bar{X} - \sum_{h=1}^{L} P_h \bar{X} - \sum_{h=1}^{L} P_h \bar{X} e_x\right) + (1-\alpha)\left[\sum_{h=1}^{L} P_h \bar{X} + \sum_{h=1}^{L} P_h \bar{X} e_x\right]\left(e_x - \frac{e_x^2}{2} + \frac{e_x^3}{3} - \cdots\right)\left[\frac{\bar{X}-\sum_{h=1}^{L} P_h \bar{X} - \sum_{h=1}^{L} P_h \bar{X} e_x}{\sum_{h=1}^{L} P_h \bar{X}^2(1+e_x)^2}\right]$$

$$\left(\bar{y}_{prc}^{*} - \bar{Y}\right) = \left\{\sum_{h=1}^{L} P_h\bar{Y} + \sum_{\mathrm{h=1}}^{\mathrm{L}} P_h\bar{Y}e_y + (1-\alpha)\sum_{h=1}^{L} P_h\left(e_x - \frac{e_x^2}{2} + \frac{e_x^3}{3} - \cdots\right)\right.$$
$$+ \beta\left(\bar{X} - \sum_{h=1}^{L} P_h\bar{X} - \sum_{\mathrm{h=1}}^{\mathrm{L}} P_h\bar{X}e_x\right)$$
$$+ (1-\alpha)\left[\sum_{h=1}^{L} P_h\bar{X} + \sum_{\mathrm{h=1}}^{\mathrm{L}} P_h\bar{X}e_x\right]\left(e_x - \frac{e_x^2}{2} + \frac{e_x^3}{3}\right.$$
$$\left.\left. - \cdots\right)\left[\frac{\bar{X} - \sum_{h=1}^{L} P_h\bar{X} - \sum_{\mathrm{h=1}}^{\mathrm{L}} P_h\bar{X}e_x}{\sum_{\boldsymbol{h=1}}^{\boldsymbol{L}} P_h\bar{X}^2(1+e_x)^2}\right] - \bar{Y}\right\}$$

$$= \left\{\sum_{\mathrm{h=1}}^{\mathrm{L}} P_h\bar{Y}e_y + (1-\alpha)\sum_{h=1}^{L} P_h\left(e_x - \frac{e_x^2}{2} + \frac{e_x^3}{3} - \cdots\right) - \beta\left(\sum_{\mathrm{h=1}}^{\mathrm{L}} P_h\bar{X}e_x\right) + \right.$$
$$\left.(1-\alpha)\left[\sum_{h=1}^{L} P_h\bar{X} + \sum_{\mathrm{h=1}}^{\mathrm{L}} P_h\bar{X}e_x\right]\left(e_x - \frac{e_x^2}{2} + \frac{e_x^3}{3} - \cdots\right)\left[\frac{-\sum_{\mathrm{h=1}}^{\mathrm{L}} P_h\bar{X}e_x}{\sum_{\boldsymbol{h=1}}^{\boldsymbol{L}} P_h\bar{X}^2(1+e_x)^2}\right]\right\}$$

$$\mathrm{Bias}\left(\bar{y}_{prc}^{*}\right) = \left\{\sum_{\mathrm{h=1}}^{\mathrm{L}} P_h\bar{Y}E(e_y) + (1-\alpha)\sum_{h=1}^{L} P_h\left(E(e_x) - \frac{1}{2}E(e_x^2) + \frac{1}{3}E(e_x^3) - \right.\right.$$
$$\left.\cdots\right) - \beta\sum_{\mathrm{h=1}}^{\mathrm{L}} P_h\bar{X}E(e_x) - (1-\alpha)\left[\sum_{h=1}^{L} P_h\bar{X} + \sum_{\mathrm{h=1}}^{\mathrm{L}} P_h\bar{X}E(e_x)\right]\left(E(e_x) - \right.$$
$$\left.\left.\frac{1}{2}E(e_x^2) + \frac{1}{3}E(e_x^3) - \cdots\right)\left[\frac{\sum_{\mathrm{h=1}}^{\mathrm{L}} P_h\bar{X}E(e_x)}{\sum_{\boldsymbol{h=1}}^{\boldsymbol{L}} P_h\bar{X}^2(1+e_x)^2}\right]\right\} \quad (23)$$

Now, let the expectation of the error terms as be given as follows

$E\left(e_y\right) = E(e_x) = 0$

$E(e_x^2) = \lambda_{2h}S_{xh}^2 + \theta_h S_{xh2}^2 \quad = A_{hQ}$

$E\left(e_y^2\right) = \lambda_{2h}S_{yh}^2 + \theta_h S_{yh2}^2 \quad = B_{hQ}$

$E\left(e_y e_x\right) = \lambda_{2h}\rho_{xyh}S_{xh}S_{yh} \quad = P_{hQ}$

Substituting into Eq. 22 gives

$$Bias\left(\bar{y}_{prc}^{*}\right) = -\frac{1}{2}(1-\alpha)\sum_{h=1}^{L} P_h \; A_{hQ} \quad (24)$$

