Pak. J. Statist.
2023 Vol. 39(3), 195-212

GENERALIZED MULTIVARIATE SHRINKAGE ESTIMATORS IN MPSS SAMPLING

Muqaddas Javed[1§] and **Zahoor Ahmad**[2]

[1] Department of Statistics, University of Gujrat, Gujrat, Pakistan.
[2] Department of Mathematics and Statistics, University of Lahore Sargodha, Pakistan
[§] Corresponding author Email: muqaddas.javed@uog.edu.pk

ABSTRACT

In this paper, we suggested generalized shrinkage regression, ratio and regression-cum-ratio estimators for population mean in multi-phase stratified systematic (MPSS) sampling design using multi-auxiliary information when information on all auxiliary variables is not available for population. The expressions of mean square error and bias are derived for suggested estimators. The extension of these estimators in bivariate and multivariate is also discussed and some important special cases are deduced from the general class. An empirical and simulation studies are conducted to assess the performance of proposed design and estimators and found suggested MPSS design perform better than estimators of multi-phase simple random sampling estimators.

KEYWORDS

Shrinkage estimator; regression type estimator; stratified random sampling, multi-phase sampling; systematic sampling, mixed sampling design.

1. INTRODUCTION

The prime objective of survey statisticians is to introduce such an estimating methodology for population parameters which provide precise results. Mean estimation has always been a mandatory question in the theory of survey sampling. Provided a sampling design, many estimation and selection procedures has been developed by researchers to improve the precision and efforts are still continuing. The use of auxiliary information for the accuracy and precision of estimators is always appreciated and many ratio, product and regression estimators are available in the literature using single, two and multiple auxiliary variables.

In practical surveys, the problem is to estimate population means of variables of interest. For example, in a typical socio-economic survey conducted in rural areas in Indo-Pak subcontinent, the multiple variables of interests may be size of household, monthly income and expenditure of the household, number of unemployed persons, number of illiterates, number of persons engaged in agriculture, amount of land owned, leased and leased out, number of cattle owned etc. In some situations, the auxiliary information may be available through the past census data or conveniently collected. For example, in a village land survey, the information on the variables such as area of the village, cultivable area, grazing grounds etc. may be easily obtained through the past census data and may be used to estimate the means of variables of interest.

Bowley (1926) and Neyman (1934) have provided the foundation for the use of auxiliary information and Olkin (1958) was the first who used multi-auxiliary information for the estimation of study variables when information on all auxiliary variables is available for population. Later on Raj (1965) achieved higher precision of difference estimator using information of several auxiliary variates and showed that difference estimator is comparable to ratio estimator.

Sen (1972) used multi-auxiliary information for the development of multivariate ratio estimators using two-phase sampling. Sahoo and Sahoo (1993) proposed a class of estimators using the information of two auxiliary variables under two-phase sampling. Ahmed (2003) has used multi-auxiliary information for chain based general estimators under multiphase sampling, while Paradhan (2005) has put forward chain regression estimators using information of several auxiliary variates in two phase sampling design. Tikkiwal and Ghiya (2004) have worked on a generalized class of composite estimators for small domains using auxiliary information under different sampling designs.

If we have information on multi-auxiliary variables, practically sometimes either information for all these auxiliary is available from population or available for some variables or not available for all auxiliary variables. These three cases are first time discussed by Samiuddin and Hanif (2007) and categorized their estimators in the following three cases: i) estimators when information on all auxiliary variables is known for population (Full Information Case (FIC)), ii) estimators when information on some auxiliary variables is known for population (Partial Information Case (PIC)), and iii) estimators when information on all auxiliary variables is unknown for population (No Information Case (NIC)). Hanif et al. (2009) proposed generalized multivariate ratio estimator in multiphase sampling using multi- auxiliary variables considering FIC and NIC. Regression, ratio, regression-in-regression and regression-cum-ratio estimation methods are used for estimating population mean of single/several study variable(s) in two-phase/multi-phase sampling using multi-auxiliary variables for FIC, PIC and NIC by Ahmad et al. (2009a, 2009b, 2010a, 2010b, 2010c, 2013)., Some other useful contributions on the application of two phase and auxiliary information in estimation methods include Srivastava (1971), Das and Tripathi (1978), Khare and Srivastava (1981), Srivastava and Jhajj (1983), Upadhyaya and Singh (1983), Sukhatme et al. (1984), Mukhopadhyay (2000), Cochran (1977), Kadilar and Cingi (2006), Javed, et al. (2014), Noor-ul-Amin et al. (2016), Zaman and Bulut (2019), Abid et al. (2018), Zaman and Bulut (2020) and Iqbal, et al. (2020).

In this paper, we suggest generalized shrinkage regression, ratio and regression-cum-ratio estimators for estimation of population mean of the variables of interest. Estimators are developed under multiphase stratified systematic sampling design using multi-auxiliary information when information on all auxiliary variables is unknown for populations that usually occur in practical situations.

After introducing the topic in section 1, sampling scheme, useful notations and mathematical expectations are discussed in section 2. Generalized shrinkage regression, ratio and regression-cum-ratio estimators are suggested in section 3. The extension of suggested estimators in bivariate and multivariate is also discussed and special cases are deduced in section 3. In the last section the performance of the suggested estimators is discussed based on empirical study.

2. MULTI-PHASE STRATIFIED SYSTEMATIC SAMPLING

Let a population of size N is stratified in to L homogeneous strata. N_h, $n_{h(k)}$ and $n_{h(m)}$ denote the h^{th} population stratum size, k^{th}-phase stratum sample size and m^{th}-phase stratum sample size respectively. Sample sizes are allocated proportionally at both phases and samples are selected using systematic sampling from every stratum at each phase.

Let y_j and x_i denote the j^{th} study and i^{th} auxiliary variables for $j=1,...,p$ and $i=1,...,q$ with population means $\bar{Y}_j$ and $\bar{X}_i$; variances $S^2_{y_j}$ and $S^2_{x_i}$; and coefficient of variation (CV's) C_{y_j} and C_{x_i} respectively. The covariance between j^{th} study and i^{th} auxiliary variable is denoted by $S_{y_j x_i}$. Further suppose $\bar{y}_{j_{ss(k)}}$ ($\bar{y}_{j_{ss(m)}}$) and $\bar{x}_{i_{ss(k)}}$ ($\bar{x}_{i_{ss(m)}}$) denote the k^{th} (m^{th}) phase sample means of j^{th} study and i^{th} auxiliary variables respectively.

To derive the vector of bias and variance covariance matrices of proposed estimators, we define the following absolute sampling errors and then require expressions of mathematical expectations under MPSS sampling design. Let the sampling errors $\bar{e}_{y_{j_{ss(m)}}} = \bar{y}_{j_{ss(m)}} - \bar{Y}_j$, $\bar{e}_{x_{i_{ss(k)}}} = \bar{x}_{i_{ss(k)}} - \bar{X}_i$ and $\bar{e}_{x_{i_{ss(m)}}} = \bar{x}_{i_{ss(m)}} - \bar{X}_i$, and further it is assumed that $E_m\left(\bar{e}_{y_{j_{ss(m)}}}\right) = E_m\left(\bar{e}_{x_{i_{ss(m)}}}\right) = E_k\left(\bar{e}_{x_{i_{ss(k)}}}\right) = 0$. Also,

