

## A new calibration estimator in stratified double sampling

Nursel Koyuncu <sup>a \*</sup> and Cem Kadilar <sup>a †</sup>

### Abstract

In the present article, we consider a new calibration estimator of the population mean in the stratified double sampling. We get more efficient calibration estimator using new calibration weights compared to the straight estimator. In addition, the estimators derived are analyzed for different populations by a simulation study. The simulation study shows that new calibration estimator is highly efficient than the existing estimator.

**Keywords:** Calibration, Auxiliary information, Stratified double sampling.

*2000 AMS Classification:* 62D05

### 1. Introduction

When the auxiliary information is available, the calibration estimator is widely used in the sampling literature to improve the estimates. Many authors, such as Deville and Sarndal [2], Estevao and Sarndal [3], Arnab and Singh [1], Farrell and Singh [4], Kim et al.[5], Kim and Park [6], Koyuncu and Kadilar etc.[8], defined some calibration estimators using different constraints. In the stratified random sampling, calibration approach is used to get optimum strata weights. Tracy et al.[9] defined calibration estimators in the stratified random sampling and stratified double sampling. In this study, we try to improve the calibration estimator in the stratified double sampling.

---

<sup>a</sup>Hacettepe University, Department of Statistics, Beytepe, 06800, Ankara, Turkey.

\*Email: [nkoyuncu@hacettepe.edu.tr](mailto:nkoyuncu@hacettepe.edu.tr)

†Email: [kadilar@hacettepe.edu.tr](mailto:kadilar@hacettepe.edu.tr)

## 2. Notations

Consider a finite population of  $N$  units consists of  $L$  strata such that the  $h$ th stratum consists of  $N_h$  units and  $\sum_{h=1}^L N_h = N$ . From the  $h$ th stratum of  $N_h$  units, draw a preliminary large sample of  $m_h$  units by the simple random sampling without replacement (SRSWOR) and measure the auxiliary character,  $x_{hi}$ , only. Select a sub-sample of  $n_h$  units from the given preliminary large sample of  $m_h$  units by SRSWOR and measure both the study variable,  $y_{hi}$  and auxiliary variable,  $x_{hi}$ . Let  $\bar{x}_h^* = \frac{1}{m_h} \sum_{i=1}^{m_h} x_{hi}$  and  $s_{hx}^{*2} = \frac{1}{m_h-1} \sum_{i=1}^{m_h} (x_{hi} - \bar{x}_h^*)^2$  denote the first phase sample mean and variance, respectively. Besides, assume that  $\bar{x}_h = \frac{1}{n_h} \sum_{i=1}^{n_h} x_{hi}$ ,  $s_{hx}^2 = \frac{1}{n_h-1} \sum_{i=1}^{n_h} (x_{hi} - \bar{x}_h)^2$  and  $\bar{y}_h = \frac{1}{n_h} \sum_{i=1}^{n_h} y_{hi}$ ,  $s_{hy}^2 = \frac{1}{n_h-1} \sum_{i=1}^{n_h} (y_{hi} - \bar{y}_h)^2$  denote the second phase sample means and variances for the auxiliary and study characters, respectively.

Calibration estimator, defined by Tracy et al.[9], is given by

$$(2.1) \quad \bar{y}_{st}(d) = \sum_{h=1}^L W_h^* \bar{y}_h,$$

where  $W_h^*$  are calibration weights minimizing the chi-square distance measure

$$(2.2) \quad \sum_{h=1}^L \frac{(W_h^* - W_h)^2}{Q_h W_h}$$

subject to calibration constraints defined by

$$(2.3) \quad \sum_{h=1}^L W_h^* \bar{x}_h = \sum_{h=1}^L W_h \bar{x}_h^*,$$

$$(2.4) \quad \sum_{h=1}^L W_h^* s_{hx}^2 = \sum_{h=1}^L W_h s_{hx}^{*2}.$$

The Lagrange function using calibration constraints and chi-square distance measure is given by

$$(2.5) \quad \Delta = \sum_{h=1}^L \frac{(W_h^* - W_h)^2}{Q_h W_h} - 2\lambda_1 \left( \sum_{h=1}^L W_h^* \bar{x}_h - \sum_{h=1}^L W_h \bar{x}_h^* \right) - 2\lambda_2 \left( \sum_{h=1}^L W_h^* s_{hx}^2 - \sum_{h=1}^L W_h s_{hx}^{*2} \right),$$

where  $\lambda_1$  and  $\lambda_2$  are Lagrange multipliers. Setting the derivative of  $\Delta$  with respect to  $W_h^*$  equals to zero gives

$$(2.6) \quad W_h^* = W_h + Q_h W_h (\lambda_1 \bar{x}_h + \lambda_2 s_{hx}^2).$$

