مجلهٔ پژوهش‌های آماری ایران
پژوهشکدهٔ آمار دورهٔ ۹، شمارهٔ ۲، پاییز و زمستان ۱۳۹۱، صص ۱۵۹–۱۷۷

J. Statist. Res. Iran **9** (2012): 159–177

Some Improvment in the Estimation of Population Mean in Cluster Sampling

M. Sakizadeh[†,*] and A. Gerami[‡]

† West Tehran Branch, Islamic Azad University
‡ Qazvin Branch, Islamic Azad University

Abstract. Gupta and Shabbir (2008) have suggested an alternative form of ratio-type estimator for estimating the population mean. In this paper, we introduced new estimators by mixing two, stratified and cluster sampling method. Then we improved these estimators by using auxiliary variables and introducing new estimators. For sampling in infinite populations with a high geographic dispersion, the population will be divided into some smaller sub-population which leads to dispersion reduction to some extent. This will affect the variance of the estimator. Additionally dividing the population will result in saving cost, time and eases calculations.

Keywords. Auxiliary variables; cluster sampling; estimator; stratified sampling.

MSC 2010: 62D05.

1 Introduction

A ratio estimator is commonly used when the study variable Y is highly correlated with the auxiliary variable X. When the population mean $\bar{X}$ is known, a number of modified versions of ratio estimators have been suggested by various authors. Further, many authors used some population parameters of the auxiliary variable to improve the precision of ratio estimators such as Sisodia and Dwivedi (1981), Upadhyaya and Singh (1999), Singh et al.

* Corresponding author

(2008) and others. Gupta and Shabbir (2008) have suggested a general class of ratio estimators when the population parameter of the auxiliary variable is known. In addition to these studies, Kadilara and Cingi (2003), Shabbir and Gupta (2005), Singh and Vishwakarma (2006-2008), Koyuncu and Kadilara (2009) extended the suggested estimators in a simple random sampling to stratified random sampling. In this study, we derive the correct expression of the MSE in Gupta and Shabbir (2008) and with an overall view of the Koyuncua and Kadilara (2010) we suggest similar estimators in the cluster sampling and extend these estimator, for large population. Then we consider a finite population of size N clusters from which a sample of size n cluster is drawn according to the simple random sampling without replacement.

In this paper a general regression-ratio type estimator of mean population is considered which generates a large class of estimator in cluster sampling. It has been shown that the suggested estimator is more efficient than the regression estimator. In Section 4, two new estimators are suggested in large population which divided into some smaller sub-population and are compared by the mean square of error. Consequently numerical examples are given in Section 5.

Let y_{ij} and x_{ij} be the values of the study and the auxiliary variables respectively on the ith cluster ($i = 1, 2, \ldots, N$) and the jth unit ($j = 1, 2, \ldots, M$). We use the following notation in this paper.

$$y_i = \sum_{j=1}^{M} y_{ij} \quad \bar{Y} = \frac{1}{N}\sum_{i=1}^{N} y_i \quad \bar{\bar{Y}} = \frac{1}{M}\bar{Y} \quad \bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i \quad \bar{\bar{y}} = \frac{1}{M}\bar{y}$$

$$x_i = \sum_{j=1}^{M} x_{ij} \quad \bar{X} = \frac{1}{N}\sum_{i=1}^{N} x_i \quad \bar{\bar{X}} = \frac{1}{M}\bar{X} \quad \bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \quad \bar{\bar{x}} = \frac{1}{M}\bar{x}$$

$$S_Y^2 = \frac{1}{MN-1}\sum_{i=1}^{N}\sum_{j=1}^{M}(y_{ij} - \bar{\bar{Y}})^2 \quad S_X^2 = \frac{1}{MN-1}\sum_{i=1}^{N}\sum_{j=1}^{M}(x_{ij} - \bar{\bar{X}})^2$$

$$S_{XY} = \frac{1}{MN-1}\sum_{i=1}^{N}\sum_{j=1}^{M}(y_{ij} - \bar{\bar{Y}})(x_{ij} - \bar{\bar{X}}) \qquad f = \frac{n}{N}$$

$$\gamma = \frac{1-f}{n} \cdot \frac{(NM-1)}{M^2(N-1)}$$

$$\rho_Y = \frac{\sum_{i=1}^{N}\sum_{j=1}^{M}\sum_{k\neq j}^{M}(y_{ij}-\bar{\bar{Y}})(y_{ik}-\bar{\bar{Y}})}{(M-1)(NM-1)S_Y^2}$$

$$\rho_X = \frac{\sum_{i=1}^{N}\sum_{j=1}^{M}\sum_{k\neq j}^{M}(x_{ij}-\bar{\bar{X}})(x_{ik}-\bar{\bar{X}})}{(M-1)(NM-1)S_X^2}$$

$$\rho_{XY} = \frac{\sum_{i=1}^{N}\sum_{j=1}^{M}(y_{ij}-\bar{\bar{Y}})(x_{ij}-\bar{\bar{X}})}{(NM-1)S_X S_Y}$$

$$\rho_{XY}^{'} = \frac{\sum_{i=1}^{N}\sum_{j=1}^{M}\sum_{k\neq j}^{M}(y_{ij}-\bar{\bar{Y}})(x_{ik}-\bar{\bar{X}})}{(M-1)(NM-1)S_X S_Y}$$

2 Estimators in the Cluster Sampling

For estimating the population mean $\bar{\bar{Y}}$, a regression type estimator of Gupta and Shabbir (2008) can be given as the following:

$$\bar{\bar{y}}_{mc(0)} = \frac{1}{M}\left\{\omega_1\bar{y} + \omega_2(\bar{X}-\bar{x})\right\} = \omega_1\bar{\bar{y}} + \omega_2(\bar{\bar{X}}-\bar{\bar{x}}), \tag{1}$$

when ω_1 and ω_2 are constants that have no restriction. The Bias and MSE of $\bar{\bar{y}}_{mc(0)}$ are respectively given by:

$$\begin{aligned} B(\bar{\bar{y}}_{mc(0)}) &= E(\bar{\bar{y}}_{mc(0)}) - \bar{\bar{Y}} = \bar{\bar{Y}}(\omega_1 - 1), \\ \text{MSE}(\bar{\bar{y}}_{mc(0)}) &= \text{Var}(\bar{\bar{y}}_{mc(0)}) + \left\{B(\bar{\bar{y}}_{mc(0)})\right\}^2 \\ &= \omega_1^2\text{Var}(\bar{\bar{y}}) + \omega_2^2\text{Var}(\bar{\bar{x}}) - 2\omega_1\omega_2\text{Cov}(\bar{\bar{y}},\bar{\bar{x}}) + \bar{\bar{Y}}^2(\omega_1-1)^2, \end{aligned} \tag{2}$$