$$\left.\begin{aligned}
&E_m\left(\bar{e}^2_{y_{j_{ss(m)}}}\right) = \sum_{h=1}^{L} \theta_{h(m)} \eta_{y_{hj}} S^2_{y_{h_j}} = S^2_{y_j},\ E_m\left(\bar{e}^2_{x_{i_{ss(m)}}}\right) = \sum_{h=1}^{L} \theta_{h(m)} \eta_{x_{hi}} S^2_{x_{h_i}} = S^2_{x_i},\\
&E_m\left(\bar{e}_{y_{j_{ss(m)}}} \bar{e}_{x_{i_{ss(m)}}}\right) = \sum_{h=1}^{L} \theta_{h(m)} \sqrt{\eta_{y_{hj}} \eta_{x_{hi}}}\, \rho_{y_{h_j} x_{h_i}} S_{y_{h_j}} S_{x_{h_i}} = S_{y_j x_i},\\
&E_k\left(\bar{e}_{y_{j_{ss(k)}}} \bar{e}_{x_{i_{ss(k)}}}\right) = \sum_{h=1}^{L} \theta^*_{h(k)} \sqrt{\eta^*_{y_{hj}} \eta^*_{x_{hi}}}\, \rho_{y_{h_j} x_{h_i}} S_{y_{h_j}} S_{x_{h_i}} = S^*_{y_j x_i},\ E_k\left(\bar{e}_{x_{i_{ss(k)}}} \bar{e}_{x_{l_{ss(k)}}}\right)\\
&\qquad = E_m\left(\bar{e}_{x_{i_{ss(m)}}} \bar{e}_{x_{l_{ss(m)}}}\right) = 0\\
&\text{for } (i \neq l)\\
&E_m\left(\bar{e}_{y_{j_{ss(m)}}} \bar{e}_{x_{i_{ss(k)}}}\right) = E_k E_{m/k}\left(\bar{e}_{y_{j_{ss(m)}}} \bar{e}_{x_{i_{ss(k)}}}\right) = E_k E_{m/k}\left[\bar{e}_{y_{j_{ss(m)}}}\left(\bar{e}_{x_{i_{ss(k)}}} - \bar{e}_{x_{i_{ss(m)}}}\right)\right]\\
&\qquad = E_k\left(\bar{e}_{y_{j_{ss(k)}}} \bar{e}_{x_{i_{ss(k)}}}\right) - E_m\left(\bar{e}_{y_{j_{ss(m)}}} \bar{e}_{x_{i_{ss(m)}}}\right),
\end{aligned}\right| \quad (1)$$

where

$$\theta_{h(k)} = \frac{N_h\left(N_h-1\right)}{N^2 n_{h(k)}},\ \ \theta_{h(m)} = \frac{N_h\left(N_h-1\right)}{N^2 n_{h(m)}},\ \ \eta_{y_{hj}} = \left\{1+\left(n_{h(m)}-1\right)\rho_{y_{hj}}\right\},$$

$$\eta_{x_{hi}} = \left\{1+\left(n_{h(m)}-1\right)\rho_{x_{hi}}\right\},\ \ \eta^*_{y_{hj}} = \left\{1+\left(n_{h(k)}-1\right)\rho_{y_{hj}}\right\}$$

and $\eta^*_{x_{hi}} = \left\{1+\left(n_{h(k)}-1\right)\rho_{x_{hi}}\right\}$.

For bivariate case, let

$$\left.\begin{array}{l}
\bar{Y}_{(2\times 1)} = \left[\bar{Y}_j\right]_{(2\times 1)},\ \bar{D}_{y_{(2\times 1)}} = \left[\bar{e}_{y_{j_{ss(m)}}}\right]_{(2\times 1)},\ B_{(2\times q)} = \left[\beta_{ij}\right]_{(2\times q)}, \\
\qquad \Delta_{(2\times l)} = \left[\dfrac{\bar{Y}_j}{\bar{X}_i}\delta_{ij}\right]_{(2\times l)} \\
\bar{D}_{x_{1(q\times 1)}} = \left[\left(\bar{e}_{x_{i_{ss(k)}}} - \bar{e}_{x_{i_{ss(m)}}}\right)\right]_{(q\times 1)},\ \bar{D}_{x_{2(l\times 1)}} = \left[\left(\bar{e}_{x_{i_{ss(k)}}} - \bar{e}_{x_{i_{ss(m)}}}\right)\right]_{(l\times 1)} \\
\bar{D}_{y'_{(2\times l)}} = \left[\delta_{ij}\bar{e}_{y_{j_{ss(m)}}}\right]_{(2\times l)},\ \bar{D}_{x_{(l\times 1)}} = \left[\dfrac{1}{\bar{X}_i}\left(\bar{e}_{x_{i_{ss(m)}}} - \bar{e}_{x_{i_{ss(k)}}}\right)\right]_{(l\times 1)}
\end{array}\right|, \qquad (2)$$

$$\left.\begin{array}{l}
E_m\left(\bar{D}_y\right) = E_k\left(\bar{D}_{x_1}\right) = E_k\left(\bar{D}_{x_2}\right) = E_m\left(\bar{D}_{x_1}\right) = E_m\left(\bar{D}_{x_2}\right) = 0 \\
\Sigma_y = E_m\left(\bar{D}_y\bar{D}'_y\right) = \left[S_{y_iy_j}\right]_{(2\times 2)},\ \Sigma_{x_1} = E_m\left(\bar{D}_{x_1}\bar{D}'_{x_1}\right) = \left[S_{x_ix_j}\right]_{(q\times q)} \\
\Sigma_{x_2} = E_m\left(\bar{D}_{x_2}\bar{D}'_{x_2}\right) = \left[S_{x_ix_j}\right]_{(l\times l)}
\end{array}\right\}, \qquad (3)$$

$$\left.\begin{array}{l}
E_m\left(\bar{D}_y\bar{D}'_{x_1}\right) = \Sigma_{yx_1} = \left[\left(S^*_{y_jx_i} - S_{y_jx_i}\right)\right]_{(2\times q)},\ E_m\left(\bar{D}_y\bar{D}'_{x_2}\right) \\
\qquad = \Sigma_{yx_2} = \left[\left(S^*_{y_jx_i} - S_{y_jx_i}\right)\right]_{(2\times l)} \\
E_m\left(\bar{D}_{x_1}\bar{D}'_y\right) = \Sigma_{x_1y} = \left[\left(S^*_{y_jx_i} - S_{y_jx_i}\right)\right]_{(q\times 2)},\ E_m\left(\bar{D}_{x_2}\bar{D}'_y\right) = \Sigma_{x_2y} \\
\qquad = \left[\left(S^*_{y_jx_i} - S_{y_jx_i}\right)\right]_{(l\times 2)} \\
E\left(\bar{D}_{y'_{(2\times l)}}\bar{D}_{x_{(l\times 1)}}\right) = \Sigma_{y'x_{(2\times 1)}} = \left[\sum_{i=q+1}^{s}\dfrac{\delta_{ij}}{\bar{X}_i}\left(S^*_{y_jx_i} - S_{y_jx_i}\right)\right]_{(2\times 1)},\ E_m\left(\bar{D}_{x_1}\bar{D}'_{x_2}\right) \\
\qquad = E_m\left(\bar{D}_{x_2}\bar{D}'_{x_1}\right) = 0
\end{array}\right\}. \qquad (4)$$

3. GENERALIZED SHRINKAGE REGRESSION, RATIO AND REGRESSION-CUM-RATIO ESTIMATORS FOR POPULATION MEAN UNDER MPSS DESIGN

In this section we proposed regression, ratio and regression-cum-ratio estimators for the estimation of population mean. Estimators are suggested when information on population parameters for auxiliary variables are unknown. The estimators are suitable when the relationship between study and auxiliary variables is linear. Further correlation between auxiliary variables must not be significant. Special cases are also discussed in this section.