Substituting (2.6) in (2.3) and (2.4) respectively, we get

$$\begin{bmatrix} \left( \sum_{h=1}^L Q_h W_h \bar{x}_h^2 \right) & \left( \sum_{h=1}^L Q_h W_h \bar{x}_h s_{hx}^2 \right) \\ \left( \sum_{h=1}^L Q_h W_h \bar{x}_h s_{hx}^2 \right) & \left( \sum_{h=1}^L Q_h W_h s_{hx}^4 \right) \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix} = \begin{bmatrix} \sum_{h=1}^L W_h \bar{x}_h^* - \sum_{h=1}^L W_h \bar{x}_h \\ \sum_{h=1}^L W_h s_{hx}^{*2} - \sum_{h=1}^L W_h s_{hx}^2 \end{bmatrix}$$

Solving the system of equations for lambdas, we obtain

$$\lambda_1 = \frac{\left( \sum_{h=1}^L Q_h W_h s_{hx}^4 \right) \left( \sum_{h=1}^L W_h \bar{x}_h^* - \sum_{h=1}^L W_h \bar{x}_h \right) - \left( \sum_{h=1}^L W_h s_{hx}^{*2} - \sum_{h=1}^L W_h s_{hx}^2 \right) \left( \sum_{h=1}^L Q_h W_h \bar{x}_h s_{hx}^2 \right)}{\left( \sum_{h=1}^L Q_h W_h s_{hx}^4 \right) \left( \sum_{h=1}^L Q_h W_h \bar{x}_h^2 \right) - \left( \sum_{h=1}^L Q_h W_h \bar{x}_h s_{hx}^2 \right)^2},$$

$$\lambda_2 = \frac{\left( \sum_{h=1}^L W_h s_{hx}^{*2} - \sum_{h=1}^L W_h s_{hx}^2 \right) \left( \sum_{h=1}^L Q_h W_h \bar{x}_h^2 \right) - \left( \sum_{h=1}^L Q_h W_h \bar{x}_h s_{hx}^2 \right) \left( \sum_{h=1}^L W_h \bar{x}_h^* - \sum_{h=1}^L W_h \bar{x}_h \right)}{\left( \sum_{h=1}^L Q_h W_h s_{hx}^4 \right) \left( \sum_{h=1}^L Q_h W_h \bar{x}_h^2 \right) - \left( \sum_{h=1}^L Q_h W_h \bar{x}_h s_{hx}^2 \right)^2}.$$

Substituting these values into (2.6), we get the weights as given by

$$W_h^* = W_h + \frac{(Q_h W_h \bar{x}_h) \left[ \left( \sum_{h=1}^L Q_h W_h s_{hx}^4 \right) \left( \sum_{h=1}^L W_h (\bar{x}_h^* - \bar{x}_h) \right) - \left( \sum_{h=1}^L W_h (s_{hx}^{*2} - s_{hx}^2) \right) \left( \sum_{h=1}^L Q_h W_h \bar{x}_h s_{hx}^2 \right) \right]}{\left( \sum_{h=1}^L Q_h W_h s_{hx}^4 \right) \left( \sum_{h=1}^L Q_h W_h \bar{x}_h^2 \right) - \left( \sum_{h=1}^L Q_h W_h \bar{x}_h s_{hx}^2 \right)^2}$$

$$+ \frac{(Q_h W_h s_{hx}^2) \left[ \left( \sum_{h=1}^L W_h (s_{hx}^{*2} - s_{hx}^2) \right) \left( \sum_{h=1}^L Q_h W_h \bar{x}_h^2 \right) - \left( \sum_{h=1}^L Q_h W_h \bar{x}_h s_{hx}^2 \right) \left( \sum_{h=1}^L W_h (\bar{x}_h^* - \bar{x}_h) \right) \right]}{\left( \sum_{h=1}^L Q_h W_h s_{hx}^4 \right) \left( \sum_{h=1}^L Q_h W_h \bar{x}_h^2 \right) - \left( \sum_{h=1}^L Q_h W_h \bar{x}_h s_{hx}^2 \right)^2}$$