where, the $V(\bar{\bar{y}})$, $V(\bar{\bar{x}})$ and $\text{Cov}(\bar{\bar{y}}, \bar{\bar{x}})$ in equation (2) are respectively given by:

$$\begin{aligned}
\text{Var}(\bar{\bar{y}}) &= \frac{1-f}{n(N-1)M^2}\sum_{i=1}^{N}(y_i - \bar{Y})^2 \\
&= \frac{1-f}{n(N-1)M^2}\sum_{i=1}^{N}\left\{(y_{i1} - \bar{\bar{Y}}) + (y_{i2} - \bar{\bar{Y}}) + \ldots + (y_{iM} - \bar{\bar{Y}})\right\}^2 \\
&= \frac{1-f}{n(N-1)M^2}\left\{\sum_{i=1}^{N}\sum_{j=1}^{M}(y_{ij} - \bar{\bar{Y}})^2 + \sum_{i=1}^{N}\sum_{j=1}^{M}\sum_{k\neq j}^{M}(y_{ij} - \bar{\bar{Y}})(y_{ik} - \bar{\bar{Y}})\right\} \\
&= \frac{1-f}{n(N-1)M^2}\left\{(NM-1)S_Y^2 + (M-1)(NM-1)\rho_Y S_Y^2\right\}.
\end{aligned}$$

So,

$$\text{Var}(\bar{\bar{y}}) = \gamma S_Y^2\{1 + (M-1)\rho_Y\}, \qquad \text{Var}(\bar{\bar{x}}) = \gamma S_X^2\{1 + (M-1)\rho_X\}.$$

Also,

$$\begin{aligned}
\text{Cov}(\bar{\bar{x}}, \bar{\bar{y}}) &= \frac{1-f}{n(N-1)M^2}\sum_{i=1}^{N}(x_i - \bar{X})(y_i - \bar{Y}) \\
&= \frac{1-f}{n(N-1)M^2}\sum_{i=1}^{N}\Bigg[\left\{(x_{i1} - \bar{\bar{X}}) + \cdots + (x_{iM} - \bar{\bar{X}})\right\} \\
&\quad \times \left\{(y_{i1} - \bar{\bar{Y}}) + \ldots + (y_{iM} - \bar{\bar{Y}})\right\}\Bigg] \\
&= \frac{1-f}{n(N-1)M^2} \\
&\quad \times \left\{\sum_{i=1}^{N}\sum_{j=1}^{M}(x_{ij} - \bar{\bar{X}})(y_{ij} - \bar{\bar{Y}}) + \sum_{i=1}^{N}\sum_{j=1}^{M}\sum_{k\neq j}^{M}\left(y_{ij} - \bar{\bar{Y}}\right)\left(x_{ik} - \bar{\bar{X}}\right)\right\} \\
&= \frac{1-f}{n(N-1)M^2}\left\{(NM-1)S_{XY} + (M-1)(NM-1)\rho'_{XY}S_X S_Y\right\}.
\end{aligned}$$

So,

$$\text{Cov}(\bar{\bar{x}}, \bar{\bar{y}}) = \gamma S_X S_Y\left\{\rho_{XY} + (M-1)\rho'_{XY}\right\}.$$

Now by considering $\mathrm{Cov}(\bar{\bar{x}}, \bar{\bar{y}})$, $\mathrm{Var}(\bar{\bar{x}})$ and $\mathrm{Var}(\bar{\bar{y}})$, the value of $\mathrm{MSE}(\bar{\bar{y}}_{mc(0)})$ will be as follow:

$$\mathrm{MSE}(\bar{\bar{y}}_{mc(0)}) = \gamma\Big[\omega_1^2 S_Y^2\{1+(M-1)\rho_Y\} + \omega_2^2 S_X^2\{1+(M-1)\rho_X\} - 2\omega_1\omega_2 S_X S_Y\{\rho_{XY}+(M-1)\rho'_{XY}\}\Big] + \bar{\bar{Y}}^2(\omega_1-1)^2. \quad (3)$$

In ordet to minimizing $\mathrm{MSE}(\bar{\bar{y}}_{mc(0)})$, we use the following notations:

$$\nu_y = \frac{\bar{\bar{y}} - \bar{\bar{Y}}}{\bar{\bar{Y}}}, \qquad \nu_x = \frac{\bar{\bar{x}} - \bar{\bar{X}}}{\bar{\bar{X}}},$$

$$E(\nu_y^2) = u_y^2, \qquad E(\nu_x^2) = u_x^2, \qquad E(\nu_x\nu_y) = u_{xy}.$$

It is obvious that $E(\nu_y) = 0,\ E(\nu_x) = 0$.

So, from (2) $\mathrm{MSE}(\bar{\bar{y}}_{mc(0)})$ is given by:

$$\mathrm{MSE}(\bar{\bar{y}}_{mc(0)}) = \omega_1^2\bar{\bar{Y}}^2u_y^2 + \omega_2^2\bar{\bar{X}}^2u_x^2 - 2\omega_1\omega_2\bar{\bar{Y}}\bar{\bar{X}}u_{xy} + \bar{\bar{Y}}^2(\omega_1-1)^2. \quad (4)$$

The value of ω_1 and ω_2 which minimize (4) are given as follow:

$$\omega_1^* = \frac{u_x^2}{u_x^2u_y^2 + u_x^2 - u_{xy}^2}, \qquad \omega_2^* = \frac{\bar{\bar{Y}}}{\bar{\bar{X}}}\left(\frac{u_{xy}}{u_x^2u_y^2 + u_x^2 - u_{xy}^2}\right).$$

Substituting these optimum values in equation (4), the minimum MSE of $\bar{\bar{y}}_{mc(0)}$ is given by

$$\mathrm{MSE}(\bar{\bar{y}}_{mc(0)})_{min} = \bar{\bar{Y}}^2\frac{\left(u_y^2 - \dfrac{u_{yx}^2}{u_x^2}\right)}{\left(u_y^2 - \dfrac{u_{yx}^2}{u_x^2}\right)+1}. \quad (5)$$

When ω_1 and ω_2 are any constants and λ and η are either constants or functions of known parameters, the ratio type estimator suggested by Gupta and Shabbir (2008) is given by

$$\bar{\bar{y}}_{mc} = \left\{\omega_1\bar{\bar{y}} + \omega_2(\bar{\bar{X}} - \bar{\bar{x}})\right\}\left(\frac{\lambda + \eta\bar{\bar{X}}}{\lambda + \eta\bar{\bar{x}}}\right) \quad (6)$$

It is worth mentioning that when values of ω_1, ω_2, η and λ are conveniently chosen, many common estimators can be obtained such as the classical ratio estimator $\bar{y}_0$, the regression type estimator $\bar{\bar{y}}_{mc(0)}$, the estimators suggested by Singh and Tailor (2003), Sisodia and Dwivedi (1981), Upadhyaya and Singh (1999) etc. In addition to these estimators, some new estimators, which are also generated from equation (6), are given in Table 1.