3.1 Generalized Shrinkage Regression Estimator

Using notations given in section 2, we can suggest the following generalized regression estimator

$$t_{reg(1)} = \lambda\left[\bar{y}_{ss(m)} + \sum_{i=1}^{q}\beta_i\left(\bar{x}_{i_{ss(k)}} - \bar{x}_{i_{ss(m)}}\right)\right] = \lambda t_{reg}\ , \tag{5}$$

where λ and β_i be the optimizing constants. To derive the expression for bias and MSE, we can write t_{reg} in sampling error form as

$$t_{reg} = \bar{Y} + \bar{e}_{y_{ss(m)}} + \sum_{i=1}^{q}\beta_i\left(\bar{e}_{x_{i_{ss(k)}}} - \bar{e}_{x_{i_{ss(m)}}}\right).$$

For MSE of t_{reg} , first we can write

$$MSE\left(t_{reg}\right) = E_m\left(t_{reg} - \bar{Y}\right)^2 = E_m\left[\bar{e}_{y_{ss(m)}} + \sum_{i=1}^{q}\beta_i\left(\bar{e}_{x_{i_{ss(k)}}} - \bar{e}_{x_{i_{ss(m)}}}\right)\right]^2 \tag{6}$$

$$= E_m\left[\bar{e}_{y_{ss(m)}}\left\{\bar{e}_{y_{y_{ss(m)}}} + \sum_{i=1}^{q}\beta_i\left(\bar{e}_{x_{i_{ss(k)}}} - \bar{e}_{x_{i_{ss(m)}}}\right)\right\}\right]$$

$$= E_m\left(\bar{e}^2_{y_{y_{ss(m)}}}\right) + \sum_{i=1}^{q}\beta_i E_m\left\{\bar{e}_{y_{y_{ss(m)}}}\left(\bar{e}_{x_{i_{ss(k)}}} - \bar{e}_{x_{i_{ss(m)}}}\right)\right\}.$$

Using relevant results of expectations from (1), we have

$$MSE\left(t_{reg}\right) = S_y^2 + \sum_{i=1}^{q}\beta_i\left(S^*_{yx_i} - S_{yx_i}\right). \tag{7}$$

For the optimum value of i^{th} component of β , differentiating (7) with respect to each component and solving the q equation for $\beta's$, we have

$$\beta_i = \frac{-E_m\left\{\bar{e}_{y_{ss(m)}}\left(\bar{e}_{x_{i_{ss(k)}}} - \bar{e}_{x_{i_{ss(m)}}}\right)\right\}}{E_m\left(\bar{e}_{x_{i_{ss(k)}}} - \bar{e}_{x_{i_{ss(m)}}}\right)^2} = \frac{-\left(S^*_{yx_i} - S_{yx_i}\right)}{S^*_{x_i} + S_{x_i}}\ ; i = 1, 2 \ldots q\ . \tag{8}$$

Then optimum MSE of t_{reg} is

$$MSE\left(t_{reg}\right)=S_y^2-\sum_{i=1}^{q}\left(S_{yx_i}^{*}-S_{yx_i}\right)^2\Big/\left(S_{x_i}^{*}+S_{x_i}\right).$$

Now for MSE of $t_{reg(1)}$, using Shahbaz and Hanif (2009) shrinkage theorem, we have $MSE\left(t_{reg(1)}\right)=MSE\left(t_{reg}\right)\left[1+\bar{Y}^{-2}MSE\left(t_{reg}\right)\right]^{-1}$ with $\lambda_{opt}=\left[1+\bar{Y}^{-2}MSE\left(t_{reg}\right)\right]^{-1}$, where $MSE\left(t_{reg}\right)=Var\left(t_{reg}\right)$, as t_{reg} is an unbiased estimator.

The bivariate version of shrinkage regression estimator can be written as

$$t_{reg(2)}=\begin{bmatrix}\lambda_1 & 0\\ 0 & \lambda_2\end{bmatrix}_{2\times2}\begin{bmatrix}t_{1reg}\\ t_{2reg}\end{bmatrix}_{2\times1}=\Lambda_1 t_{reg},$$

where $t_{j_{reg}}=\bar{y}_{j_{ss(m)}}+\sum_{i=1}^{q}\beta_{ij}\left(\bar{x}_{i_{ss(k)}}-\bar{x}_{i_{ss(m)}}\right)$ for $j=1,2.$

We, rewrite $t_{j_{reg}}$ in sampling errors as

$$t_{jreg}=\left[\left(\bar{Y}_j+\bar{e}_{y_{jss(m)}}\right)-\sum_{i=1}^{q}\beta_{ij}\left(\bar{e}_{x_{i_{ss(k)}}}-\bar{e}_{x_{i_{ss(m)}}}\right)\right].\ \left(j=1,2\right)$$

or $$t_{reg}=\bar{Y}_{(2\times1)}+\bar{D}_{y_{(2\times1)}}-B_{(2\times q)}\bar{D}_{x_{1(q\times1)}}$$

The expression of variance covariance matrix is as

$$\Sigma_{t_{reg}}=E_m\left[\bar{D}_{y_{(2\times1)}}-B_{(2\times q)}\bar{D}_{x_{1(q\times1)}}\right]\left[\bar{D}_{y_{(2\times1)}}-B_{(2\times q)}\bar{D}_{x_{1(q\times1)}}\right]'$$

$$=E\left(\bar{D}_{y_{(2\times1)}}\bar{D}'_{y_{(1\times2)}}\right)-E\left(\bar{D}_{y_{(2\times1)}}\bar{D}'_{x_{1(1\times q)}}\right)B'_{(q\times2)}-B_{(2\times q)}E\left(\bar{D}_{x_{1(q\times1)}}\bar{D}'_{y_{(1\times2)}}\right)$$

$$+B_{(2\times q)}E\left(\bar{D}_{x_{1(q\times1)}}\bar{D}'_{x_{1(1\times q)}}\right)B'_{(q\times2)}$$

Using the results of we get the variance covariance matrix as:

$$\Sigma_{t_{reg}}=\Sigma_{y_{(2\times2)}}-\Sigma_{yx_{1(2\times q)}}B'_{(q\times2)}-B_{(2\times q)}\Sigma_{x_1y_{(q\times2)}}+B_{(2\times q)}\Sigma_{x_{1(q\times q)}}B'_{(q\times2)}. \qquad (9)$$

Hence $MSE\left(\mathrm{t}_{reg}\right)=\Sigma_{\mathrm{t}_{reg(2\times2)}}$

The expression of optimum value of Λ_1 and MSE of $\mathrm{t}_{reg(2)}$ using the multivariate Shrinkage Estimators theorem (see Ahmad and Hanif (2016)) are

$$\Lambda_1=\bar{Y}\bar{Y}'\left\{\bar{Y}\bar{Y}'+MSE\left(t_{reg}\right)\right\}^{-1};\ MSE\left(t_{reg(2)}\right)=\bar{Y}\bar{Y}'-\bar{Y}\bar{Y}'\left\{\bar{Y}\bar{Y}'+MSE\left(t_{reg}\right)\right\}^{-1}\bar{Y}\bar{Y}'$$

where $MSE\left(t_{reg}\right)=E\left(t_{reg}-\bar{Y}\right)\left(t_{reg}-\bar{Y}\right)'=B\left(t_{reg}\right)B'\left(t_{reg}\right)+\Sigma_{t_{reg}}$ and

$B\left(t_{reg}\right)=E\left(t_{reg}-\bar{Y}\right)$, and $\Sigma_{t_{reg}}=E\left(t_{reg}-E\left(t_{reg}\right)\right)\left(t_{reg}-E\left(t_{reg}\right)\right)'=Var\left(t_{reg}\right)$,

as $B\left(t_{reg}\right)=0$. The expression of variance covariance matrix $\Sigma_{t_{reg}}$ is given in (9).

The multivariate version of $t_{reg(2)}$ can be defined by replacing 2 by p in bivariate version.