Writing these weights in (2.1), we get the calibration estimator as

$$\bar{y}_{st}(d) = \left( \sum_{h=1}^L W_h \bar{y}_h \right) + \beta_{1(d)} \left( \sum_{h=1}^L W_h (\bar{x}_h^* - \bar{x}_h) \right) + \beta_{2(d)} \left( \sum_{h=1}^L W_h (s_{hx}^{*2} - s_{hx}^2) \right),$$

where betas are given by

$$\beta_{1(d)} = \frac{\left( \sum_{h=1}^L Q_h W_h s_{hx}^4 \right) \left( \sum_{h=1}^L Q_h W_h \bar{x}_h \bar{y}_h \right) - \left( \sum_{h=1}^L Q_h W_h \bar{x}_h s_{hx}^2 \right) \left( \sum_{h=1}^L Q_h W_h \bar{y}_h s_{hx}^2 \right)}{\left( \sum_{h=1}^L Q_h W_h s_{hx}^4 \right) \left( \sum_{h=1}^L Q_h W_h \bar{x}_h^2 \right) - \left( \sum_{h=1}^L Q_h W_h \bar{x}_h s_{hx}^2 \right)^2},$$

$$\beta_{2(d)} = \frac{(\sum_{h=1}^L Q_h W_h \bar{x}_h^2)(\sum_{h=1}^L Q_h W_h \bar{y}_h s_{hx}^2) - (\sum_{h=1}^L Q_h W_h \bar{x}_h s_{hx}^2)(\sum_{h=1}^L Q_h W_h \bar{x}_h \bar{y}_h)}{(\sum_{h=1}^L Q_h W_h s_{hx}^4)(\sum_{h=1}^L Q_h W_h \bar{x}_h^2) - (\sum_{h=1}^L Q_h W_h \bar{x}_h s_{hx}^2)^2}.$$

### 3. Suggested Estimator

Motivated by Tracy et al.[9], we consider a new calibration estimator as

$$(3.1) \quad \bar{y}_{st}(dnew) = \sum_{h=1}^L \Omega_h \bar{y}_h.$$

Using the chi-square distance

$$(3.2) \quad \sum_{h=1}^L \frac{(\Omega_h - W_h)^2}{Q_h W_h},$$

and subject to calibration constraints defined by Koyuncu[7]

$$(3.3) \quad \sum_{h=1}^L \Omega_h \bar{x}_h = \sum_{h=1}^L W_h \bar{x}_h^*,$$

$$(3.4) \quad \sum_{h=1}^L \Omega_h s_{hx}^2 = \sum_{h=1}^L W_h s_{hx}^{*2},$$

$$(3.5) \quad \sum_{h=1}^L \Omega_h = \sum_{h=1}^L W_h,$$

we can write the Lagrange function given by

$$\begin{aligned} \Delta = \sum_{h=1}^L \frac{(\Omega_h - W_h)^2}{Q_h W_h} - 2\lambda_1 \left( \sum_{h=1}^L \Omega_h \bar{x}_h - \sum_{h=1}^L W_h \bar{x}_h^* \right) - 2\lambda_2 \left( \sum_{h=1}^L \Omega_h s_{hx}^2 - \sum_{h=1}^L W_h s_{hx}^{*2} \right) \\ - 2\lambda_3 \left( \sum_{h=1}^L \Omega_h - \sum_{h=1}^L W_h \right), \end{aligned}$$

Setting  $\frac{\partial \Delta}{\partial \Omega_h} = 0$ , we obtain

$$(3.6) \quad \Omega_h = W_h + Q_h W_h (\lambda_1 \bar{x}_h + \lambda_2 s_{hx}^2 + \lambda_3).$$

Substituting (3.6) in (3.3)-(3.5), respectively, we get the following system of equations

$$\begin{bmatrix} \left(\sum_{h=1}^L Q_h W_h \bar{x}_h^2\right) & \left(\sum_{h=1}^L Q_h W_h \bar{x}_h s_{hx}^2\right) & \left(\sum_{h=1}^L Q_h W_h \bar{x}_h\right) \\ \left(\sum_{h=1}^L Q_h W_h \bar{x}_h s_{hx}^2\right) & \left(\sum_{h=1}^L Q_h W_h s_{hx}^4\right) & \left(\sum_{h=1}^L Q_h W_h s_{hx}^2\right) \\ \left(\sum_{h=1}^L Q_h W_h \bar{x}_h\right) & \left(\sum_{h=1}^L Q_h W_h s_{hx}^2\right) & \left(\sum_{h=1}^L Q_h W_h\right) \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{bmatrix} = \begin{bmatrix} \sum_{h=1}^L W_h \bar{x}_h^* - \sum_{h=1}^L W_h \bar{x}_h \\ \sum_{h=1}^L W_h s_{hx}^{*2} - \sum_{h=1}^L W_h s_{hx}^2 \\ 0 \end{bmatrix}$$