Expressing equation (6) in terms of ν_x , ν_y we have

$$\bar{\bar{y}}_{mc} = \left\{\omega_1 \bar{\bar{Y}}(1+\nu_y) - \omega_2 \bar{\bar{X}}\nu_x\right\}(\alpha\nu_x + 1)^{-1} \tag{7}$$

where $\alpha = \dfrac{\eta\bar{\bar{X}}}{\lambda + \eta\bar{\bar{X}}}$.

By expanding the terms of Equation (7) up to the first order of approximation we have:

$$\begin{aligned} B(\bar{\bar{y}}_{mc}) &\approx \omega_1\bar{\bar{Y}}(1 - \alpha u_{xy} + \alpha^2 u_x^2) + \omega_2\bar{\bar{X}}(\alpha u_x^2) - \bar{\bar{Y}} \\ \text{MSE}(\bar{\bar{y}}_{mc}) &\approx (\omega_1 - 1)^2\bar{\bar{Y}}^2 + \omega_1^2\bar{\bar{Y}}^2(u_y^2 - 4\alpha u_{xy} + 3\alpha^2 u_x^2) \\ &\quad - 2\omega_1\bar{\bar{Y}}^2(\alpha^2 u_x^2 - \alpha u_{xy}) + \omega_2^2\bar{\bar{X}}^2 u_x^2 - 2\alpha\omega_2\bar{\bar{X}}\bar{\bar{Y}}u_x^2 \\ &\quad - 2\omega_1\omega_2\bar{\bar{X}}\bar{\bar{Y}}(u_{xy} - 2\alpha u_x^2). \end{aligned} \tag{8}$$

The value for ω_1 and ω_2 which minimizing $\text{MSE}(\bar{\bar{y}}_{mc})$ are given as follow:

$$\omega_1^* \approx \frac{1 - \alpha^2 u_x^2}{1 + u_y^2 - \dfrac{u_{xy}^2}{u_x^2} - \alpha^2 u_x^2}, \qquad \omega_2^* \approx \frac{\bar{\bar{Y}}}{\bar{\bar{X}}}\left\{\frac{(u_{xy} - 2\alpha u_x^2)(1 - \alpha^2 u_x^2)}{u_x^2 + u_x^2 u_y^2 - u_{xy}^2 - \alpha^2 u_x^4)} + \alpha\right\}. \tag{9}$$

On substituting those values for the ω_1^* and ω_2^* in equation (8), by accepting condition $\left(u_x^2 + u_x^2 u_y^2 - u_{yx}^2 - \alpha^2 u_x^4\right) > 0$, the minimum MSE of $\bar{\bar{y}}_{mc}$ can be written as follow:

$$\text{MSE}(\bar{\bar{y}}_{mc})_{min} \approx \bar{\bar{Y}}^2 \frac{\left(u_y^2 - \dfrac{u_{yx}^2}{u_x^2}\right)(1 - \alpha^2 u_x^2)}{\left(u_y^2 - \dfrac{u_{yx}^2}{u_x^2}\right) + 1 - \alpha^2 u_x^2}. \tag{10}$$

Before comparison the efficiencies of the estimators, it is necessary to define ratio and regression estimator in cluster sampling as follow:

$$\bar{\bar{y}}_{reg} = \bar{\bar{y}} + B(\bar{\bar{X}} - \bar{\bar{x}}) \qquad \bar{\bar{y}}_0 = \frac{\bar{\bar{y}}}{\bar{\bar{x}}}\bar{\bar{X}} \qquad B = \frac{S_{XY}}{S_X^2}$$

$$\bar{\bar{y}}_{reg(mcs)} = \sum_{h=1}^{L} \left\{ \bar{\bar{y}}_h + B_h(\bar{\bar{X}}_h - \bar{\bar{x}}_h) \right\} \qquad \bar{\bar{y}}_{0(acs)} = \sum_{h=1}^{L} \frac{\bar{\bar{y}}_h}{\bar{\bar{x}}_h} \bar{\bar{X}}_h \qquad B_h = \frac{S_{XYh}}{S_{Xh}^2}$$

$$\bar{\bar{y}}_{reg(acs)} = \bar{\bar{y}}_{st} + B_{st}(\bar{\bar{X}}_{st} - \bar{\bar{x}}) \qquad \bar{\bar{y}}_{0(mcs)} = \frac{\bar{\bar{y}}_{st}}{\bar{\bar{x}}_{st}} \bar{\bar{X}} \qquad B_{st} = \sum_{h=1}^{L} \frac{S_{XYh}}{S_{Xh}^2}$$

where $\bar{\bar{y}}_{mcs}$ and $\bar{\bar{y}}_{acs}$ stands for seprate and combined estimators, respectively. In the following section, we compare the efficiencies of the estimators.

3 Efficiency Comparisons in the Cluster Sampling

$\bar{\bar{y}}_{mc}$ is more efficient than $\bar{\bar{y}}_{mc(0)}$ if and only if

$$\text{MSE}(\bar{\bar{y}}_{mc(0)})_{min} - \text{MSE}(\bar{\bar{y}}_{mc})_{min} > 0$$

or

$$\bar{\bar{Y}}^2 \frac{u_y^2 - \dfrac{u_{yx}^2}{u_x^2}}{u_y^2 - \dfrac{u_{yx}^2}{u_x^2} + 1} - \bar{\bar{Y}}^2 \frac{\left(u_y^2 - \dfrac{u_{yx}^2}{u_x^2}\right)(1 - \alpha^2 u_x^2)}{\left(u_y^2 - \dfrac{u_{yx}^2}{u_x^2}\right) + 1 - \alpha^2 u_x^2} > 0. \tag{11}$$

Because of condition $\left(u_y^2 - \dfrac{u_{yx}^2}{u_x^2} + 1 - \alpha^2 u_x^2\right) > 0$, the inequality (11) reduces to

$$u_x^2 \left(u_y^2 - \frac{u_{yx}^2}{u_x^2}\right) > 0. \tag{12}$$

If the condition (12) is satisfied, $\bar{\bar{y}}_{mc}$ is more efficient than $\bar{\bar{y}}_{mc(0)}$. Note that condition (12) is always satisfied. Consequently we can say that this estimator is always more efficient than $\bar{\bar{y}}_{mc(0)}$. Also $\bar{\bar{y}}_{mc}$ is more efficient than $\bar{\bar{y}}_{reg}$ if and only if

$$\text{MSE}(\bar{\bar{y}}_{reg}) - \text{MSE}(\bar{\bar{y}}_{mc})_{min} > 0$$

or

$$\bar{\bar{Y}}^2 u_y^2 \left(1 - \frac{u_{yx}^2}{u_x^2 u_y^2}\right) - \bar{\bar{Y}}^2 \frac{\left(u_y^2 - \dfrac{u_{yx}^2}{u_x^2}\right)(1 - \alpha^2 u_x^2)}{\left(u_y^2 - \dfrac{u_{yx}^2}{u_x^2}\right) + 1 - \alpha^2 u_x^2} > 0. \tag{13}$$

Because of condition $\left(u_y^2 - \frac{u_{yx}^2}{u_x^2} + 1 - \alpha^2 u_x^2\right) > 0$, the inequality (13) reduces to

$$\left(u_y^2 - \frac{u_{yx}^2}{u_x^2}\right) > 0. \tag{14}$$

If the condition (14) is satisfied, $\bar{\bar{y}}_{mc}$ is more efficient than $\bar{\bar{y}}_{reg}$. Note that condition (14) is always satisfied for $\bar{\bar{y}}_{mc}$. So we can say that $\bar{\bar{y}}_{mc}$ is always more efficient than $\bar{\bar{y}}_{reg}$.