3.2 Generalized Shrinkage Ratio Estimator

Using notations given in section 2, we can suggest the following generalized ratio estimator

$$t_{ra(1)}=\delta\,\bar{y}_{ss(m)}\prod_{i=1}^{q}\left(\bar{x}_{i_{ss(k)}}/\bar{x}_{i_{ss(m)}}\right)^{\alpha_i}=\delta\,t_{ra},$$

where δ and α_i be the optimizing unknown constants. Following the procedure of deriving MSE of above section, we can write

$$MSE\left(t_{ra}\right)=S_y^2+\sum_{i=1}^{q}\alpha_i\frac{\bar{Y}}{\bar{X}_i}\left(S^*_{yx_i}-S_{yx_i}\right).$$

and expression of Bias is as

$$Bias\left(t_{ra}\right)=\sum_{i=1}^{q}\frac{\alpha_i}{\bar{X}_i}\left(S^*_{y_jx_i}-S_{y_jx_i}\right).$$

Following the previous section, the optimum value of i^{th} component of α is,

$$\alpha_i=\frac{-\bar{X}_iE_m\left\{\bar{e}_{y_{ss(m)}}\left(\bar{e}_{x_{i_{ss(k)}}}-\bar{e}_{x_{i_{ss(m)}}}\right)\right\}}{\bar{Y}E_m\left(\bar{e}_{x_{i_{ss(k)}}}-\bar{e}_{x_{i_{ss(m)}}}\right)^2}=\frac{-\bar{X}_i\left(S^*_{yx_i}-S_{yx_i}\right)}{\bar{Y}\left(S^*_{x_i}+S_{x_i}\right)};\,i=1,2\ldots q. \tag{10}$$

Then optimum MSE and Bias of t_{ra} is

$$MSE\left(t_{ra}\right)=S_y^2-\sum_{i=1}^{q}\left(S^*_{yx_i}-S_{yx_i}\right)^2\Big/\left(S^*_{x_i}+S_{x_i}\right),$$

same as for regression case but it is approximately derived.

$$Bias\left(t_{ra}\right)=-\frac{1}{\bar{Y}}\sum_{i=1}^{q}\left(S^*_{yx_i}-S_{yx_i}\right)^2\left(S^*_{x_i}+S_{x_i}\right)^{-1}.$$

Now for MSE and Bias of $t_{ra(1)}$, again using Shahbaz and Hanif (2009) and Ahmad and Hanif (2016) shrinkage theorems, we have

$$MSE\left(t_{ra(1)}\right)=\bar{Y}^2\left[MSE\left(t_{ra}\right)-\left\{Bias\left(t_{ra}\right)\right\}^2\right]\left[\bar{Y}^2+MSE\left(t_{ra}\right)+2\bar{Y}Bias\left(t_{ra}\right)\right]^{-1},$$

$$Bias\left(t_{ra(1)}\right)=\delta_{opt}\left\{Bias\left(t_{ra}\right)+\bar{Y}\right\}-\bar{Y} \quad ,$$

where $\delta_{opt}=\left[\bar{Y}^2+\bar{Y}Bias(t_{ra})\right]\left[\bar{Y}^2+2\bar{Y}Bias(t_{ra})+MSE\left(t_{ra}\right)\right]^{-1}$.

The bivariate version of shrinkage regression estimator can be written as

$$t_{ra(2)}=\begin{bmatrix}\delta_1 & 0\\ 0 & \delta_2\end{bmatrix}_{2\times 2}\begin{bmatrix}t_{1ra}\\ t_{2ra}\end{bmatrix}_{2\times 1}=\Lambda_2 t_{ra} \, ,$$

where $t_j=\bar{y}_{j_{ss(m)}}\prod_{i=1}^{q}\left(\frac{\bar{x}_{i_{ss(k)}}}{\bar{x}_{i_{ss(m)}}}\right)^{\alpha_{ij}}$ for $j=1,2.$

Using the above method, expression of $MSE\left(t_{ra}\right)$ and $B\left(t_{ra}\right)$ is

$$MSE\left(t_{ra}\right)=\Sigma_{y'x_{(2\times 1)}}\Sigma'_{y'x_{(1\times 2)}}+\Sigma_{t_{(2\times 2)}}$$

and

$$B\left(t_{ra}\right)=E\left(\bar{D}_{y'_{(2\times q)}}\bar{D}_{x_{(q\times 1)}}\right)=\Sigma_{y'x_{(2\times 1)}} \, ,$$

where $\Sigma_t=\Sigma_{y_{(2\times 2)}}+\Delta_{(2\times q)}\Sigma_{x_2y_{(q\times 2)}}+\Sigma_{yx_{2(2\times q)}}\Delta'_{(q\times 2)}+\Delta_{(2\times q)}\Sigma_{x_{2(q\times q)}}\Delta'_{(q\times 2)}$

The expression of $MSE\left(t_{ra(2)}\right)$ and $Bias\left(t_{ra(2)}\right)$ are

$$MSE\left(t_{ra(2)}\right)=\bar{Y}\bar{Y}'-\left\{\bar{Y}\bar{Y}'+\bar{Y}B'\left(t_{ra}\right)\right\}$$
$$\left\{\bar{Y}\bar{Y}'+MSE\left(t_{ra}\right)+2\bar{Y}B'\left(t_{ra}\right)\right\}^{-1}\left\{\bar{Y}\bar{Y}'+\bar{Y}B'\left(t_{ra}\right)\right\}'$$

and

$$Bias\left(t_{ra(2)}\right)=\left(\Lambda B\left(t_{ra}\right)+\left(\Lambda-I\right)\bar{Y}\right)\left(B'\left(t_{ra}\right)\Lambda+\bar{Y}'\left(\Lambda-I\right)\right),$$

where $\Lambda_2=\left\{\bar{Y}\bar{Y}'+\bar{Y}B'\left(t_{ra}\right)\right\}\left\{\bar{Y}\bar{Y}'+MSE\left(t_{ra}\right)+2\bar{Y}B'\left(t_{ra}\right)\right\}^{-1}$.

The multivariate can be defined by replacing 2 by p in bivariate version.

3.3 Generalized Shrinkage Regression-cum-Ratio Estimator

Using notations given in section 2, we can suggest the following generalized regression-cum-ratio estimator

$$t_{rera(1)}=\gamma\left[\bar{y}_{ss(m)}+\sum_{i=1}^{l}\beta_i\left(\bar{x}_{i_{ss(k)}}-\bar{x}_{i_{ss(m)}}\right)\right]\prod_{i=l+1}^{q}\left(\frac{\bar{x}_{i_{ss(k)}}}{\bar{x}_{i_{ss(m)}}}\right)^{\alpha_i}=\gamma t_{rera} \, , \quad (14)$$

where γ, β_i and α_i are unknown optimizing constants. Following the procedure of deriving MSE of above sections, we can write

$$E_m\left(t_{rera}-\bar{Y}\right)^2=E_m\left[\bar{e}_{y_{ss(m)}}+\sum_{i=1}^{l}\beta_i\left(\bar{e}_{x_{i_{ss(k)}}}-\bar{e}_{x_{i_{ss(m)}}}\right)+\sum_{i=l+1}^{q}\frac{\bar{Y}\delta_i}{\bar{X}_i}\left(\bar{e}_{x_{i_{ss(k)}}}-\bar{e}_{x_{i_{ss(m)}}}\right)\right]^2$$

and

$$MSE\left(t_{rera}\right)=S_y^2+\sum_{i=1}^{l}\beta_i\left(S_{yx_i}^*-S_{yx_i}\right)+\sum_{i=l+1}^{q}\frac{\bar{Y}\delta_i}{\bar{X}_i}\left(S_{yx_i}^*-S_{yx_i}\right).$$