And the mean square error is obtain as follows

$$MSE\left(\bar{y}_{prc}^{*}\right) = E\left\{\sum_{\mathrm{h=1}}^{\mathrm{L}} P_h\bar{Y}e_y + (1-\alpha)\sum_{h=1}^{L} P_h\left(e_x - \frac{e_x^2}{2}\right) - \beta\left(\sum_{\mathrm{h=1}}^{\mathrm{L}} P_h\bar{X}e_x\right) - \right.$$
$$\left.(1-\alpha)\left[\sum_{h=1}^{L} P_h\bar{X} + \sum_{\mathrm{h=1}}^{\mathrm{L}} P_h\bar{X}e_x\right]\left(e_x - \frac{e_x^2}{2}\right)\left[\frac{\sum_{\mathrm{h=1}}^{\mathrm{L}} P_h\bar{X}e_x}{\sum_{\boldsymbol{h=1}}^{\boldsymbol{L}} P_h\bar{X}^2(1+e_x)^2}\right]\right\}^2$$

$$= E\left\{\sum_{\mathrm{h=1}}^{\mathrm{L}} P_h\bar{Y}e_y + (1-\alpha)\sum_{h=1}^{L} P_h\left(e_x - \frac{e_x^2}{2}\right) - \beta\left(\sum_{\mathrm{h=1}}^{\mathrm{L}} P_h\bar{X}e_x\right)\right.$$
$$\left. - (1-\alpha)\left[\bar{X} + \sum_{\mathrm{h=1}}^{\mathrm{L}} P_h\bar{X}e_x\right]\left(e_x - \frac{e_x^2}{2}\right)\left[\frac{\sum_{\mathrm{h=1}}^{\mathrm{L}} P_h\bar{X}e_x}{\sum_{\boldsymbol{h=1}}^{\boldsymbol{L}} P_h\bar{X}^2(1+e_x)^2}\right]\right\}^2$$

$$= E\left\{\sum_{\text{h=1}}^{\text{L}} P_h \bar{Y} e_y + (1-\alpha)\left(e_x - \frac{e_x^2}{2}\right)\left(\sum_{h=1}^{L} P_h - \left[\bar{X} + \sum_{\text{h=1}}^{\text{L}} P_h \bar{X} e_x\right]\left[\frac{\sum_{\text{h=1}}^{\text{L}} P_h \bar{X} e_x}{\sum_{\boldsymbol{h=1}}^{\boldsymbol{L}} P_h \bar{X}^2 (1+e_x)^2}\right]\right) - \beta\left(\sum_{\text{h=1}}^{\text{L}} P_h \bar{X} e_x\right)\right\}^2$$

$$= E\left[\sum_{\text{h=1}}^{\text{L}} P_h \bar{Y} e_y + (1-\alpha)\left(e_x - \frac{e_x^2}{2}\right)\left(\sum_{h=1}^{L} P_h - \left[\bar{X} + \sum_{\text{h=1}}^{\text{L}} P_h \bar{X} e_x\right]\left[\frac{\sum_{\text{h=1}}^{\text{L}} P_h \bar{X} e_x}{\sum_{\boldsymbol{h=1}}^{\boldsymbol{L}} P_h \bar{X}^2 (1+e_x)^2}\right]\right) - \beta\left(\sum_{\text{h=1}}^{\text{L}} P_h \bar{X} e_x\right)\right]\left[\sum_{\text{h=1}}^{\text{L}} P_h \bar{Y} e_y + (1-\alpha)\left(e_x - \frac{e_x^2}{2}\right)\left(\sum_{h=1}^{L} P_h - \left[\bar{X} + \sum_{\text{h=1}}^{\text{L}} P_h \bar{X} e_x\right]\left[\frac{\sum_{\text{h=1}}^{\text{L}} P_h \bar{X} e_x}{\sum_{\boldsymbol{h=1}}^{\boldsymbol{L}} P_h \bar{X}^2 (1+e_x)^2}\right]\right) - \beta\left(\sum_{\text{h=1}}^{\text{L}} P_h \bar{X} e_x\right)\right]$$

$$= E\left\{\sum_{h=1}^{L} \bar{Y}^2 e_y^2 - \beta \sum_{h=1}^{L} P_h^2 \bar{Y}_h \bar{X}_h e_y e_x + \beta \sum_{h=1}^{L} P_h \bar{Y}_h \bar{X}_h e_y e_x + \beta^2 \sum_{h=1}^{L} {P_h}^2 {\bar{X}_h}^2 {e_x}^2\right\}$$

$$= \left\{\sum_{h=1}^{L} \bar{Y}^2 E(e_y^2) - \beta \sum_{h=1}^{L} P_h^2 \bar{Y}_h \bar{X}_h E(e_y e_x) + \beta \sum_{h=1}^{L} P_h \bar{Y}_h \bar{X}_h E(e_y e_x) + \beta^2 \sum_{h=1}^{L} {P_h}^2 {\bar{X}_h}^2 E({e_x}^2)\right\}$$