Solving the system of equations for lambdas, we obtain

$$\lambda_1 = \frac{A}{D}, \lambda_2 = \frac{B}{D}, \lambda_3 = \frac{C}{D},$$

where

$$\begin{aligned} A = & \left(\sum_{h=1}^L W_h (\bar{x}_h^* - \bar{x}_h)\right) \left[ \left(\sum_{h=1}^L Q_h W_h\right) \left(\sum_{h=1}^L Q_h W_h s_{hx}^4\right) - \left(\sum_{h=1}^L Q_h W_h s_{hx}^2\right)^2 \right] \\ & + \left(\sum_{h=1}^L W_h (s_{hx}^{*2} - s_{hx}^2)\right) \left[ \left(\sum_{h=1}^L Q_h W_h \bar{x}_h\right) \left(\sum_{h=1}^L Q_h W_h s_{hx}^2\right) \right. \\ & \left. - \left(\sum_{h=1}^L Q_h W_h\right) \left(\sum_{h=1}^L Q_h W_h s_{hx}^2 \bar{x}_h\right) \right] \end{aligned}$$

$$\begin{aligned} B = & \left(\sum_{h=1}^L W_h (s_{hx}^{*2} - s_{hx}^2)\right) \left[ \left(\sum_{h=1}^L Q_h W_h\right) \left(\sum_{h=1}^L Q_h W_h \bar{x}_h^2\right) - \left(\sum_{h=1}^L Q_h W_h \bar{x}_h\right)^2 \right] \\ & - \left(\sum_{h=1}^L W_h (\bar{x}_h^* - \bar{x}_h)\right) \left[ \left(\sum_{h=1}^L Q_h W_h s_{hx}^2 \bar{x}_h\right) \left(\sum_{h=1}^L Q_h W_h\right) \right. \\ & \left. - \left(\sum_{h=1}^L Q_h W_h \bar{x}_h\right) \left(\sum_{h=1}^L Q_h W_h s_{hx}^2\right) \right] \end{aligned}$$

$$\begin{aligned} C = & \left(\sum_{h=1}^L W_h (\bar{x}_h^* - \bar{x}_h)\right) \left[ \left(\sum_{h=1}^L Q_h W_h s_{hx}^2\right) \left(\sum_{h=1}^L Q_h W_h s_{hx}^2 \bar{x}_h\right) - \left(\sum_{h=1}^L Q_h W_h \bar{x}_h\right) \left(\sum_{h=1}^L Q_h W_h s_{hx}^4\right) \right] \\ & + \left(\sum_{h=1}^L W_h (s_{hx}^{*2} - s_{hx}^2)\right) \left[ \left(\sum_{h=1}^L Q_h W_h \bar{x}_h\right) \left(\sum_{h=1}^L Q_h W_h \bar{x}_h s_{hx}^2\right) \right. \\ & \left. - \left(\sum_{h=1}^L Q_h W_h \bar{x}_h^2\right) \left(\sum_{h=1}^L Q_h W_h s_{hx}^2\right) \right] \end{aligned}$$

$$\begin{aligned}
D = & \left( \sum_{h=1}^L Q_h W_h \right) \left( \sum_{h=1}^L Q_h W_h s_{hx}^4 \right) \left( \sum_{h=1}^L Q_h W_h \bar{x}_h^2 \right) - \left( \sum_{h=1}^L Q_h W_h \bar{x}_h \right)^2 \left( \sum_{h=1}^L Q_h W_h s_{hx}^4 \right) \\
& - \left( \sum_{h=1}^L Q_h W_h \right) \left( \sum_{h=1}^L Q_h W_h s_{hx}^2 \bar{x}_h \right)^2 - \left( \sum_{h=1}^L Q_h W_h s_{hx}^2 \right)^2 \left( \sum_{h=1}^L Q_h W_h \bar{x}_h^2 \right) \\
& + 2 \left( \sum_{h=1}^L Q_h W_h \bar{x}_h \right) \left( \sum_{h=1}^L Q_h W_h s_{hx}^2 \right) \left( \sum_{h=1}^L Q_h W_h \bar{x}_h s_{hx}^2 \right)
\end{aligned}$$