4 Suggested Estimators in Cluster Sampling in Large Populations

4.1 Notations

In order to have less calculations, at first we should divide the population into L sub-populations. For example in country, we devide the country to some smaller sections with less dispersion. For instance a sub-population consists of some deprived states and some rich states then each sub-society is clustered. In order to ease, we choose the same size for each cluster. So that the hth sub-population contains N_h clusters that each cluster has M members. As a result the specified population consist of N clusters.

$$N = \sum_{h=1}^{L} N_h, \quad W_h = \frac{N_h}{\sum_{i=1}^{L} N_h} \qquad h = 1, 2, \ldots, L$$

Now we choose a simple random sample with n_h cluster size without replacment from hth sub-population. So, $n = \sum_{h=1}^{L} n_h$.

Also let y_{ijh}, x_{ijh} denote the observed value respectively for Y, X of the hth sub-population, for the ith cluster and jth unit, $i = 1, 2, \ldots, N_h,\ j = 1, 2, \ldots, M,\ h = 1, 2, \ldots, L$.

Also we define the following notations:

$$y_{ih} = \sum_{j=1}^{M} y_{ijh} \qquad \bar{Y}_h = \frac{1}{N_h}\sum_{i=1}^{N_h} y_{ih} \qquad \bar{\bar{Y}}_h = \frac{1}{M}\bar{Y}_h$$

$$\bar{y}_h = \frac{1}{n_h}\sum_{i=1}^{n_h} y_{ih} \qquad \bar{\bar{y}}_h = \frac{1}{M}\bar{y}_h \qquad \bar{\bar{y}}_{st} = \sum_{h=1}^{L} W_h \bar{\bar{y}}_h$$

$$x_{ih} = \sum_{j=1}^{M} x_{ijh} \qquad \bar{X}_h = \frac{1}{N_h}\sum_{i=1}^{N_h} x_{ih} \qquad \bar{\bar{X}}_h = \frac{1}{M}\bar{X}_h$$

$$\bar{x}_h = \frac{1}{n_h}\sum_{i=1}^{n_h} x_{ih} \qquad \bar{\bar{x}}_h = \frac{1}{M}\bar{x}_h \qquad \bar{\bar{x}}_{st} = \sum_{h=1}^{L} W_h \bar{\bar{x}}_h$$

$$S^2_{XYh} = \frac{1}{MN_h - 1}\sum_{i=1}^{N_h}\sum_{j=1}^{M}(y_{ijh} - \bar{\bar{Y}}_h)(x_{ijh} - \bar{\bar{X}}_h)$$

$$S^2_{Yh} = \frac{1}{MN_h - 1}\sum_{i=1}^{N_h}\sum_{j=1}^{M}(y_{ijh} - \bar{\bar{Y}}_h)^2$$

$$\rho_{Yh} = \frac{\sum_{i=1}^{N_h}\sum_{j=1}^{M}\sum_{k\neq j}^{M}(y_{ijh} - \bar{\bar{Y}}_h)(y_{ikh} - \bar{\bar{Y}}_h)}{(M-1)(N_hM-1)S^2_{Yh}}$$

$$S^2_{Xh} = \frac{1}{MN_h - 1}\sum_{i=1}^{N_h}\sum_{j=1}^{M}(x_{ijh} - \bar{\bar{X}}_h)^2$$

$$\rho_{Xh} = \frac{\sum_{i=1}^{N_h}\sum_{j=1}^{M}\sum_{k\neq j}^{M}(x_{ijh} - \bar{\bar{X}}_h)(x_{ikh} - \bar{\bar{X}}_h)}{(M-1)(N_hM-1)S^2_{Xh}}$$

$$\rho_{XYh} = \frac{\sum_{i=1}^{N_h}\sum_{j=1}^{M}(y_{ijh} - \bar{\bar{Y}}_h)(x_{ijh} - \bar{\bar{X}}_h)}{(N_hM-1)S_{Xh}S_{Yh}}$$

$$\rho'_{XYh} = \frac{\sum_{i=1}^{N_h}\sum_{j=1}^{M}\sum_{k\neq j}^{M}(y_{ijh} - \bar{\bar{Y}}_h)(x_{ikh} - \bar{\bar{X}}_h)}{(M-1)(N_hM-1)S_{Xh}S_{Yh}}$$

4.2 Suggested Separate Estimators in Cluster Sampling in Large Populations

To estimate the population mean $\bar{\bar{Y}}$, a separate type of the regression-ratio estimator can be given in the following when constants ω_{1h} and ω_{2h} have no restrictions:

$$\bar{\bar{y}}_{mcs} = \sum_{h=1}^{L} W_h \left\{ \omega_{1h}\bar{\bar{y}}_h + \omega_{2h}(\bar{\bar{X}}_h - \bar{\bar{x}}_h) \right\} \left(\frac{\lambda_h + \eta_h \bar{\bar{X}}_h}{\lambda_h + \eta_h \bar{\bar{x}}_h} \right), \tag{15}$$

where η_h and λ_h are either real numbers or the functions of the known parameters of the auxiliary variable for the hth stratum. To obtain the bias and the MSE, let us define $\nu_{yh} = \dfrac{\bar{\bar{y}}_h - \bar{\bar{Y}}_h}{\bar{\bar{Y}}_h}$ and $\nu_{xh} = \dfrac{\bar{\bar{x}}_h - \bar{\bar{X}}_h}{\bar{\bar{X}}_h}$. Then utilizing

$$E(\nu_{yh}) = 0, \quad E(\nu_{xh}) = 0, \quad E(\nu_{yh}^2)u_{yh}^2, \quad E(\nu_{xh}^2) = u_{xh}^2, \quad E(\nu_{xh}\nu_{yh}) = u_{xyh}.$$