$$= \left\{\sum_{h=1}^{L} \bar{Y}^2 \lambda_{2h} \frac{S_{yh}^2}{\bar{Y}^2} + \theta_h \frac{S_{yh2}^2}{\bar{Y}^2} - \beta \sum_{h=1}^{L} P_h^2 \bar{Y}_h \bar{X}_h \rho_{xyh} \lambda_{2h} \frac{S_{yh}}{\bar{Y}} \frac{S_{xh}}{\bar{X}} + \beta \sum_{h=1}^{L} P_h \bar{Y}_h \bar{X}_h \rho_{xyh} \lambda_{2h} \frac{S_{yh}}{\bar{Y}} \frac{S_{xh}}{\bar{X}} + \beta^2 \sum_{h=1}^{L} {P_h}^2 {\bar{X}_h}^2 \lambda_{2h} \frac{S_{xh}^2}{\bar{X}^2} + \theta_h \frac{S_{xh2}^2}{\bar{X}^2}\right\}$$

$$= \beta^2 \sum_{h=1}^{K} {P_h}^2 A_{hQ} + \sum_{h=1}^{L} B_{hQ} - \beta \sum_{h=1}^{L} P_h^2 \; T_{hQ} + \beta \sum_{h=1}^{K} P_h \; T_{hQ} \tag{25}$$

Calibration of Rajest et al (2020) Existing Estimator

Recall that the calibration estimator of Eq.20 is given by

$\bar{y}_{prc}^{*} = \sum_{h=1}^{L} P_h \bar{y}_h^{**} + \sum_{h=1}^{L} P_h Q_h \, \bar{x}_h^{*} \bar{y}_h^{**} \left[\frac{\bar{X} - \sum_{h=1}^{L} P_h \bar{x}_h^{*}}{\sum_{h=1}^{L} P_h Q_h \bar{x}_h^{*2}}\right]$

and the existing estimator by Rajest et al (2020) is given by

$t^{*} = \sum_{h=1}^{L} P_h \left[\bar{y}_h^{*} + \alpha \log\left(\frac{\bar{x}_h^{*}}{\bar{X}_h}\right)\right]$

Writing $t^{*}$ as a calibration equation, we have

$$t_c^{*} = \sum_{h=1}^{L} w_h \, \bar{y}_h^{***} \quad (26)$$

where in this case, $\bar{y}_h^{***} = \bar{y}_h^{*} + \alpha \log\left(\frac{\bar{x}_h^{*}}{\bar{X}_h}\right)$

Therefore Eq. 26 becomes

$$t_c^{*} = \sum_{h=1}^{L} P_h \left[\bar{y}_h^{*} + \alpha \log\left(\frac{\bar{x}_h^{*}}{\bar{X}_h}\right)\right] + \sum_{h=1}^{L} P_h Q_h \, \bar{x}_h^{*} \left[\bar{y}_h^{*} + \alpha \log\left(\frac{\bar{x}_h^{*}}{\bar{X}_h}\right)\right] \left[\frac{\bar{X} - \sum_{h=1}^{L} P_h \bar{x}_h^{*}}{\sum_{h=1}^{L} P_h Q_h \bar{x}_h^{*2}}\right]$$

$$= \sum_{h=1}^{L} P_h \bar{y}_h^{*} + \alpha \sum_{h=1}^{L} P_h \log\left(\frac{\bar{x}_h^{*}}{\bar{X}_h}\right) + \left[\sum_{h=1}^{L} P_h \bar{x}_h^{*} \, \bar{y}_h^{*} + \alpha \sum_{h=1}^{L} P_h \bar{x}_h^{*} \log\left(\frac{\bar{x}_h^{*}}{\bar{X}_h}\right)\right] \left[\frac{\bar{X} - \sum_{h=1}^{L} P_h \bar{x}_h^{*}}{\sum_{h=1}^{L} P_h Q_h \bar{x}_h^{*2}}\right]$$

$= \sum_{h=1}^{L} P_h \bar{y}_h^{*} + \alpha \sum_{h=1}^{L} P_h \log\left(\frac{\bar{x}_h^{*}}{\bar{X}^{*}}\right) + \alpha(\bar{X} - \sum_{h=1}^{L} P_h \bar{x}_h^{*}) +$
$\alpha \sum_{h=1}^{L} P_h \bar{x}_h^{*} \log\left(\frac{\bar{x}_h^{*}}{\bar{X}^{*}}\right) \left[\frac{\bar{X} - \sum_{h=1}^{L} P_h \bar{x}_h^{*}}{\sum_{h=1}^{L} P_h \bar{x}_h^{*2}}\right]$

Let $\bar{y}_h^{*} = \bar{Y}\left(1 + e_y\right)$, $\bar{x}_h^{*} = \bar{X}\,(1 + e_x)$