The optimum values of β_i and α_i are same as given in (8) and (13) respectively but now$;\, i=1,2,...,l$ for β and $;\, i=l+1,l+2,...,q$ for α. Then optimum MSE of t_{rera} is

$$MSE\left(t_{rera}\right)=S_{y_j}^2-\sum_{i=1}^{l}\left(S_{y_jx_i}^*-S_{y_jx_i}\right)^2\Big/\left(S_{x_i}^*+S_{x_i}\right)-\sum_{i=l+1}^{q}\left(S_{y_jx_i}^*-S_{y_jx_i}\right)^2\Big/\left(S_{x_i}^*+S_{x_i}\right)$$

or

$$MSE\left(t_{rera}\right)=S_{y_j}^2-\sum_{i=1}^{q}\left(S_{y_jx_i}^*-S_{y_jx_i}\right)^2\Big/\left(S_{x_i}^*+S_{x_i}\right),$$

it is same as for regression. The expression of Bias is

$$Bias\left(t_{rera}\right)=-\frac{1}{\bar{Y}}\sum_{i=l+1}^{q}\left(S_{y_jx_i}^*-S_{y_jx_i}\right)^2\left(S_{x_i}^*+S_{x_i}\right)^{-1}.$$

Now for MSE of $t_{rera(1)}$, again using Shahbaz and Hanif (2009) shrinkage theorem, we have

$$MSE\left(t_{rera(1)}\right)=\bar{Y}^2\left[MSE\left(t_{rera}\right)-\left\{Bias\left(t_{rera}\right)\right\}^2\right]\left[\bar{Y}^2+MSE\left(t_{rera}\right)+2\bar{Y}Bias\left(t_{rera}\right)\right]^{-1}$$

and

$$Bias\left(t_{rera(1)}\right)=\gamma_{opt}\left\{Bias\left(t_{rera}\right)+\bar{Y}\right\}-\bar{Y},$$

where $\gamma_{opt}=\left[\bar{Y}^2+\bar{Y}Bias(t_{rera})\right]\left[\bar{Y}^2+2\bar{Y}Bias(t_{rera})+MSE\left(t_{rera}\right)\right]^{-1}$.

The bivariate version of shrinkage regression-cum-ratio estimator can be written as

$$t_{rera(2)}=\begin{bmatrix}\gamma_1 & 0\\ 0 & \gamma_2\end{bmatrix}_{2\times2}\begin{bmatrix}t_{1rera}\\ t_{2rera}\end{bmatrix}_{2\times1}=\Lambda_3 t_{rera},$$

where $t_j=\left\{\bar{y}_{j_{ss(m)}}+\sum_{i=1}^{l}\beta_{ij}\left(\bar{x}_{i_{ss(k)}}-\bar{x}_{i_{ss(m)}}\right)\right\}\prod_{i=l+1}^{q}\left(\frac{\bar{x}_{i_{ss(k)}}}{\bar{x}_{i_{ss(m)}}}\right)^{\alpha_{ij}}$ for $j=1,2.$

Using previous approach, the expressions of MSE and Bias of $t_{rera(2)}$ can be obtained as

$$MSE\left(t_{rera}\right)=B\left(t_{rera}\right)B'\left(t_{rera}\right)+\Sigma_{t_{(2\times2)}}=\Sigma_{y'x_{(2\times1)}}\Sigma'_{y'x_{(1\times2)}}+\Sigma_{t_{(2\times2)}},$$

and

$$B\left(t_{rera}\right)=-E\left(\bar{D}_{y'_{(2\times l)}}\bar{D}_{x_{(l\times1)}}\right)=\Sigma_{y'x_{(2\times1)}} \text{ where}$$

$$\Sigma_t=\Sigma y_{(2\times2)}-\Sigma yx_{1(2\times q)}B'_{(q\times2)}+\Sigma yx_{2(2\times l)}\Delta'_{(l\times2)}-B_{(2\times q)}\Sigma x_1y_{(q\times2)}$$
$$+B_{(2\times q)}\Sigma x_{1(q\times q)}B'_{(q\times2)}+\Delta_{(2\times l)}\Sigma x_2y_{(l\times2)}+\Delta_{(2\times l)}\Sigma x_{2(l\times l)}\Delta'_{(l\times2)}$$

and the expression of $MSE\left(t_{rera(2)}\right)$ and $Bias\left(t_{rera(2)}\right)$ are

$$MSE\left(t_{rera(2)}\right)=\bar{Y}\bar{Y}'-\left\{\bar{Y}\bar{Y}'+\bar{Y}B'\left(t_{rera}\right)\right\}$$
$$\left\{\bar{Y}\bar{Y}'+MSE\left(t_{rera}\right)+2\bar{Y}B'\left(t_{rera}\right)\right\}^{-1}\left\{\bar{Y}\bar{Y}'+\bar{Y}B'\left(t_{rera}\right)\right\}'$$

and

$$Bias\left(t_{rera(2)}\right)=\left(\Lambda B\left(t_{rera}\right)+\left(\Lambda-I\right)\bar{Y}\right)\left(B'\left(t_{rera}\right)\Lambda+\bar{Y}'\left(\Lambda-I\right)\right),$$

where $\Lambda_3=\left\{\bar{Y}\bar{Y}'+\bar{Y}B'\left(t_{rera}\right)\right\}\left\{\bar{Y}\bar{Y}'+MSE\left(t_{rera}\right)+2\bar{Y}B'\left(t_{rera}\right)\right\}^{-1}$.

The multivariate version can be defined by replacing 2 by p in bivariate version.

Note: The bivariate regression-cum-ratio estimator reduced to bivariate regression estimator for $\alpha_{ij}=0$ and becomes bivariate ratio estimator for $\beta_{ij}=0$. Further if both α_{ij} and β_{ij} are zero then it reduces to bivariate version of mean per unit estimator.

4. EMPIRICAL STUDY

The performance of the suggested generalized shrinkage regression, ratio and regression-cum-ratio estimators has been observed using the data of district census report (1998), Punjab, Pakistan. This data is already used by Hanif, et al. (2009) for multivariate ratio estimators and univeriate ratio estimators under multi-phase sampling design. The empirical study is conducted for univariate cases which have more application as compared to multivariate estimators. The description of variables and detail of parameters is given in Appendix A. The suggested estimators are compared the estimators suggested by Hanif et al. (2009) and Ahmad and Hanif (2010) for no information case. As these estimators are suggested in multiphase design but the process of selecting the units was simple random sampling. Through this comparison we have the aim to observe the performance of the suggested estimators under the newly developed mixed sampling design (MPSS), further the efficiency of the estimators is observed due to shrinkage strategy. The pair wise comparison of phases is made to observe the performance of suggested estimators into different phases.

We considered five districts (Jhang, Faisalabad, Gujrat, Kasur, Sialkot) as natural populations from the above stated census report. We considered a variable of interest, denoted by Y and five auxiliary variables denoted by X's to for empirical study. As the expressions of MSE's depend on unknown population parameters, the required are provided in Table A-1 given in Appendix.

Table A-2, contains the MSE's of the suggested generalized shrinkage regression, ratio and regression-cum-ratio estimators under multiphase stratified systematic (MPSS) design. As the expressions of MSE's of all the three suggested estimators are same so the numerical results are not separately given. The percent relative efficiency (PRE) of suggested estimators is shown in Table A-3 in comparison with Hanif, et al. (2009) and Ahmad and Hanif (2010) ratio and regression estimators under multiphase sampling design for no information case. Further under multiphase design the results of univariate ratio estimator of Hanif et al. (2009) and univariate regression estimator of Ahmad and Hanif (2010) for no information case are same. These results clearly state that the performances of suggested generalized shrinkage regression, ratio and regression-cum-ratio estimators under MPSS design is much higher than multiphase design.

The pair-wise comparison also suggests that under multiphase design increase of phase can significantly decrease the performance where under MPSS design increase of phases not create much effect on the efficiency of estimators.