By applying the standard techniques mentioned in the previous section, bias and MSE are, respectively, as follow:

$$B(\bar{\bar{y}}_{mcs}) \approx \sum_{h=1}^{L} W_h \left\{ \omega_{1h}\bar{\bar{Y}}_h(1 - \alpha_h u_{xyh} + \alpha_h^2 u_{xh}^2) + \omega_{2h}\bar{\bar{X}}_h(\alpha_h u_{xh}^2) \right\} - \bar{\bar{Y}},$$

$$\begin{aligned} \text{MSE}(\bar{\bar{y}}_{mcs}) \approx \sum_{h=1}^{L} W_h^2 \Big\{ & (\omega_{1h} - 1)^2 \bar{\bar{Y}}_h^2 + \omega_{1h}^2 \bar{\bar{Y}}_h^2 (u_{yh}^2 - 4\alpha_h u_{xyh} + 3\alpha_h^2 u_{xh}^2) \\ & - 2\alpha_h \omega_{2h} \bar{\bar{X}}_h \bar{\bar{Y}}_h u_{xh}^2 - 2\omega_{1h} \bar{\bar{Y}}_h^2 (\alpha_h^2 u_{xh}^2 - \alpha_h u_{xyh}) \\ & + \omega_{2h}^2 \bar{\bar{X}}_h^2 u_{xh}^2 - 2\omega_{1h}\omega_{2h} \bar{\bar{X}}_h \bar{\bar{Y}}_h (u_{xyh} - 2\alpha_h u_{xh}^2) \Big\}, \end{aligned} \tag{16}$$

where $\alpha_h = \dfrac{\eta_h \bar{\bar{X}}_h}{\lambda_h + \eta_h \bar{\bar{X}}_h}$.
It can be shown that optimum values of ω_{1h} and ω_{2h} which minimize $\text{MSE}(\bar{\bar{y}}_{mcs})$

are.

$$\omega_{1h}^{*} \approx \frac{1-\alpha_h^2 u_{xh}^2}{1+u_{yh}-\dfrac{u_{yxh}^2}{u_{xh}^2}-\alpha_h^2 u_{xh}^2}$$

$$\omega_{2h}^{*} \approx \frac{\bar{\bar{Y}}_h}{\bar{\bar{X}}_h}\left\{\frac{(u_{xyh}-2\alpha_h u_{xh}^2)(1-\alpha_h^2 u_{xh}^2)}{u_{xh}^2+u_{xh}^2 u_{yh}^2-u_{yxh}^2-\alpha_h^2 u_{xh}^4)}+\alpha_h\right\}. \tag{17}$$

On substituting those values for the ω_{1h}^{*} and ω_{2h}^{*} in equation (16), by accepting condition
$\left(u_{xh}^2+u_{xh}^2u_{yh}^2-u_{yxh}^2-\alpha_h^2u_{xh}^4\right) > 0$, the minimum MSE of $\bar{\bar{y}}_{mcs}$ can be written as follows

$$\text{MSE}(\bar{\bar{y}}_{mcs})_{min} \approx \sum_{h=1}^{L} W_h^2 \bar{\bar{Y}}_h^2 \frac{\left(u_{yh}^2-\dfrac{u_{yxh}^2}{u_{xh}^2}\right)(1-\alpha_h^2u_{xh}^2)}{\left(u_{yh}^2-\dfrac{u_{yxh}^2}{u_{xh}^2}\right)+(1-\alpha_h^2u_{xh}^2)}. \tag{18}$$

4.2.1 Efficiency Comparisons for Separate Estimators

In this section, we compare the efficiencies of the seprate estimators as follows:

$$\text{MSE}(\bar{\bar{y}}_{mcs(0)})_{min} - \text{MSE}(\bar{\bar{y}}_{mcs})_{min} > 0$$

$$\iff \sum_{h=1}^{L} W_h^2 \bar{\bar{Y}}_h^2 \left\{\frac{\left(u_{yh}^2-\dfrac{u_{yxh}^2}{u_{xh}^2}\right)}{\left(u_{yh}^2-\dfrac{u_{yxh}^2}{u_{xh}^2}\right)+1}-\frac{\left(u_{yh}^2-\dfrac{u_{yxh}^2}{u_{xh}^2}\right)(1-\alpha_h^2u_{xh}^2)}{\left(u_{yh}^2-\dfrac{u_{yxh}^2}{u_{xh}^2}\right)+(1-\alpha_h^2u_{xh}^2)}\right\} > 0. \tag{19}$$

Because of condition $\left(u_{yh}^2-\dfrac{u_{yxh}^2}{u_{xh}^2}+1-\alpha_h^2u_{xh}^2\right) > 0$, for each h, the inequality (19) reduces to

$$u_{xh}^2\left(u_{yh}^2-\frac{u_{yxh}^2}{u_{xh}^2}\right) > 0, \text{ for each } h. \tag{20}$$

If the condition (20) is satisfied, $\bar{\bar{y}}_{mcs}$ is more efficient than $\bar{\bar{y}}_{mcs(0)}$. Note that condition (20) is always satisfied. So we can say that this estimator is always more efficient than $\bar{\bar{y}}_{mcs(0)}$. Also,

$$\text{MSE}(\bar{\bar{y}}_{reg(mcs)}) - \text{MSE}(\bar{\bar{y}}_{mcs})_{min} > 0,$$

$$\iff \quad \sum_{h=1}^{L} W_h^2 \bar{\bar{Y}}_h^2 u_{yh}^2 \left(1 - \frac{u_{yxh}^2}{u_{xh}^2 u_{yh}^2}\right)$$

$$- \sum_{h=1}^{L} W_h^2 \bar{\bar{Y}}_h^2 \frac{\left(u_{yh}^2 - \dfrac{u_{yxh}^2}{u_{xh}^2}\right)(1 - \alpha_h^2 u_{xh}^2)}{\left(u_{yh}^2 - \dfrac{u_{yxh}^2}{u_{xh}^2}\right) + 1 - \alpha_h^2 u_{xh}^2} > 0. \tag{21}$$

Because of condition $\left(u_{yh}^2 - \dfrac{u_{yxh}^2}{u_{xh}^2} + 1 - \alpha_h^2 u_{xh}^2\right) > 0$, for each h, the inequality (21) reduces to

$$\left(u_{yh}^2 - \frac{u_{yxh}^2}{u_{xh}^2}\right) > 0. \tag{22}$$

Note that condition (22) is always satisfied for $\bar{\bar{y}}_{mcs}$. So we can say $\bar{\bar{y}}_{mcs}$ is always more efficient than $\bar{\bar{y}}_{reg(mcs)}$. In the followinsg section we suggest a new estimator and its properties will be considered.