Such that

$t_c^{*} = \sum_{h=1}^{L} P_h \bar{Y}\left(1 + e_y\right) + \alpha \sum_{h=1}^{L} P_h \log\left(\frac{\bar{X}\,(1+e_x)}{\bar{X}^{*}}\right) + \beta(\bar{X} - \sum_{h=1}^{L} P_h \bar{X}\,(1 + e_x)) +$
$\alpha \sum_{h=1}^{L} P_h \bar{X}\,(1 + e_x) \log\left(\frac{\bar{X}\,(1+e_x)}{\bar{X}^{*}}\right) \left[\frac{\bar{X} - \sum_{h=1}^{L} P_h \bar{X}\,(1+e_x)}{\sum_{h=1}^{L} P_h \bar{X}^{2}(1+e_x)^{2}}\right]$

Also, assume that $\bar{X}^{*} = \bar{X}$,

$$= \sum_{h=1}^{L} P_h \bar{Y} + \sum_{h=1}^{L} P_h \bar{Y} e_y + \alpha \sum_{h=1}^{L} P_h \log(1 + e_x) + \beta\left(\bar{X} - \sum_{h=1}^{L} P_h \bar{X} - \sum_{h=1}^{L} P_h \bar{X} e_x\right) + \alpha\left[\sum_{h=1}^{L} P_h \bar{X} + \sum_{h=1}^{L} P_h \bar{X} e_x\right] \log(1 + e_x) \left[\frac{\bar{X} - \sum_{h=1}^{L} P_h \bar{X} - \sum_{h=1}^{L} P_h \bar{X} e_x}{\sum_{h=1}^{L} P_h \bar{X}^{2}(1 + e_x)^{2}}\right]$$

$= \sum_{h=1}^{L} P_h \bar{Y} + \sum_{\mathrm{h=1}}^{\mathrm{L}} P_h \bar{Y} e_y \; + \alpha \sum_{h=1}^{L} P_h \log(1+e_x) + \beta\left(\bar{X} - \sum_{h=1}^{L} P_h \bar{X} - \sum_{\mathrm{h=1}}^{\mathrm{L}} P_h \bar{X} e_x\right) + \alpha\left[\sum_{h=1}^{L} P_h \bar{X} + \sum_{\mathrm{h=1}}^{\mathrm{L}} P_h \bar{X} e_x\right] \log(1+e_x) \left[\frac{\bar{X} - \sum_{h=1}^{L} P_h \bar{X} - \sum_{\mathrm{h=1}}^{\mathrm{L}} P_h \bar{X} e_x}{\sum_{h=1}^{L} P_h \bar{X}^2 (1+e_x)^2}\right]$

But $\log(1+e_x) = \; e_x - \frac{e_x^2}{2} + \frac{e_x^3}{3} - \cdots$

$$= \sum_{h=1}^{L} P_h \bar{Y} + \sum_{\mathrm{h=1}}^{\mathrm{L}} P_h \bar{Y} e_y \; + \alpha \sum_{h=1}^{L} P_h \left(e_x - \frac{e_x^2}{2} + \frac{e_x^3}{3} - \cdots\right) + \beta\left(\bar{X} - \sum_{h=1}^{L} P_h \bar{X} - \sum_{\mathrm{h=1}}^{\mathrm{L}} P_h \bar{X} e_x\right) + \alpha\left[\sum_{h=1}^{L} P_h \bar{X} + \sum_{\mathrm{h=1}}^{\mathrm{L}} P_h \bar{X} e_x\right]\left(e_x - \frac{e_x^2}{2} + \frac{e_x^3}{3} - \cdots\right)\left[\frac{\bar{X} - \sum_{h=1}^{L} P_h \bar{X} - \sum_{\mathrm{h=1}}^{\mathrm{L}} P_h \bar{X} e_x}{\sum_{h=1}^{L} P_h \bar{X}^2 (1+e_x)^2}\right]$$

$$(t_c^* - \bar{Y}) = \left\{\sum_{h=1}^{L} P_h \bar{Y} + \sum_{\mathrm{h=1}}^{\mathrm{L}} P_h \bar{Y} e_y \; + \alpha \sum_{h=1}^{L} P_h \left(e_x - \frac{e_x^2}{2} + \frac{e_x^3}{3} - \cdots\right) + \beta\left(\bar{X} - \sum_{h=1}^{L} P_h \bar{X} - \sum_{\mathrm{h=1}}^{\mathrm{L}} P_h \bar{X} e_x\right) + \alpha\left[\sum_{h=1}^{L} P_h \bar{X} + \sum_{\mathrm{h=1}}^{\mathrm{L}} P_h \bar{X} e_x\right]\left(e_x - \frac{e_x^2}{2} + \frac{e_x^3}{3} - \cdots\right)\left[\frac{\bar{X} - \sum_{h=1}^{L} P_h \bar{X} - \sum_{\mathrm{h=1}}^{\mathrm{L}} P_h \bar{X} e_x}{\sum_{h=1}^{L} P_h \bar{X}^2 (1+e_x)^2}\right] - \bar{Y}\right\}$$