The pair-wise comparison of each phase was given in comparison tables. The comparison of phases for each estimator is almost same for each district. Critical review of Table A-2 shows that each proposed estimator provide small MSE's on different phase for each district. Such as in the district Jhang T_{15} give minimum MSE than all other pair wise combinations, similarly T_{13} perform better in district Faisalabad, T_{15} in district Gujrat, T_{12} in District Kasur and T_{23} was efficient in district Sialkot. This comparison clearly define that each district has its own variation and on the basis of their own variations the suitable phases varies for each district but the use of large number of phases does not create much effect as in district Jhang T_{15} and T_{45} variation is much closer and similarly observed in other districts. It is also noting point that performance of estimators under proposed sampling designs (MPSS) in these districts was much better than the estimators under multiphase simple random sampling design.

To observe the design and shrinkage effect we compute the results of MSE's in Table A-4 by ignoring shrinkage concept and in Table A-5 we find percent relative efficiency of suggested estimator. In this comparison it is again clearly shown that the performance of the suggested estimators under MPSS design is better than the estimators suggested by Hanif et al. (2009) and Ahmad and Hanif (2010) under multiphase design. Further we observe the shrinkage effect be comparing Table A-3 and Table A-5. This comparison also shows that shrinkage estimators performance is better than non-shrinkage estimators but both tables gives much better result than multiphase simple random sampling design.

5. SIMULATION STUDY

We used the census data that is already used in empirical study for simulation study after increasing the size by selecting a sample using simple random sampling with replacement. By the way the distribution of data remains the same. Also some random noise is added to the data values to get closer to the real data. The population sizes are 10000, 8000, 7000, 5000 and 7500 respectively. The detail of simulated population parameters is given in Table A-6. We generated 4000 independent samples from a magnified population and computed the estimates using pair-wise phases. The percent relative bias (PRB) is given in Table A-7. From Table A-7, column (1) contains the PRB of suggested generalized shrinkage regression estimator, column (2) contains for ratio estimator and column (3) contains for regression-cum-ratio estimator. Critically observing the results of this table we conclude that there is a slightly difference exist between these three suggested estimators. The PRB of column (3) regression-cum-ratio is much low than others, so we can say that this estimator's performance is better than the other two. These results also suggested that increase in phase does not decrease the estimator's performance.

Now we can say that the suggested generalized shrinkage regression, ratio and regression-cum-ratio estimators under MPSS design are much better than ratio or regression estimators under multiphase simple random sampling design.

6. CONCLUSION

In conclusion, multi-phase stratified systematic sampling design is a better choice than multi-phase simple random sampling design while selecting a sample from district census reports of Punjab Pakistan. As it is observed that for different districts, different pair of phases provides better performance. So there is further need to investigate in general that for what type of population which pair is preferable. The shrinkage estimators are better than simple ones in term of efficiency but not suitable if the objective is to reduce the bias.

This paper fills the gap of univariate and multivariate shrinkage estimators under multi-phase stratified systematic sampling design. The additional advantage is the use of multi-auxiliary variables to get benefit of practically available information on several auxiliary variables in term of gain in efficiency and reduction in bias. The construction of regression-cum-ratio estimator has an extra advantage over regression and ratio estimator that it can be applied when some auxiliary variable is suitable for regression and some are suitable for ratio method. As the estimators are general in nature, the auxiliary variables that could be negatively correlated with response variable can also be accommodated.

The practitioners can use the developed estimators to estimate the mean of single or several correlated outcome variables in the presence of several auxiliary variables with flexible set of estimation methods (regression, ratio, regression-cum-ratio) under a more practically useable sampling design i.e. multi-phase stratified systematic sampling design. The methodology developed in this paper is equally important and useful for any other practical situation like this and can be tried for different settings because the suggested sampling design and estimation methods have a large scope of application due to their general settings.

Conflict of Interest:

Authors have no conflicts of interest to disclose.

Funding Statement:

This research did not receive any funding.

REFERENCES

1. Abid, M., Nazir, H.Z., Riaz, M., Lin, Z. and Tahir, H.M. (2018). Improved ratio estimators using some robust measures. *Hacettepe Journal of Mathematics and Statistics*, 47(5), 1375-93.
2. Ahmad, Z. and Hanif, M. (2016). A note of multivariate shrinkage estimators. *Journal of ISOSS*, 2(1), 33-40.
3. Ahmad, Z., Hanif, M. and Ahmad, M. (2009). Generalized Regression-cum-Ratio Estimators for two phase sampling using multi-auxiliary variables. *Pakistan Journal of Statistics*, 25(2), 93-106.
4. Hanif, M., Ahmad, Z. and Ahmad, M. (2009). Generalized Multivariate Ratio Estimators for Multi-Phase Sampling using Multi-Auxiliary Variables. *Pakistan Journal of Statistics*, 25(4), 615-629.
5. Ahmad, Z., Hanif, M. and Ahmad, M. (2010a). Generalized Multi-Phase Multivariate Ratio Estimators for Partial Information Case using Multi-Auxiliary Variables. *Communications of the Korean Statistical Society*, 17(5), 625-637.
6. Ahmad, Z., Hanif, M. and Ahmad, M. (2010b). Generalized Regression-in Regression Estimators for Two-Phase Sampling using Multi-Auxiliary Variables. *Journal of Statistical Theory and Applications*, 9(3), 443-457.
7. Ahmad, Z., Hussin, A.G. and Hanif, M. (2010). Generalized multivariate regression estimators for multiphase sampling using multi-Auxiliary variables. *Pakistan Journal of Statistics*, 26(4), 569-583.
8. Ahmad, Z. and Hanif, M. (2010). Generalized Multi-Phase Multivariate Regression Estimators for Partial Information Case using Multi-Auxiliary Variables. *World Applied Sciences Journal*, 10(3), 370-379.
9. Ahmad, Z., Hanif, M. and Ahmad, M. (2013). Generalized Regression Estimators under Two Phase Sampling for Partial Information Case. *Pakistan Journal of Statistics*, 29(1), 99-110.
10. Ahmed, M.S. (2003). General Chain estimators under multi-phase sampling. *J. Applied Statist. Sci.*, 12(4), 243-250.
11. Bowley, A.L. (1926). Measurements of precision attained in sampling. *Bull. Inst. Inte. Statist.*, 22, 1-62.
12. Cochran, W.G. (1977). *Sampling Techniques.* John Wiley and Sons, New York.
13. Das, A.K. and Tripathi, T.P. (1978). Use of auxiliary information in estimating the finite population variance. *Sankhya C*, 40, 139-148.
14. Hanif, M., Ahmad, Z. and Ahmad, M. (2009). Generalized Multivariate Ratio Estimator using Multi-Auxiliary Variables for Multi-Phase Sampling. *Pakistan Journal of Statistics*, 25 (4), 615-629.
15. Javed, M., Ahmad, Z. and Hanif, M. (2014). Replication Variance Estimation under Two-phase Sampling in the Presence of Non-response. *Statistica*, *74*(3), 247.
16. Kadilar, C. and Cingi, H. (2006). Improvement in variance estimation using auxiliary information. *Hacettepe Journal of Mathematics and Statistics*, 35(1), 111-115.
17. Iqbal, K., Moeen, M., Ali, A. and Iqbal, A. (2020). Mixture regression cum ratio estimators of population mean under stratified random sampling. *Journal of Statistical Computation and Simulation*, 90(5), 854-868.