4.3 Suggested Combined Estimators in the Cluster Sampling

In this section, using the prior value of certain population parameter(s) of the auxiliary variable and following Gupta and Shabbir (2008), we suggest some classes of estimators in the cluster sampling and study their properties considering the effect of the population parameter(s) to the MSE. In cluster sampling, the combined version of the estimator suggested by Gupta and Shabbir (2008) can be given by:

$$\bar{\bar{y}}_{acs} = \left\{\omega_1 \bar{\bar{y}} + \omega_2 (\bar{\bar{X}}_{st} - \bar{\bar{x}}_{st})\right\} \left(\frac{\lambda_{st} + \eta_{st} \bar{\bar{X}}_{st}}{\lambda_{st} + \eta_{st} \bar{\bar{x}}_{st}}\right), \tag{23}$$

where η_{st} and λ_{st} are either values or the functions of the known parameters of the auxiliary variable X, such as $\lambda_{st(2)} = \sum_{h=1}^{L} \rho_h$ (there are other examples in Table 1). Similar to equation (8), the Bias and MSE of $\bar{\bar{y}}_{acs}$ are, respectively, given by

$$\begin{aligned} B(\bar{\bar{y}}_{acs}) &\approx \omega_1 \bar{\bar{Y}}(1 - \alpha_{st}u_{11} + \alpha_{st}^2 u_{20}) + \omega_2 \bar{\bar{X}}(\alpha_{st}u_{20}) - \bar{\bar{Y}} \\ \text{MSE}(\bar{\bar{y}}_{acs}) &\approx (\omega_1 - 1)^2 \bar{\bar{Y}}^2 + \omega_1^2 \bar{\bar{Y}}^2 (u_{02} - 4\alpha_{st}u_{11} + 3\alpha_{st}^2 u_{20}) \\ &\quad + \omega_2^2 \bar{\bar{X}}^2 u_{20} - 2\omega_1 \bar{\bar{Y}}^2 (\alpha_{st}^2 u_{20} - \alpha_{st}u_{11}) \\ &\quad - 2\alpha_{st}\omega_2 \bar{\bar{X}}\bar{\bar{Y}} u_{20} - 2\omega_1\omega_2 \bar{\bar{X}}\bar{\bar{Y}}(u_{11} - 2\alpha_{st}u_{20}), \end{aligned} \tag{24}$$

where

$$u_{02} = E\left(\frac{\bar{\bar{y}}_{st} - \bar{\bar{Y}}}{\bar{\bar{Y}}}\right)^2, \qquad u_{20} = E\left(\frac{\bar{\bar{x}}_{st} - \bar{\bar{X}}}{\bar{\bar{X}}}\right)^2,$$

$$u_{11}^2 = E\left(\frac{\bar{\bar{y}}_{st} - \bar{\bar{Y}}}{\bar{\bar{Y}}}\right)\left(\frac{\bar{\bar{x}}_{st} - \bar{\bar{X}}}{\bar{\bar{X}}}\right), \qquad \alpha_{st} = \frac{\eta_h \bar{\bar{X}}_h}{\lambda_h + \eta_h \bar{\bar{X}}_h}.$$

We obtain the optimum values as

$$\omega_1^* = \frac{1 - \alpha_{st}^2 u_{20}}{1 + u_{02} - \dfrac{u_{11}^2}{u_{20}} - \alpha_{st}^2 u_{20}},$$

$$\omega_2^* = \frac{\bar{\bar{Y}}}{\bar{\bar{X}}}\left\{\frac{(u_{11} - 2\alpha_{st}u_{20})(1 - \alpha_{st}^2 u_{20})}{u_{20} + u_{20}u_{02} - u_{11}^2 - \alpha_{st}^2 u_{20}^2} + \alpha_{st}\right\}. \tag{25}$$

On substituting those values for the ω_1^* and ω_2^* in equation (24), by accepting condition $u_{20} + u_{20}u_{02} - u_{11}^2 - \alpha_{st}^2 u_{20}^2 > 0$, the minimum MSE of $\bar{\bar{y}}_{acs}$ can be written as follows:

$$\text{MSE}(\bar{\bar{y}}_{acs})_{min} \approx \bar{\bar{Y}}^2 \frac{\left(u_{02} - \dfrac{u_{11}^2}{u_{20}}\right)(1 - \alpha_{st}^2 u_{20})}{\left(u_{02} - \dfrac{u_{11}^2}{u_{20}}\right) + 1 - \alpha_{st}^2 u_{20}}. \tag{26}$$

4.3.1 Efficiency Comparisons for Combined Estimators

In this section, we compare the efficiencies of the combined estimators as follow:

$$\text{MSE}(\bar{\bar{y}}_{acs(0)})_{min} - \text{MSE}(\bar{\bar{y}}_{acs})_{min} > 0,$$

$$\iff \quad \bar{\bar{Y}}^2 \frac{u_{02}^2 - \dfrac{u_{11}^2}{u_{20}^2}}{u_{02}^2 - \dfrac{u_{11}^2}{u_{20}^2} + 1} - \bar{\bar{Y}}^2 \frac{\left(u_{02}^2 - \dfrac{u_{11}^2}{u_{20}^2}\right)(1 - \alpha_{st}^2 u_{20}^2)}{\left(u_{02}^2 - \dfrac{u_{11}^2}{u_{20}^2}\right) + 1 - \alpha_{st}^2 u_{20}^2} > 0. \tag{27}$$

Because of condition $\left(u_{02}^2 - \dfrac{u_{11}^2}{u_{20}^2} + 1 - \alpha_{st}^2 u_{20}^2\right) > 0$, the inequality (27) reduces to

$$u_{20}^2 \left(u_{02}^2 - \frac{u_{11}^2}{u_{20}^2}\right) > 0. \tag{28}$$

Note that condition (28) is always satisfied for $\bar{\bar{y}}_{acs}$. So we can say $\bar{\bar{y}}_{acs}$ is always more efficient than $\bar{\bar{y}}_{acs(0)}$. Also,

$$\text{MSE}(\bar{\bar{y}}_{reg(acs)}) - \text{MSE}(\bar{\bar{y}}_{acs})_{min} > 0,$$

$$\iff \quad \bar{\bar{Y}}^2 u_{02}^2 \left(1 - \frac{u_{11}^2}{u_{20}^2 u_{02}^2}\right) - \bar{\bar{Y}}^2 \frac{\left(u_{02}^2 - \dfrac{u_{11}^2}{u_{20}^2}\right)(1 - \alpha_{st}^2 u_{20}^2)}{\left(u_{02}^2 - \dfrac{u_{11}^2}{u_{20}^2}\right) + 1 - \alpha_{st}^2 u_{20}^2} > 0. \tag{29}$$

Because of condition $\left(u_{02}^2 - \dfrac{u_{11}^2}{u_{20}^2} + 1 - \alpha_{st}^2 u_{20}^2\right) > 0$, the inequality (29) reduces to

$$\left(u_{02}^2 - \frac{u_{11}^2}{u_{20}^2}\right) > 0. \tag{30}$$

Note that condition (30) is always satisfied for $\bar{\bar{y}}_{acs}$. So we can say $\bar{\bar{y}}_{acs}$ is always more efficient than $\bar{\bar{y}}_{reg(acs)}$.