Substituting these lambdas in (3.6) and then (3.1), we get

$$\bar{y}_{st}(dnew) = \bar{y}_{st} + \beta_{1(dnew)} \left( \sum_{h=1}^L W_h (\bar{x}_h^* - \bar{x}_h) \right) + \beta_{2(dnew)} \left( \sum_{h=1}^L W_h (s_{hx}^{*2} - s_{hx}^2) \right),$$

where  $\beta_{1(dnew)} = \frac{A^*}{D}$  and  $\beta_{2(dnew)} = \frac{B^*}{D}$

$$\begin{aligned}
A^* = & \left( \sum_{h=1}^L Q_h W_h \bar{x}_h \bar{y}_h \right) \left[ \left( \sum_{h=1}^L Q_h W_h \right) \left( \sum_{h=1}^L Q_h W_h s_{hx}^4 \right) - \left( \sum_{h=1}^L Q_h W_h s_{hx}^2 \right)^2 \right] \\
& - \left( \sum_{h=1}^L Q_h W_h s_{hx}^2 \bar{y}_h \right) \left[ \left( \sum_{h=1}^L Q_h W_h s_{hx}^2 \bar{x}_h \right) \left( \sum_{h=1}^L Q_h W_h \right) - \left( \sum_{h=1}^L Q_h W_h \bar{x}_h \right) \left( \sum_{h=1}^L Q_h W_h s_{hx}^2 \right) \right] \\
& + \left( \sum_{h=1}^L Q_h W_h \bar{y}_h \right) \left[ \left( \sum_{h=1}^L Q_h W_h s_{hx}^2 \right) \left( \sum_{h=1}^L Q_h W_h s_{hx}^2 \bar{x}_h \right) - \left( \sum_{h=1}^L Q_h W_h \bar{x}_h \right) \left( \sum_{h=1}^L Q_h W_h s_{hx}^4 \right) \right] \\
B^* = & \left( \sum_{h=1}^L Q_h W_h \bar{x}_h \bar{y}_h \right) \left[ \left( \sum_{h=1}^L Q_h W_h \bar{x}_h \right) \left( \sum_{h=1}^L Q_h W_h s_{hx}^2 \right) - \left( \sum_{h=1}^L Q_h W_h \right) \left( \sum_{h=1}^L Q_h W_h s_{hx}^2 \bar{x}_h \right) \right] \\
& + \left( \sum_{h=1}^L Q_h W_h s_{hx}^2 \bar{y}_h \right) \left[ \left( \sum_{h=1}^L Q_h W_h \right) \left( \sum_{h=1}^L Q_h W_h \bar{x}_h^2 \right) - \left( \sum_{h=1}^L Q_h W_h \bar{x}_h \right)^2 \right] \\
& + \left( \sum_{h=1}^L Q_h W_h \bar{y}_h \right) \left[ \left( \sum_{h=1}^L Q_h W_h \bar{x}_h \right) \left( \sum_{h=1}^L Q_h W_h \bar{x}_h s_{hx}^2 \right) - \left( \sum_{h=1}^L Q_h W_h \bar{x}_h^2 \right) \left( \sum_{h=1}^L Q_h W_h s_{hx}^2 \right) \right]
\end{aligned}$$

#### 4. Theoretical Variance

We can write the estimators  $\bar{y}_{st}(d)$  and  $\bar{y}_{st}(dnew)$  as follows:

$$(4.1) \quad \bar{y}_{st}(\alpha) = \sum_{h=1}^L W_h \bar{y}_h + \beta_{1(\alpha)} \sum_{h=1}^L W_h (\bar{x}_h - \bar{x}_h^*) + \beta_{2(\alpha)} \sum_{h=1}^L W_h (s_{hx}^2 - s_{hx}^{*2})$$

where  $\alpha = d, dnew$ . To find the variance of estimators, let us define following equations:

$$e_{0h} = \frac{(\bar{y}_h - \bar{Y}_h)}{\bar{Y}_h}, e_{1h} = \frac{(\bar{x}_h - \bar{X}_h)}{\bar{X}_h}, e_{1h}^* = \frac{(\bar{x}_h^* - \bar{X}_h)}{\bar{X}_h}, e_{2h} = \frac{(S_{hx}^2 - S_{hx}^2)}{S_{hx}^2}$$

and  $e_{2h}^* = \frac{(S_{hx}^{*2} - S_{hx}^2)}{S_{hx}^2} \bar{y}_h = \bar{Y}_h(1 + e_{0h}), \bar{x}_h = \bar{X}_h(1 + e_{1h}), \bar{x}_h^* = \bar{X}_h(1 + e_{1h}^*),$

$$S_{hx}^2 = S_{hx}^2(1 + e_{2h}), S_{hx}^{*2} = S_{hx}^2(1 + e_{2h}^*),$$

$$E(e_{0h}^2) = \lambda_{nh}C_{yh}^2, E(e_{1h}^2) = \lambda_{nh}C_{xh}^2, E(e_{0h}e_{1h}) = \lambda_{nh}C_{yxh}, E(e_{0h}e_{1h}^*) = \lambda_{mh}C_{yxh},$$