18. Khare, B.B. and Srivastava, S.R. (1981). A generalized regression ratio estimator for the population mean using two auxiliary variables. *The Aligarh Journal of Statistics*, 1(1), 43-51.
19. Mukhopadhyay, P. (2000). *Theory and methods of survey sampling*, Prentice Hall of India Pvt. Ltd., New Delhi.
20. Neyman, J. (1934). On the two different aspects of representative method: The method of stratified sampling and the method of purposive selection. *J. Roy. Statist. Soc.*, 97, 558-606.
21. Noor-Ul-Amin, M., Shahbaz, M.Q. and Kadılar, C. (2016). Ratio estimators for population mean using robust regression in double sampling. *Gazi University Journal of Science*, 29(4), 793-8.
22. Olkin, I. (1958). Multivariate ratio estimation for finite populations, *Biometrika*, 45,154-165.
23. Pardhan, B.K. (2005). A chain regression estimator in two phase sampling using multi-auxiliary information, *Bull. Malays. Math. Sci. Soc.*, 28(1), 81-86.
24. Raj, D. (1965). On a method of using multi-auxiliary information in sample surveys. *J. Amer. Statist. Assoc.*, 60(309), 270-277.
25. Sahoo, J. and Sahoo, L.N. (1993). A class of estimators in two-phase sampling using two auxiliary variables. *J. Ind. Statist. Assoc.*, 31, 107-114.
26. Sahoo, J., Sahoo, L.N. and Mohanty, S. (1993). A regression approach to estimation in two-phase sampling using two auxiliary variables. *Current Sciences*, 65(1), 73-75.
27. Samiuddin, M. and Hanif, M. (2007). Estimation of population mean in single and two phase sampling with or without additional information. *Pakistan Journal of Statistics*, 23(2), 99-118.
28. Sen, A.R. (1972). Successive Sampling with $p(p \geq 1)$ auxiliary variables. *Ann. Math. Statist.*, 43(6), 2031-2034.
29. Shahbaz, M.Q. and Hanif, M. (2009). A General Shrinkage Estimator in Survey Sampling. *World Applied Sciences Journal*, 7(5), 593-596.
30. Srivastava, S.K. (1971). A generalized estimator for the mean of a finite population using multi-auxiliary information. *J. Amer. Statist. Assoc.*, 66, 404-407.
31. Srivastava, S.K. and Jhajj, H.S. (1983). A class of estimators of the population mean using multi-auxiliary information, *Cal. Stat. Assoc. Bull.*, 32, 47-56.
32. Sukhatme, P.V., Sukhatme, B.V., Sukhatme, S. and Ashok, C. (1984). *Sampling Theory of Surveys with Applications*. University Press, Ames, Iowa, USA.
33. Tikkiwal, G.C. and Ghiya, A. (2004). A generalized class of composite estimators with application to crop acreage estimation for small domains. *Statistics in Transition*, 1(6), 697-711.
34. Upadhyaya, L.N. and Singh, H.P. (1983). Use of auxiliary information in the estimation of population variance. *Mathematical Forum*, 4(2), 33-36.
35. Zaman, T. and Bulut, H. (2019). Modified ratio estimators using robust regression methods. *Communications in Statistics-Theory and Methods*, 48(8), 2039-2048.
36. Zaman, T. and Bulut, H. (2020). Modified regression estimators using robust regression methods and covariance matrices in stratified random sampling. *Communications in Statistics-Theory and Methods*, 49(14), 3407-3420.

APPENDIX

Table-A
Variables Description

Y	Population of currently married
X_1	Population of both sexes
X_2	Population of primary but below matric
X_3	Population of matric and above
X_4	Population of 18 years old and above
X_5	Population of women 15-49 years old

Source: Districts Population Census Report (1998), Pakistan.

Table A-1
Parameters of Populations for Calculating the MSE's of Suggested Estimators

	Districts				
	Jhang	**Faisalabad**	**Gujrat**	**Kasur**	**Sailkot**
N	368	283	204	181	269
n_1	184	142	102	91	135
n_2	92	71	51	45	67
n_3	46	35	26	23	34
n_4	23	18	13	11	17
n_5	12	9	6	6	8
Y	860.110	1511.260	1101.280	1393.200	1058.740
σ_y	511.908	788.380	533.041	767.636	685.019
σ_{x1}	5626.450	5426.030	3507.160	5515.420	4787.250
σ_{x2}	455.060	1677.920	940.480	1095.690	1172.710
σ_{x3}	170.670	525.670	381.690	357.890	603.220
σ_{x4}	2455.170	6289.710	8139.680	2719.210	2461.590
σ_{x5}	1064.480	1482.170	830.010	1355.640	1151.320
ρ_{yx1}	0.428	0.943	0.995	0.998	0.999
ρ_{yx2}	0.912	0.927	0.941	0.758	0.983
ρ_{yx3}	0.659	0.599	0.764	0.879	0.931
ρ_{yx4}	0.484	0.731	0.490	0.989	0.996
ρ_{yx5}	0.425	0.501	0.996	0.799	0.939
ρ_{x1x2}	0.416	0.641	0.954	0.764	0.983
ρ_{x1x3}	0.421	0.782	0.796	0.889	0.931
ρ_{x1x4}	0.317	0.513	0.509	0.993	0.997
ρ_{x1x5}	0.275	0.819	0.996	0.802	0.939
ρ_{x2x3}	0.824	0.708	0.892	0.798	0.959
ρ_{x2x4}	0.475	0.359	0.5	0.764	0.985
ρ_{x2x5}	0.432	0.559	0.958	0.614	0.928
ρ_{x3x4}	0.59	0.543	0.42	0.896	0.939
ρ_{x3x5}	0.464	0.685	0.797	0.719	0.887
ρ_{x4x5}	0.325	0.436	0.505	0.797	0.938

Table A-2
MSE's of Suggested Shrinkage Regression, Ratio and Regression-Cum-Ratio Estimators for Pair Wise of Phases

	District				
	Jhang	**Faisalabad**	**Gujrat**	**Kasur**	**Sailkot**
T12 (k=1, m=2)	53980.170	125593.900	58108.320	114995.100	93676.220
T13 (k=1, m=3)	56768.130	114878.800	56383.240	120830.900	96113.280
T14 (k=1, m=4)	45355.040	133748.200	60255.550	133630.800	98830.870
T15 (k=1, m=5)	44880.140	128877.900	51229.140	122358.400	103141.500
T23 (k=2, m=3)	57629.550	125294.500	62426.840	125499.700	86075.820
T24 (k=2, m=4)	55137.720	119580.300	63544.590	120313.500	93605.820
T25 (k=2, m=5)	51090.440	123834.200	66374.680	131205.700	90619.570
T34 (k=3, m=4)	57288.820	130866.000	58398.020	126581.500	95844.660
T35 (k=3, m=5)	49872.370	128968.400	57283.290	135940.400	98105.820
T45 (k=4, m=5)	47124.140	133624.500	64186.620	126678.500	100722.600

Table A-3
Relative Efficiency of Estimator Proposed by Hanif, et al. (2009) and Ahmad and Hanif (2010) Over Suggested Shrinkage Estimators

	District				
	Jhang	**Faisalabad**	**Gujrat**	**Kasur**	**Sailkot**
T12 (k=1, m=2)	393.25547	169.20844	164.11278	176.60842	10.20032
T13 (k=1, m=3)	2154.80341	1036.68036	553.50054	746.33369	43.14108
T14 (k=1, m=4)	15922.70155	5028.62476	1830.53154	2872.26451	175.86232
T15 (k=1, m=5)	102781.45492	31780.95647	8220.70585	13233.58522	701.88002
T23 (k=2, m=3)	6199.71287	2216.79595	1799.24319	1734.07603	146.58641
T24 (k=2, m=4)	29038.44196	11087.42366	5308.05107	7458.62740	515.91418
T25 (k=2, m=5)	156509.04490	54459.69915	16951.55015	28051.39677	2093.77165
T34 (k=3, m=4)	67847.06339	21056.12497	18326.27801	15739.05971	1311.38264
T35 (k=3, m=5)	316996.34026	95314.35466	54019.99110	58400.49992	4642.93682
T45 (k=4, m=5)	552776135.54327	182600.57965	144998.67516	134257.56533	11070.16487