$= \left\{\sum_{\mathrm{h=1}}^{\mathrm{L}} P_h \bar{Y} e_y \; \alpha \sum_{h=1}^{L} P_h \left(e_x - \frac{e_x^2}{2} + \frac{e_x^3}{3} - \cdots\right) + \beta\left(\sum_{\mathrm{h=1}}^{\mathrm{L}} P_h \bar{X} e_x\right) \alpha\left[\sum_{h=1}^{L} P_h \bar{X} + \sum_{\mathrm{h=1}}^{\mathrm{L}} P_h \bar{X} e_x\right]\left(e_x - \frac{e_x^2}{2} + \frac{e_x^3}{3} - \cdots\right)\left[\frac{-\sum_{\mathrm{h=1}}^{\mathrm{L}} P_h \bar{X} e_x}{\sum_{h=1}^{L} P_h \bar{X}^2 (1+e_x)^2}\right]\right\}$

$\mathrm{Bias}(t_c^*) = \left\{\sum_{\mathrm{h=1}}^{\mathrm{L}} P_h \bar{Y} E(e_y) \; + \alpha \sum_{h=1}^{L} P_h \left(E(e_x) - \frac{1}{2} E(e_x^2) + \frac{1}{3} E(e_x^3) - \cdots\right) - \beta \sum_{\mathrm{h=1}}^{\mathrm{L}} P_h \bar{X} E(e_x) - \alpha\left[\sum_{h=1}^{L} P_h \bar{X} + \sum_{\mathrm{h=1}}^{\mathrm{L}} P_h \bar{X} E(e_x)\right]\left(E(e_x) - \frac{1}{2} E(e_x^2) + \frac{1}{3} E(e_x^3) - \cdots\right)\left[\frac{\sum_{\mathrm{h=1}}^{\mathrm{L}} P_h \bar{X} E(e_x)}{\sum_{h=1}^{L} P_h \bar{X}^2 (1+e_x)^2}\right]\right\}$

Substituting the expectation of the error terms gives

$$Bias(t_c^*) = -\frac{1}{2} \alpha \sum_{h=1}^{L} P_h \, A_{hQ} \quad (27)$$

$$MSE(t_c^*) = E\left\{\sum_{h=1}^{L} P_h \bar{Y} e_y \ + \alpha \sum_{h=1}^{L} P_h \left(e_x - \frac{e_x^2}{2}\right) - \beta\left(\sum_{h=1}^{L} P_h \bar{X} e_x\right) - \alpha\left[\sum_{h=1}^{L} P_h \bar{X} + \sum_{h=1}^{L} P_h \bar{X} e_x\right]\left(e_x - \frac{e_x^2}{2}\right)\left[\frac{\sum_{h=1}^{L} P_h \bar{X} e_x}{\sum_{h=1}^{L} P_h \bar{X}^2 (1+e_x)^2}\right]\right\}^2$$

$$= E\left\{\sum_{h=1}^{L} P_h \bar{Y} e_y \ + \alpha \sum_{h=1}^{L} P_h \left(e_x - \frac{e_x^2}{2}\right) - \beta\left(\sum_{h=1}^{L} P_h \bar{X} e_x\right) - \alpha\left[\bar{X} + \sum_{h=1}^{L} P_h \bar{X} e_x\right]\left(e_x - \frac{e_x^2}{2}\right)\left[\frac{\sum_{h=1}^{L} P_h \bar{X} e_x}{\sum_{h=1}^{L} P_h \bar{X}^2 (1+e_x)^2}\right]\right\}^2$$

$$= E\left\{\sum_{h=1}^{L} \bar{Y}^2 e_y^2 - \beta \sum_{h=1}^{L} P_h^2 \bar{Y}_h \bar{X}_h e_y e_x + \beta \sum_{h=1}^{L} P_h \bar{Y}_h \bar{X}_h e_y e_x + \beta^2 \sum_{h=1}^{L} {P_h}^2 {\bar{X}_h}^2 {e_x}^2\right\}$$

$$= \left\{\sum_{h=1}^{L} \bar{Y}^2 \lambda_{2h} \frac{S_{yh}^2}{\bar{Y}^2} + \theta_h \frac{S_{yh2}^2}{\bar{Y}^2} - \beta \sum_{h=1}^{L} P_h^2 \bar{Y}_h \bar{X}_h \lambda_{2h} \frac{S_{yh}}{\bar{Y}} \frac{S_{xh}}{\bar{X}} + \beta \sum_{h=1}^{L} P_h \bar{Y}_h \bar{X}_h \lambda_{2h} \frac{S_{yh}}{\bar{Y}} \frac{S_{xh}}{\bar{X}} + \beta^2 \sum_{h=1}^{L} {P_h}^2 {\bar{X}_h}^2 \lambda_{2h} \frac{S_{xh}^2}{\bar{X}^2} + \theta_h \frac{S_{xh2}^2}{\bar{X}^2}\right\}$$

$$= \sum_{h=1}^{L} \ B_{hQ} - \beta \sum_{h=1}^{L} P_h^2 \ T_{hQ} + \beta \sum_{h=1}^{K} P_h \ T_{hQ} + \beta^2 \sum_{h=1}^{K} {P_h}^2 \ A_{hQ} \tag{28}$$

## 3. Empirical Study

In this section we have carried out an empirical study for which we have considered two natural population data set.