$$E(e_{2h}^2) = \lambda_{nh}(\lambda_{04h} - 1), E(e_{2h}^2) = \lambda_{mh}(\lambda_{04h} - 1), E(e_{2h}e_{2h}^*) = \lambda_{mh}(\lambda_{04h} - 1),$$

$$E(e_{0h}e_{2h}) = \lambda_{nh}C_{yh}\lambda_{12h}, E(e_{0h}e_{2h}^*) = \lambda_{mh}C_{yh}\lambda_{12h}, E(e_{1h}^2) = \lambda_{mh}C_{xh}^2, E(e_{1h}e_{1h}^*) = \lambda_{mh}C_{xh}^2,$$

$$E(e_{1h}e_{2h}) = \lambda_{nh}C_{xh}\lambda_{03h}, E(e_{1h}^*e_{2h}) = \lambda_{mh}C_{xh}\lambda_{03h}, E(e_{1h}e_{2h}^*) = \lambda_{mh}C_{xh}\lambda_{03h}, E(e_{1h}^*e_{2h}^*) = \lambda_{mh}C_{xh}\lambda_{03h}$$

where  $\lambda_{nh} = \frac{1}{n_h} - \frac{1}{N_h}, \lambda_{mh} = \frac{1}{m_h} - \frac{1}{N_h}, C_{yh} = \frac{S_{yh}}{\bar{Y}_h}, C_{xh} = \frac{S_{xh}}{\bar{X}_h}, C_{yxh} = \frac{S_{yxh}}{\bar{Y}_h\bar{X}_h},$

$$\mu_{rsh} = \frac{\mu_{rsh}^{\frac{r}{2}} \mu_{02h}^{\frac{s}{2}}}{\mu_{20h}^{\frac{r}{2}} \mu_{02h}^{\frac{s}{2}}} \text{ and } \mu_{rsh} = \frac{\sum_{i=1}^{N_h} (Y_{hi} - \bar{Y}_h)^r (X_{hi} - \bar{X}_h)^s}{N_h - 1}.$$

Expressing (4.1) in terms of  $e$ 's, we have

(4.2)

$$\bar{y}_{st}(\alpha) = \sum_{h=1}^L W_h [\bar{Y}_h(1+e_{0h}) + \beta_{1(\alpha)}\bar{X}_h((1+e_{1h}) - (1+e_{1h}^*)) + \beta_{2(\alpha)}S_{hx}^2((1+e_{2h}) - (1+e_{2h}^*))]$$

(4.3)

$$\bar{y}_{st}(\alpha) - \sum_{h=1}^L W_h \bar{Y}_h = \sum_{h=1}^L W_h [\bar{Y}_h e_{0h} + \beta_{1(\alpha)}\bar{X}_h(e_{1h} - e_{1h}^*) + \beta_{2(\alpha)}S_{hx}^2(e_{2h} - e_{2h}^*)]$$

Squaring both sides of (4.3),

$$(4.4) \quad \left( \bar{y}_{st}(\alpha) - \sum_{h=1}^L W_h \bar{Y}_h \right)^2 = \sum_{h=1}^L W_h^2 [\bar{Y}_h e_{0h} + \beta_{1(\alpha)}\bar{X}_h(e_{1h} - e_{1h}^*) + \beta_{2(\alpha)}S_{hx}^2(e_{2h} - e_{2h}^*)]^2$$

$$= \sum_{h=1}^L W_h^2 [\bar{Y}_h^2 e_{0h}^2 + \beta_{1(\alpha)}^2 \bar{X}_h^2 (e_{1h} - e_{1h}^*)^2 + \beta_{2(\alpha)}^2 S_{hx}^4 (e_{2h} - e_{2h}^*)^2$$

$$+ 2\beta_{1(\alpha)}\bar{X}_h\bar{Y}_h e_{0h}(e_{1h} - e_{1h}^*) + 2\beta_{2(\alpha)}\bar{Y}_h S_{hx}^2 e_{0h}(e_{2h} - e_{2h}^*)$$

$$+ 2\beta_{1(\alpha)}\beta_{2(\alpha)}\bar{X}_h S_{hx}^2 (e_{1h} - e_{1h}^*)(e_{2h} - e_{2h}^*)]$$

and taking expectations, we get the variance of  $\bar{y}_{st}(\alpha)$  as

$$(4.5) \quad Var(\bar{y}_{st}(\alpha)) = \sum_{h=1}^L W_h^2 [\bar{Y}_h^2 \lambda_{nh} C_{yh}^2 + (\lambda_{nh} - \lambda_{mh}) \beta_{1(\alpha)}^2 \bar{X}_h^2 C_{xh}^2 + (\lambda_{nh} - \lambda_{mh}) \beta_{2(\alpha)}^2 S_{hx}^4 (\lambda_{04h} - 1)$$