The suggested optimum estimators depend on the population parameter $\bar{\bar{Y}}$. In application we can use $\bar{\bar{y}}_{mc}$ (or $\bar{\bar{y}}_{mcs}, \bar{\bar{y}}_{acs}$) insted of $\bar{\bar{Y}}$ to compute the optimum estimators.

Table 1. Some members of suggested estimators

			Separate Estimators			Combined Estimators		
	$\bar{\bar{y}}_{mc}$		$\bar{\bar{y}}_{mcs}$			$\bar{\bar{y}}_{acs}$		
	η	λ	η_h	λ_h		η_{st}	λ_{st}	
$\bar{\bar{y}}_{mc(0)}$	0	1	$\bar{\bar{y}}_{mcs(0)}$	0	1	$\bar{\bar{y}}_{acs(0)}$	0	1
$\bar{\bar{y}}_{mc(1)}$	1	ρ_{XY}	$\bar{\bar{y}}_{mcs(1)}$	1	ρ_h	$\bar{\bar{y}}_{acs(1)}$	1	$\sum_{h=1}^{L} W_h \rho_h$
$\bar{\bar{y}}_{mc(2)}$	1	C_X	$\bar{\bar{y}}_{mcs(2)}$	1	C_{Xh}	$\bar{\bar{y}}_{acs(2)}$	1	$\sum_{h=1}^{L} W_h C_{Xh}$
$\bar{\bar{y}}_{mc(3)}$	1	$\beta_{2(X)}$	$\bar{\bar{y}}_{mcs(3)}$	1	$\beta_{2h}(X)$	$\bar{\bar{y}}_{acs(3)}$	1	$\sum_{h=1}^{L} W_h \beta_{2h}(X)$
$\bar{\bar{y}}_{mc(4)}$	$\beta_{2(X)}$	C_X	$\bar{\bar{y}}_{mcs(4)}$	$\beta_{2h}(X)$	C_{Xh}	$\bar{\bar{y}}_{acs(4)}$	$\sum_{h=1}^{L} W_h \beta_{2h}(X)$	$\sum_{h=1}^{L} W_h C_{Xh}$
$\bar{\bar{y}}_{mc(5)}$	C_X	$\beta_{2(X)}$	$\bar{\bar{y}}_{mcs(5)}$	C_{Xh}	$\beta_{2h}(X)$	$\bar{\bar{y}}_{acs(5)}$	$\sum_{h=1}^{L} W_h C_{Xh}$	$\sum_{h=1}^{L} W_h \beta_{2h}(X)$

5 Numerical Examples

5.1 Cluster Sampling

For numerical comparisons of estimators in the simple random sampling, consider the cost and income data for 1388 A.H from Statistical Centre of Iran. In order to reduce computation, 200 households in the population, we consider 40 clusters each consists of five households. Clustering which has been done by Statistical Centre of Iran is based on households addresses. We consider cost as the random variable under consideration, the main variable Y, and number of household members as the auxiliary variable X. We choose 10 randomly selected clusters, and the total sample of 50 households. The following information is obtained data.

$N = 40$	$n = 10$
$\bar{Y} = 91945181.72$	$\bar{X} = 3.575$
$C_Y^2 = 1.251288889$	$C_X^2 = 0.1418899476$
$S_{XY} = 16731039.05$	$\rho_{XY} = 0.1207989844$
$\beta_2(X) = 0.08760911251$	$\beta_1(X) = 2.699611204$

Although Gupta and Shabbir (2008) claim that various transformations of the auxiliary variable do not affect the value of the minimum MSE, we show that the specific values of λ and η play a role on the minimum MSE. For this reason, we decided to calculate the minimum MSE values of $\bar{\bar{y}}_{mc}$ using different values of λ and η as shown in Table 1. The minimum MSE values for the members of $\bar{\bar{y}}_{mc}$, given in Table 1, have been obtained using equation (10). Besides MSEs of the classical ratio estimator $\bar{\bar{y}}_0$ and the regression estimator $\bar{\bar{y}}_{reg}$ have been obtained. These values are given in Table 2. From Table 2, we observe that $\bar{\bar{y}}_{mc(3)}$ is the most efficient estimator for the cluster sampling data. From this result, we can say that the best efficiency is obtained when η and λ are defined as 1 and $\beta_{2(X)}$. As a result, $\bar{\bar{y}}_{mc(i)}$ $(i = 1, 2, 3, 4, 5)$ is more efficient than $\bar{\bar{y}}_0$, $\bar{\bar{y}}_{reg}$ and $\bar{\bar{y}}_{mc(0)}$.

5.2 Cluster Sampling in Large Population

We use the data set from previous section, data are divided into four smaller sub-population that clusters in a geographic area (neighborhood) are put in a sub-population. The summary statistics of the data are given in Table 3. We used the Neyman allocation method for determining the sample sizes of each stratum (see cochran, 1977). The minimum MSE values for the members

Table 2. MSE values of proposed estimators

Separate estimators	
$\bar{\bar{y}}_{mc}$	MSE
$\bar{\bar{y}}_{mc(0)}$	$1.872216196\ \ 10^{14}$
$\bar{\bar{y}}_{mc(1)}$	$1.872118922\ \ 10^{14}$
$\bar{\bar{y}}_{mc(2)}$	$1.872131137\ \ 10^{14}$
$\bar{\bar{y}}_{mc(3)}$	$1.872117147\ \ 10^{14}$
$\bar{\bar{y}}_{mc(4)}$	$1.872194807\ \ 10^{14}$
$\bar{\bar{y}}_{mc(5)}$	$1.872124562\ \ 10^{14}$
$\bar{\bar{y}}_{reg}$	$1.914617585\ \ 10^{14}$
$\bar{\bar{y}}_0$	$4.960385258\ \ 10^{15}$

Table 3. Summary of data (data for the stratified random sampling)

$n_1 = 3$	$N_1 = 10$
$n_2 = 3$	$N_2 = 10$
$n_3 = 3$	$N_3 = 10$
$n_4 = 3$	$N_4 = 10$
$C_{Y1}^2 = 0.4109078632$	$\bar{\bar{y}}_1 = 53479269.72$
$C_{Y2}^2 = 0.550190469$	$\bar{\bar{y}}_2 = 85265883.72$
$C_{Y3}^2 = 2.187934113$	$\bar{\bar{y}}_3 = 113342806.1$
$C_{Y4}^2 = 0.3627429521$	$\bar{\bar{y}}_4 = 115692767.3$
$C_{X1}^2 = 0.2027136904$	$\bar{\bar{X}}_1 = 3.42$
$C_{X2}^2 = 0.1200219671$	$\bar{\bar{X}}_2 = 3.84$
$C_{X3}^2 = 0.1471280458$	$\bar{\bar{X}}_3 = 3.64$
$C_{X4}^2 = 0.5289629483$	$\bar{\bar{X}}_4 = 3.4$
$\rho_{XY1} = 0.32567812$	$S_{XY1} = 17191505.98$
$\rho_{XY2} = 0.2786069615$	$S_{XY2} = 23441499.41$
$\rho_{XY3} = -0.2892457541$	$S_{XY3} = -6770600.512$
$\rho_{XY4} = 1.142857143$	$S_{XY4} = 32629753.45$
$\beta_2(X_1) = 0.2535825673$	$\beta_1(X_1) = 2.50823672$
$\beta_2(X_2) = -0.1899703271$	$\beta_1(X_2) = 2.613252181$
$\beta_2(X_3) = 0.7191360431$	$\beta_1(X_3) = 2.423905932$
$\beta_2(X_4) = -0.2429771612$	$\beta_1(X_4) = 2.3275$