Table A-4
MSE's of Suggested Regression, Ratio and Regression-Cum-Ratio Estimators for Pair Wise of Phases without Shrinkage

	District				
	Jhang	**Faisalabad**	**Gujrat**	**Kasur**	**Sailkot**
T12 (k=1, m=2)	59345.270	133237.350	60295.000	12115.350	94887.520
T13 (k=1, m=3)	62738.780	125888.280	59057.320	128779.000	969886.370
T14 (k=1, m=4)	52552.240	151211.130	62555.380	140556.720	99520.660
T15 (k=1, m=5)	47320.440	138787.590	52989.570	128967.560	113896.900
T23 (k=2, m=3)	61423.130	129998.030	65086.650	134466.500	86999.990
T24 (k=2, m=4)	55137.720	132235.530	64934.480	129873.300	94535.410
T25 (k=2, m=5)	54115.380	133842.270	68137.080	138893.100	92061.070
T34 (k=3, m=4)	60509.130	139722.440	61535.820	137586.500	97154.000
T35 (k=3, m=5)	54027.570	135663.760	59476.760	140710.900	99045.660
T45 (k=4, m=5)	51118.250	141225.330	66888.070	129686.700	116788.000

Table A-5: Percent Relative Efficiency of Estimator Proposed by Hanif, et al. (2009) and Ahmad and Hanif (2010) Over without Shrinkage Suggested Estimators

	District				
	Jhang	**Faisalabad**	**Gujrat**	**Kasur**	**Sailkot**
T12 (k=1, m=2)	357.703268	159.501431	158.161008	1676.311704	10.070102
T13 (k=1, m=3)	1949.737631	946.018136	528.438371	700.270782	4.275172
T14 (k=1, m=4)	13742.035843	4447.883631	1763.232595	2730.733927	174.643396
T15 (k=1, m=5)	97481.048063	29511.737541	7947.595933	12555.407841	635.600776
T23 (k=2, m=3)	5816.809775	2136.588839	1725.715903	1618.440444	145.029270
T24 (k=2, m=4)	29038.441960	10026.332915	5194.434898	6909.607810	510.841070
T25 (k=2, m=5)	147760.506680	50387.469340	16513.089745	26498.819229	2060.987201
T34 (k=3, m=4)	64236.226864	19721.462422	17391.794730	14480.154565	1293.709194
T35 (k=3, m=5)	292616.506165	90610.342932	52027.763718	56420.556751	4598.880193
T45 (k=4, m=5)	509585128.598886	172772.909477	139142.523667	131143.339980	9547.348940

Table A-6: Parameters of Simulated Populations for PRB of Suggested Estimators

	Districts				
	Jhang	**Faisalabad**	**Gujrat**	**Kasur**	**Sailkot**
N	10000	8000	7000	5000	7500
n_1	5000	4000	3500	2500	3750
n_2	2500	2000	1750	1250	1875
n_3	1250	1000	875	625	938
n_4	625	500	438	313	469
n_5	313	250	219	156	234
Y	865.376	1507.438	1096.517	1387.577	1062.090
σ_y	493.397	776.801	518.077	765.396	672.087
σ_{x1}	5744.959	5387.612	3466.628	5602.730	4692.195
σ_{x2}	465.110	1697.969	921.828	1037.277	1153.150
σ_{x3}	170.721	519.998	371.974	362.581	583.563
σ_{x4}	2472.054	6536.105	7206.235	2761.372	2416.695
σ_{x5}	1136.245	1456.886	838.735	1370.214	1125.382
ρ_{yx1}	0.523	0.891	0.938	0.980	0.963
ρ_{yx2}	0.861	0.912	0.961	0.716	0.970
ρ_{yx3}	0.621	0.583	0.752	0.832	0.936
ρ_{yx4}	0.514	0.755	0.522	0.951	0.986
ρ_{yx5}	0.403	0.505	0.977	0.732	0.951
ρ_{x1x2}	0.422	0.660	0.953	0.770	0.972
ρ_{x1x3}	0.34	0.581	0.701	0.815	0.94
ρ_{x1x4}	0.361	0.388	0.415	0.922	0.91
ρ_{x1x5}	0.211	0.757	0.891	0.813	0.926
ρ_{x2x3}	0.732	0.711	0.821	0.811	0.915
ρ_{x2x4}	0.402	0.291	0.443	0.755	0.93
ρ_{x2x5}	0.44	0.521	0.901	0.57	0.865
ρ_{x3x4}	0.6	0.535	0.362	0.785	0.909
ρ_{x3x5}	0.455	0.683	0.703	0.725	0.818
ρ_{x4x5}	0.331	0.417	0.462	0.8	0.877

Table A-7
Percent Relative Bias of suggested Regression, Ratio and Regression-cum-Ratio estimators of 4000 samples

		Districts				
		Jhang	**Faisalabad**	**Gujrat**	**Kasur**	**Sailkot**
T12 (k=1,m=2)	$t_{reg(1)}$	167.9300	175.8981	159.1297	169.7127	180.5292
	$t_{ra(1)}$	167.0400	174.0779	156.7814	166.6547	177.4033
	$t_{rera(1)}$	166.6200	175.1442	157.7327	168.3716	179.1042
T13 (k=1,m=3)	$t_{reg(1)}$	177.1776	172.2111	147.7502	163.2177	184.5097
	$t_{ra(1)}$	174.1297	166.6433	142.0024	155.3183	174.7512
	$t_{rera(1)}$	176.1551	170.0643	144.3697	159.6254	180.0264
T14 (k=1,m=4)	$t_{reg(1)}$	172.6916	153.6655	123.0060	165.3200	162.8270
	$t_{ra(1)}$	166.1666	143.7942	114.0342	157.1100	145.2866
	$t_{rera(1)}$	170.4890	150.3022	118.4891	156.2100	155.2914
T15 (k=1,m=5)	$t_{reg(1)}$	152.8424	117.1153	114.2309	53.7321	63.2727
	$t_{ra(1)}$	144.4182	107.0191	111.1017	53.0017	57.1225
	$t_{rera(1)}$	150.7813	115.3850	110.2990	52.5602	54.0105
T23 (k=2,m=3)	$t_{reg(1)}$	43.9065	47.3610	34.4209	40.0596	57.4283
	$t_{ra(1)}$	44.2430	49.0363	39.3712	41.0622	58.5163
	$t_{rera(1)}$	44.1296	47.5980	34.5835	41.0575	58.0673
T24 (k=2,m=4)	$t_{reg(1)}$	53.4929	43.9243	31.3449	40.4513	67.3999
	$t_{ra(1)}$	54.3252	48.2221	45.5304	39.5555	70.3054
	$t_{rera(1)}$	54.1084	44.5815	31.9586	39.0291	69.0919
T25 (k=2,m=5)	$t_{reg(1)}$	51.3685	36.8603	28.0014	30.1100	43.5200
	$t_{ra(1)}$	53.9050	46.8238	20.2516	25.8735	48.1476
	$t_{rera(1)}$	52.9058	38.8354	20.0510	24.1517	38.1353
T34 (k=3,m=4)	$t_{reg(1)}$	-20.7787	-25.2494	-10.9290	-3.0002	-9.5162
	$t_{ra(1)}$	-21.3033	-23.8607	-9.8911	-4.8876	-8.9901
	$t_{rera(1)}$	-20.9413	-25.7067	-9.0010	-4.1303	-9.0738
T35 (k=3,m=5)	$t_{reg(1)}$	-10.1207	-17.9397	-23.2525	-29.5263	-9.1083
	$t_{ra(1)}$	-10.4598	-13.2141	-18.1489	-21.8877	-9.5176
	$t_{rera(1)}$	-10.0248	-18.5759	-16.0371	-12.0000	-8.8770