$$+ 2(\lambda_{nh} - \lambda_{mh}) \beta_{1(\alpha)} \bar{Y}_h \bar{X}_h C_{yxh} + 2(\lambda_{nh} - \lambda_{mh}) \beta_{2(\alpha)} \bar{Y}_h S_{hx}^2 C_{yh} \lambda_{12h}$$

$$+ 2(\lambda_{nh} - \lambda_{mh}) \beta_{1(\alpha)} \beta_{2(\alpha)} S_{hx}^2 \bar{X}_h C_{xh} \lambda_{03h}]$$

The variance of  $\bar{y}_{st}(\alpha)$  in (4.5) is minimized for

$$\frac{Var(\bar{y}_{st}(\alpha))}{\partial\beta_{1(\alpha)}} = 0,$$

$$(4.6) \quad \beta_{1(\alpha)} = \frac{-\beta_{2(\alpha)}S_{hx}^2C_{xh}\lambda_{03h} - \bar{Y}_hC_{yxh}}{\bar{X}_hC_{xh}^2},$$

$$\frac{Var(\bar{y}_{st}(\alpha))}{\partial\beta_{2(\alpha)}} = 0,$$

$$(4.7) \quad \beta_{2(\alpha)} = \frac{-\beta_{1(\alpha)}\bar{X}_hC_{xh}\lambda_{03h} - \bar{Y}_hC_{yh}\lambda_{12h}}{S_{hx}^2(\lambda_{04h} - 1)},$$

Substituting (4.6) in (4.7) or vice versa, we have optimum betas as given by

$$\beta_{1(\alpha)} = \frac{S_{yh}}{S_{xh}} \frac{\lambda_{12h}\lambda_{03h} - \lambda_{11h}(\lambda_{04h} - 1)}{(\lambda_{04h} - 1) - \lambda_{03h}^2}, \beta_{2(\alpha)} = \frac{S_{yh}}{S_{xh}^2} \frac{\lambda_{11h}\lambda_{03h} - \lambda_{12h}}{(\lambda_{04h} - 1) - \lambda_{03h}^2}$$

The resulting (minimum) variance of  $\bar{y}_{st}(\alpha)$  is given by

$$(4.8) \quad Var(\bar{y}_{st}(\alpha))$$

$$= \sum_{h=1}^L W_h^2 \bar{Y}_h^2 \left[ \lambda_{nh}C_{yh}^2 - (\lambda_{nh} - \lambda_{mh}) \frac{C_{yxh}^2(\lambda_{04h} - 1) + C_{xh}^2C_{yh}^2\lambda_{12h}^2 - 2C_{yxh}C_{yh}C_{xh}\lambda_{03h}\lambda_{12h}}{C_{xh}^2[(\lambda_{04h} - 1) - \lambda_{03h}^2]} \right]$$

$$= \sum_{h=1}^L W_h^2 \bar{Y}_h^2 C_{yh}^2 \left[ \lambda_{mh} + (\lambda_{nh} - \lambda_{mh}) \left[ 1 - \lambda_{11h}^2 - \frac{(\lambda_{12h} - \lambda_{11h}\lambda_{03h})^2}{(\lambda_{04h} - 1) - \lambda_{03h}^2} \right] \right]$$

## 5. Simulation Study

To study the properties of the proposed calibration estimator, we perform a simulation study by generating four different artificial populations where  $\bar{x}_{hi}^*$  and  $\bar{y}_{hi}^*$  values are from different distributions given in Table 1. To get different level of correlations between study and auxiliary variables, we apply some transformations given in Table 2. Each population consists of three strata having 500 units. After selecting a preliminary sample of size 300 from each stratum, we select 5000 times for the second sample whose sample of sizes are 30 and 50. The correlation coefficients between study and auxiliary variables for each stratum are taken as  $\rho_{xy1} = 0.5$ ,  $\rho_{xy2} = 0.7$  and  $\rho_{xy3} = 0.9$ . The quantities,  $S_{1x} = 4.5$ ,  $S_{2x} = 6.2$ ,  $S_{3x} = 8.4$  and  $S_{1y} = S_{2y} = S_{3y}$  are taken as fixed in each stratum as in Tracy et al. [9]. We calculate the empirical mean square error and percent relative efficiency, respectively, using following formulas:



$$MSE(\bar{y}_{st}(\alpha)) = \frac{\binom{N}{n} \sum_{k=1}^n [\bar{y}_{st}(\alpha) - \bar{Y}]^2}{\binom{N}{n}}, \alpha = d, dnew$$

$$PRE = \frac{MSE(\bar{y}_{st}(d))}{MSE(\bar{y}_{st}(dnew))} * 100$$

From Table 3, the simulation study shows that new calibration estimator is quite efficient than the existing estimator.