Population -1 (Sarndal, C. E., Swenssen, B. Wretman, J. (2003))

Y: production of wheat (in tons), X: area of wheat (in hectares)

No of strata = 4.

$N_1 = 47, N_2 = 30, N_3 = 29, N_4 = 13, n_1 = 15, n_2 = 10, n_3 = 10, n_4 = 5, \bar{Y}_1 = 443.5447, \bar{Y}_2 = 68.68276, \quad \bar{Y}_3 = 17.06667, \bar{Y}_4 = 52.52308, \bar{X}_1 = 160.2362, \bar{X}_2 = 29.70345, \bar{X}_3 = 11.54667,$

$\bar{X}_4 = 23.62308, S_{1Y}^2 = 74026.75, S_{2Y}^2 = 28871.781, S_{3Y}^2 = 244.1292, S_{4Y}^2 = 4451.124, S_{1X}^2 = 8377.401, S_{2X}^2 = 315.4532, S_{3X}^2 = 91.45775, S_{3X}^2 = 682.9703, P_{1YX} = 0.9583838, P_{2YX} = 0.779071, P_{3YX} = 0.8719665, P_{4YX} = 0.9922591$

Population -2 (FBS, crops area population by districts, Islamabad:2011)

Y: 1983 Population (Millions), X: 1982 gross national product.

$$N_1 = 38, N_2 = 14, N_3 = 11, N_4 = 33, N_5 = 24, n_1 = 17, n_2 = 6, n_3 = 4, n_4 = 12, n_5 = 11,$$

$$\bar{Y}_1 = 13.03684, \bar{Y}_2 = 37.35, \bar{Y}_3 = 23.13636, \bar{Y}_4 = 79.65455, \bar{Y}_5 = 20.28333, \bar{X}_1 = 1029.158,$$

$$\bar{X}_2 = 25671.57, \bar{X}_3 = 5028.818, \bar{X}_4 = 7533.939, \bar{X}_5 = 16315.25, S_{1Y}^2 = 270.9083,$$

$$S_{2Y}^2 = 3906.929, S_{3Y}^2 = 1339.405, S_{4Y}^2 = 45082.17, S_{5Y}^2 = 368.9423, S_{1X}^2 = 3667.896,$$

$$S_{2X}^2 = 656846.1403, S_{3X}^2 = 633487.43, S_{3X}^2 = 440717912, S_{5Y}^2 = 408441212,$$

$P_{1YX} = 0.7439544, P_{2YX} = 0.969956, P_{3YX} = 0.9768227, P_{4YX} = 0.2948897, P_{5YX} = 0.901107$

The MSE expression for the existing estimators for the section 1 and 2 i.e for the case of nonresponse and measurement errors can be obtained from the section of the existing estimators by using the appropriate notations from sections 1 and 2 respectfully.

To determined the percent relative efficiency (PRE), of the estimator with respect to the usual estimator ($\bar{y}_{HH}^*, \bar{y}_{st,}^*$) we have use the given formula.

$PRE\ (\theta) = \frac{mse\ usual\ estimstor}{MSE(\theta)} \times 100, \theta = \ \bar{y}_{SR,}^* \bar{y}_{S(D)}^*\ , \bar{y}_{S(AH)}^*,\ \bar{y}_{\mathrm{SP},}^*\ t^*,\ t_c^*,\ \ \bar{y}_{\mathrm{pr},}^*, \bar{y}_{\mathrm{prc}}^*,\ t_c^*$

## 4. Results

**Table 1.** MSE and PRE of estimators when there is presence of nonresponse on both the study and auxiliary variables for population 1

| $g_h = 2$ | | |
|---|---|---|
| Estimators | MSE | PRE |
| $\bar{y}_{HH}^*$ | 551.8020 | 100.000.00 |
| $\bar{y}_{s(R)}^*$ | 61.7760 | 893.2299 |
| $\bar{y}_{s(D)}^*$ | 61.5828 | 896.0318 |
| $\bar{y}_{s(AH)}^*$ | 467.9744 | 117.9129 |
| $\bar{y}_{s(P)}^*$ ($\propto_h = 0$) | 61.32050 | 899.7985 |
| $\bar{y}_{s(P)}^*$ ($\propto_h = 1$) | 61.3900 | 898.8466 |
| $\bar{y}_{s(P)}^*$ ($\propto_h = -1$) | 61.3910 | 898.8313 |
| $t^*$ | 61.5858 | 896.0318 |
| $\bar{y}_{\mathrm{pr}}^*$ | 40.367 | 887.875 |
| $\bar{y}_{\mathrm{prc}}^*$ | 29.571 | 894.224 |
| $t_c^*$ | 29.571 | 894.224 |
| $g_h = 4$ | | |
| Estimators | MSE | PRE |
| $\bar{y}_{HH}^*$ | 567.8053 | 100.000. |
| $\bar{y}_{s(R)}^*$ | 88.6968 | 640.1639 |
| $\bar{y}_{s(D)}^*$ | 88.5978 | 640.8790 |