of $\bar{\bar{y}}_{mcs}$ and $\bar{\bar{y}}_{acs}$, are given in Table 4. These values are obtained using equations (19) and (27), respectively. In addition to these MSE values, MSE values for the estimators of the classical combined ratio $\bar{\bar{y}}_{0(acs)}$, the classical separate ratio $\bar{\bar{y}}_{0(mcs)}$, the combined regression $\bar{\bar{y}}_{reg(acs)}$, and the separate regression $\bar{\bar{y}}_{reg(mcs)}$ have also been computed. All of these MSE values are given in Table 4. From Table 4, we can see that combined estimators are more efficient than separate estimators and that $\bar{\bar{y}}_{mcs(1)}$ is the most efficient in the separate estimators for this data set. When we further examine Table 4, we see that the differences among the MSE values for the members of $\bar{\bar{y}}_{mcs}$ and $\bar{\bar{y}}_{acs}$ are small. As a result, $\bar{\bar{y}}_{acs(3)}$ is the most efficient estimator in the combined estimators and it is also more efficient than the combined regression estimator, $\bar{\bar{y}}_{reg(acs)}$. From this, we can say that the best efficiency is obtained when λ_{st} and η_{st} are defined as $\sum_{h=1}^{L} W_h \beta_{2h}(X)$ and 1 respectively.

Table 4. MSE values of proposed estimators in larg population

Separate Estimators		Combined Estimators	
$\bar{\bar{y}}_{mcs}$	MSE	$\bar{\bar{y}}_{acs(0)}$	MSE
$\bar{\bar{y}}_{mcs(0)}$	$1.219798867\ 10^{14}$	$\bar{\bar{y}}_{acs(0)}$	$9.703731132\ 10^{13}$
$\bar{\bar{y}}_{mcs(1)}$	$1.219067370\ 10^{14}$	$\bar{\bar{y}}_{acs(1)}$	$9.703482235\ 10^{13}$
$\bar{\bar{y}}_{mcs(2)}$	$1.219194405\ 10^{14}$	$\bar{\bar{y}}_{acs(2)}$	$9.703499509\ 10^{13}$
$\bar{\bar{y}}_{mcs(3)}$	$1.219075213\ 10^{14}$	$\bar{\bar{y}}_{acs(3)}$	$9.703450933\ 10^{13}$
$\bar{\bar{y}}_{mcs(4)}$	$1.219482672\ 10^{14}$	$\bar{\bar{y}}_{acs(4)}$	$9.703726231\ 10^{13}$
$\bar{\bar{y}}_{mcs(5)}$	$1.219074110\ 10^{14}$	$\bar{\bar{y}}_{acs(5)}$	$9.703455053\ 10^{13}$
$\bar{\bar{y}}_{reg(mcs)}$	$1.342568113\ 10^{14}$	$\bar{\bar{y}}_{reg(acs)}$	$9.816407646\ 10^{13}$
$\bar{\bar{y}}_{0(mcs)}$	$1.425773984\ 10^{14}$	$\bar{\bar{y}}_{0(acs)}$	$1.425773984\ 10^{14}$

Note that we get more efficient estimators when we define $\eta_{st} = 1$ and λ_{st} as the correlation coefficient in the family of estimators, $\bar{\bar{y}}_{acs(3)}$. We conclude that the minimum MSE values of the $\bar{\bar{y}}_{mc}$, $\bar{\bar{y}}_{acs}$ and $\bar{\bar{y}}_{mcs}$ can change according to the definition of λ and η using known population parameter(s). Although the above example illustrates the higher MSE of suggested estimators, the claim should be confirmed by using a simulation which is beyond the purpose of this paper.

References

Cochran, W.G. (1977). *Sampling Techniques*, John Wiley and Sons, NewYork.

Gupta, S. and Shabbir, J. (2008). On Improvement in Estimating the Population Mean in Simple Random Sampling, *J. Appl. Stat.*, **35**, 559-566.

Kadilar, C. and Cingi, H. (2003). Ratio Estimators in Stratified Random Sampling, *Biom. J.*, **45**, 218-225.

Koyuncua, N. and Kadilar, C. (2009). Ratio and Product Estimators in Stratified Random Sampling, *J. Statist. Plann. Inference*, **139**, 2552-2558.

Koyuncua, N. and Kadilara, C. (2010). On Improvement in Estimating Population Mean in Stratified Random Sampling, *J. App. Stat.*, **37**, 999-1013.

Shabbir, J. and Gupta, S. (2005). Improved Ratio Estimators in Stratified Sampling, *Amer. J. Math. Management Sci.*, **25**, 293-311.

Singh, H.P. and Tailor, R. (2003). Use of Known Correlation Coefficient in Estimating the Finite Population Mean, *Stat. Transit.*, **6**, 555-560.

Singh, H.P., Tailor, R., Singh, S. and Kim, J.M. (2008). A Mmodified Estimator of Population Mean Uusing Power Transformation, *Statist. Papers*, **49**, 37-58.

Singh, H.P. and Vishwakarma, G.K. (2006). An Efficient Variant of the Product and Ratio Estimators in Stratified Random Sampling, *Stat. Transit.*, **7**, 1311-1325.

Singh, H.P. and Vishwakarma, G.K. (2008). A Family of Estimators of Population Mean Using Auxiliary Information in Stratified Sampling, *Comm. Statist. Theory Methods*, **37**, 1038-1050.

Sisodia, B.V.S. and Dwivedi, V.K. (1981). A Modified Ratio Estimator Using Coefficient of Variation of Auxiliary Variable, *J. Indian Soc. Agricultural Statist.*, **33**, 13-18.

Upadhyaya, L.N. and Singh, H.P. (1999). Use of Transformed Auxiliary Variable in Estimating the Finite Population Mean, *Biom. J.*, **41**, 627-636.

http://www.amar.org.ir/

M. Sakizadeh
Department of General and Basic Science,
West Tehran Branch,
Islamic Azad University,
Tehran, Iran.
email: *mansoure6611@yahoo.com*

A. Gerami
Qazvin Branch,
Islamic Azad University,
Qazvin, Iran.
email: *agerami@ut.ac.ir*