## 6. Conclusion

In this study we derived new calibration weights in stratified double sampling. The performance of the weights are compared with a simulation study. We found that suggested weights perform better than existing weights.

## References

- [1] Arnab, R. and Singh, S. *A note on variance estimation for the generalized regression predictor*, Australian and New Zealand Journal of Statistics **47** (2), 231–234, 2005.
- [2] Deville, J. C. and Sarndal, C. E. *Calibration estimators in survey sampling*, Journal of the American Statistical Association **87**, 376–382, 1992.
- [3] Estevao, V. M. and Sarndal, C. E. *A functional form approach to calibration*, Journal of Official Statistics **16**, 379–399, 2000.
- [4] Farrell, P. J. and Singh, S. *Model-assisted higher order calibration of estimators of variance*, Australian and New Zealand Journal of Statistics **47** (3), 375–383, 2005.
- [5] Kim, J. M., Sungur, E. A. and Heo T. Y. *Calibration approach estimators in stratified sampling*, Statistics and Probability Letters **77** (1), 99–103, 2007.
- [6] Kim, J. K. and Park, M. *Calibration estimation in survey sampling*, International Statistical Review **78** (1), 21–29, 2010.
- [7] Koyuncu, N. *Application of calibration method to estimators in sampling theory*, (PhD. Thesis, Hacettepe University Department of Statistics, 2012).
- [8] Koyuncu, N. and Kadilar, C. *Calibration estimator using different distance measures in stratified random sampling*, International Journal of Modern Engineering Research **3** (1), 415–419, 2013.
- [9] Tracy, D. S., Singh, S. and Arnab, R. *Note on calibration in stratified and double sampling*, Survey Methodology **29**, 99–104, 2003.

**Table 1.** Parameters and Distributions of Study and Auxiliary Variables

Parameters and distributions of study variable	Parameters and distributions of auxiliary variable
	I. Population, h=1,2,3
$f(y_{hi}^*) = \frac{1}{\Gamma(1.5)} y_{hi}^{*1.5-1} e^{-y_{hi}^*}$	$f(x_{hi}^*) = \frac{1}{\Gamma(0.3)} x_{hi}^{*0.3-1} e^{-x_{hi}^*}$
	II. Population, h=1,2,3
$f(y_{hi}^*) = \frac{1}{\sqrt{2\pi}} e^{-\frac{y_{hi}^{*2}}{2}}$	$f(x_{hi}^*) = \frac{1}{\Gamma(0.3)} x_{hi}^{*0.3-1} e^{-x_{hi}^*}$

**Table 2.** Properties of Strata

Strata	Study Variable	Auxiliary Variable
<b>1. Stratum</b>	$y_{1i} = 50 + y_{1i}^*$	$x_{1i} = 15 + \sqrt{(1 - \rho_{xy1}^2)} x_{1i}^* + \rho_{xy1} \frac{S_{1x}}{S_{1y}} y_{1i}^*$
<b>2. Stratum</b>	$y_{2i} = 150 + y_{2i}^*$	$x_{2i} = 100 + \sqrt{(1 - \rho_{xy2}^2)} x_{2i}^* + \rho_{xy2} \frac{S_{2x}}{S_{2y}} y_{2i}^*$
<b>3. Stratum</b>	$y_{3i} = 50 + y_{3i}^*$	$x_{3i} = 200 + \sqrt{(1 - \rho_{xy3}^2)} x_{3i}^* + \rho_{xy3} \frac{S_{3x}}{S_{3y}} y_{3i}^*$

**Table 3.** Empirical Mean Square Error (MSE) and Percent Relative Efficiency (PRE) of Estimators

Population	Empirical MSE( $\bar{y}_{st}(d)$ )	Empirical MSE( $\bar{y}_{st}(dnew)$ )	PRE
<b>I</b> ( $m_h = 30$ )	<b>61205495678</b>	<b>65730798</b>	<b>93115.4</b>
<b>I</b> ( $m_h = 50$ )	<b>626914412</b>	<b>38472456</b>	<b>1629.515</b>
<b>II</b> ( $m_h = 30$ )	<b>6.8404e+11</b>	<b>50343013</b>	<b>1358759</b>
<b>II</b> ( $m_h = 50$ )	<b>245901177</b>	<b>35046173</b>	<b>701.6491</b>