| | | |
|---|---|---|
| $\bar{y}^*_{s(AH)}$ | 530.3673 | 107.0589 |
| $\bar{y}^*_{s(P)}$ ($\propto_h = 0$) | 88.1623 | 644.0449 |
| $\bar{y}^*_{s(P)}$ ($\propto_h = 1$) | 88.2567 | 643.3562 |
| $\bar{y}^*_{s(P)}$ ($\propto_h = -1$) | 88.2590 | 643.3396 |
| $t^*$ | 88.5978 | 640.8790 |
| $\bar{y}^*_{pr}$ | 50. 6775 | 778.7689 |
| $\bar{y}^*_{prc}$ | 40.0890 | 851.1137 |
| $t^*_c$ | 40.0890 | 851.1137 |
| $g_h = 8$ | | |
| Estimators | MSE | PRE |
| $\bar{y}^*_{HH}$ | 599.8118 | 100.000. |
| $\bar{y}^*_{s(R)}$ | 143.5385 | 420.8069 |
| $\bar{y}^*_{s(D)}$ | 140.8417 | 425.8766 |
| $\bar{y}^*_{s(AH)}$ | 655.1529 | 91.5529 |
| $\bar{y}^*_{s(P)}$ ($\propto_h = 0$) | 139.9058 | 428.7255 |
| $\bar{y}^*_{s(P)}$ ($\propto_h = 1$) | 140.0595 | 428.2550 |
| $\bar{y}^*_{s(P)}$ ($\propto_h = -1$) | 140.0662 | 428.2345 |
| $t^*$ | 140.8417 | 425.8766 |
| $\bar{y}^*_{pr}$ | 109.854 | 460.2247 |
| $\bar{y}^*_{prc}$ | 97.670 | 479.1143 |
| $t^*_c$ | 97.670 | 479.1143 |

**Table 2.** MSE and PRE of estimator when there is presence of nonresponse on both the study and auxiliary variables for population 2

| $g_h = 2$ | | |
|---|---|---|
| Estimators | MSE | PRE |
| $\bar{y}^*_{HH}$ | 192.2504 | 100.000.00 |
| $\bar{y}^*_{s(R)}$ | 288.7807 | 66.5731 |
| $\bar{y}^*_{s(D)}$ | 169.2728 | 113.5743 |
| $\bar{y}^*_{s(AH)}$ | 1025.251 | 18.7515 |
| $\bar{y}^*_{s(P)}$ ($\propto_h = 0$) | 120.9315 | 158.9756 |
| $\bar{y}^*_{s(P)}$ ($\propto_h = 1$) | 120.644 | 159.3265 |
| $\bar{y}^*_{s(P)}$ ($\propto_h = -1$) | 125.6208 | 153.0403 |
| $t^*$ | 169.2728 | 113.5743 |
| $\bar{y}^*_{pr}$ | 100.6741 | 200.2365 |
| $\bar{y}^*_{prc}$ | 97.9111 | 202.6009 |
| $t^*_c$ | 97.9111 | 202.6009 |
| $g_h = 4$ | | |
| Estimators | MSE | PRE |
| $\bar{y}^*_{HH}$ | 198.2250 | 100.000.00 |
| $\bar{y}^*_{s(R)}$ | 301.6620 | 65.7109 |

| | | |
|---|---|---|
| $\bar{y}^*_{s(D)}$ | 175.813 | 112.7476 |
| $\bar{y}^*_{s(AH)}$ | 1066.2370 | 18.5910 |
| $\bar{y}^*_{s(P)}$ ($\propto_h$= 0) | 124.3443 | 159.4162 |
| $\bar{y}^*_{s(P)}$ ($\propto_h$= 1) | 1239.130 | 159.9711 |
| $\bar{y}^*_{s(P)}$ ($\propto_h$= −1) | 129.3531 | 153.2433 |
| $t^*$ | 175.813 | 112.7476 |
| $\bar{y}^*_{pr}$ | 105.7800 | 181.8001 |
| $\bar{y}^*_{prc}$ | 99.9677 | 189.6710 |
| $t^*_c$ | 99.9677 | 189.6710 |
| $g_h = 8$ | | |
| Estimators | MSE | PRE |
| $\bar{y}^*_{HH}$ | 210.1741 | 100.000.00 |
| $\bar{y}^*_{s(R)}$ | 327.4247 | 64.1900 |
| $\bar{y}^*_{s(D)}$ | 188.7929 | 111.3252 |
| $\bar{y}^*_{s(AH)}$ | 1148.21 | 18.3045 |
| $\bar{y}^*_{s(P)}$ ($\propto_h$= 0) | 120.8596 | 160.6104 |
| $\bar{y}^*_{s(P)}$ ($\propto_h$= 1) | 129.9988 | 161.6793 |
| $\bar{y}^*_{s(P)}$ ($\propto_h$= −1) | 136.5215 | 153.9495 |
| $t^*$ | 188.7929 | 111.3252 |
| $\bar{y}^*_{pr}$ | 108.8675 | 190.1134 |
| $\bar{y}^*_{prc}$ | 100,6519 | 199.7699 |
| $t^*_c$ | 100,6519 | 199.7699 |

From Tables 1 and 2, for both populations 1 and 2 and for the values of $g = 2, 4$ *and* 8, the proposed estimators $\bar{y}^*_{pr}$ , $\bar{y}^*_{prc}$ *and* $t^*_c$ perform exceedingly better in terms of gains in efficiency compared to the existing estimators considered in this study. Again, it is also observed that the proposed calibration estimators $\bar{y}^*_{prc}$ and $t^*_c$ , which are gotten from the proposed and Rajest et al (2020) estimators respectively have the same PRE which makes calibration technique a veritable tool in smoothing process.

## 5. Conclusion

Obtaining an efficient estimator has been a major challenge in sample survey due to measurement error and nonresponse. This study has proven to be very effective in terms of gains in efficiency and has unraveled the potentials of the calibration technique as a smoothing tool in making estimators of the same class to perform exceedingly better in the same manner irrespective of the weight adjustments in the parent estimator. This is seen in the performance of the two proposed calibration estimators. Consequently, it is preferable to use the proposed estimators in real life practice when measurement error or nonresponse is detected in the survey data.

## References

Azeem M. & Hanif M. (2017). Joint influence of measurement error and non-response on estimation of population mean. Communications in Statistics-Theory and Methods. 46(4):1679–1693. https://doi.org/10.1080/03610926.2015.1026992

Chaudhary, M. K., Singh, R., Shukla, R. K., Kumar, M. & Smarandache, F. (2009). A family of estimators for estimating population mean in stratified sampling under nonresponse. Pakistan Journal of statistics and operations research, 5, 147-54

FBS. (2011). Crops area production by districts, Islamabad.

Hansen, M. H,. & Hurwitz, W. N. (1946). The problem of nonresponse in sample surveys. Journal of the American Statistical Association, 41, 517-529.

Iseh, M. J, & Bassey, K. J. (2021a). A new calibration estimator of population mean for small area with non-response. Asian Journal of Probability and Statistics, 12(2):14-51.

Iseh, M. J, & Bassey, K. J. (2021b). Calibration estimator for population mean in small sample size with Non-response. European Journal of Statistics and Probability, 9(1), 32-42.

Iseh, M. J & Bassey M. O. (2022). Calibration estimators for population mean with sub-sampling the nonrespondents under stratified sampling. Science Journal of Applied Mathematics and Statistics, 10(4), 45-56. DOIi: 10.11648/j.sjams.20221004.11

Khare, B. B. & Srivastava, S. (1997): Transformed ratio type estimators for the population mean in the presence of nonresponse. Communications in Statistics-Theory and Methods, 26, 1779-1791.

Kumar, S., Singh, H. P. & Gupta, R. (2011). A class of ratio-cum-product type estimators under double sampling in the presence of nonresponse. Hacettepe Journal of Mathematics and Statistics, 40, 115-129.

Mishra, M., Singh, B. P. & Singh, R. (2017). Estimation of population mean using two auxiliary variables in systematic sampling. Journal of Reliability and Statistical Studies, 10, 10-21.

Rao, P. S. R. S. (1986). Ratio estimation with subsampling the nonrespondents. Survey methodology, 12, 217-230.

Rajesh S,, Madhulika M. & Madhulika M. (2020). Estimation in stratified random sampling in the presence of errors. Revista Investigacion Operacional. 41, 125-137.

Sarndal, C-E., Swensson, B. & Wretman, J. (2003). Model Assisted Survey Sampling. Springer Science and Business Media, N. York.

Shalabh, S. (1997). Ratio method of estimation in the presence of measurement errors. Journal of Indian Statistics 52, 150-155.

Singh, H. P. & Kumar, S. (2010). Estimation of mean in the presence of nonresponse using two phase sampling scheme. Statistical papers, 51, 559-582.

Singh, H. P., & Karpe, N. (2009). On the estimation ratio and product of two population means using supplementary information in the presence of measurement error. Statistica 69, 27-47.

Singh, H. P., & Kumar, S. (2008). A regression approach to the estimation of the finite population mean in the presence of nonresponse. Australian & New Zealand Journal of Statistics 50, 395-408

Singh, H. P., Kumar, S. and Kozak, M. (2010). Improved estimation of finite-population mean using subsampling to deal with nonresponse in two-phase sampling scheme. Communications in Statistics-Theory and Methods, 39, 791-802.

Singh, R. & Sharma. (2015). Method of estimation in the presence of nonresponse and measurement errors simultaneously. Journal of Modern Applied Statistical Methods, 14, 1-12.

Sud, U. C., & S. K. Srivastava. (2000). Estimation of population mean in repeat surveys in the presence of measurement errors. Jour. Ind. Soc. Ag. Statist 53, 125-133.

Tabasum, R., & Khan, I. A. (2006). Double sampling ratio estimator for the population mean in the presence of nonresponse. Assam Statistical Review, 20, 73-83.

Zahid, E., & Shabbir, J. (2018). Estimation of mean in the presence of measurement error and nonresponse under stratified random sampling. Plos one, 13(2) e0191572. doi: 10.1371/journal.pone.0191572