Journal of King Saud University – Science 30 (2018) 223–228

Contents lists available at ScienceDirect

Journal of King Saud University – Science

journal homepage: www.sciencedirect.com

Original article

A multivariate regression-cum-exponential estimator for population variance vector in two phase sampling

Amber Asghar [a,c], Aamir Sanaullah [b,*], Muhammad Hanif [c]

[a] Virtual University of Pakistan, Lahore, Pakistan
[b] COMSATS Institute of Information Technology Lahore, Pakistan
[c] NCBA&E Lahore, Pakistan

ARTICLE INFO

Article history:
Received 11 October 2016
Accepted 28 January 2017
Available online 9 February 2017

Keywords:
Multivariate estimator
Multi-auxiliary variables
Two-phase sampling
Regression estimator
Variance-covariance matrix
Simulation

ABSTRACT

In this study we have proposed a multivariate regression-cum-exponential type estimator for estimating a vector of population variance. In the present study, unknown population variance vector estimation has been discussed using multi-auxiliary variables in two-phase sampling and different cases have also been derived. A comparison between existing and the proposed multivariate, bivariate and univariate estimators has been prepared with the help of a real data for estimating population variance. A simulation study for multivariate estimator using multi-auxiliary variables has also been carried out to demonstrate the performance of the estimators.

1. Introduction

Auxiliary information plays a vital role in illustrating conclusion about the parameters of population for the characteristics under study. It is used to enhance estimation of population parameters of the main variable under study. At the stage of manipulating as well as estimation, auxiliary information is used for improved estimation. Sometimes auxiliary information is known in prior of a survey and sometimes it is not known in advanced. There are many examples in survey sampling where auxiliary information is known in advance; number of banks in a city, number of employees, educational status, number of educated male and females in a city etc.

Graunt (1662) was the first who estimated the population of England using auxiliary information. Olkin (1958) suggested ratio estimator based on multi-auxiliary variables for multivariate case. John (1969) provided multivariate ratio and product type estimators for estimating the population means. Further comprehensive contribution of multivariate ratio and regression estimators using multi-auxiliary variables were taken up by Ahmad and Hanif (2010) for estimating population mean. Isaki (1983) proposed ratio and regression type estimators for estimating the population variance. Cebrian and Garcia (1997) worked on variance estimation by using auxiliary variables. Following Isaki (1983), Singh et al. (2009) proposed exponential estimator for estimating population variance and Abu-Dayyeh and Ahmed (2005) provided some multi-variate ratio and regression-type estimators in two-phase sampling and studied some properties of the proposed estimator through simulation study using real data. Kadilar and Cingi (2006) suggested the regression type estimator for estimating variance using known population variance of the auxiliary variable. Many other authors including Upadhyaya and Singh (2006), Ahmed et al. (2000), Yadav and Kadilar (2013), Singh and Solanki (2013), Ahmad et al. (2016) and Singh and Pal (2016) etc. have worked on variance estimation for using population variance of the auxiliary information. Asghar et al. (2014) provided some exponential-type estimator for variance using population means of multi-auxiliary auxiliary variables.

In fact, there is no significant work on variance estimation in the literature for estimating finite population variance under two-phase sampling using multivariate multi-auxiliary variables. Therefore to fulfill this gap we proposed a multivariate regression type exponential estimator for estimating the finite population variance using multi-auxiliary variables under two-phase sampling for full information case.

* Corresponding author.
E-mail addresses: zukhruf10@gmail.com (A. Asghar), chaamirsanaullah@yahoo.com (A. Sanaullah), drmianhanif@gmail.com (M. Hanif)

Peer review under responsibility of King Saud University.

Production and hosting by Elsevier

http://dx.doi.org/10.1016/j.jksus.2017.01.010
1018-3647/

Let $S^2_{y_{(m\times m)}}$ be the population variance and its usual unbiased estimator is defined as,

$$t_0 = [\, t_{01} \quad t_{02} \quad \ldots \quad t_{0j} \quad \ldots \quad t_{0m} \,],$$

$$t_{0j} = s^2_{y_{j(2)}}, \qquad j = 1,2,\ldots,m$$

The variance of unbiased estimator is defined as,

$$Var(t_0) = \gamma_2 S'_{(m\times 1)} S_{(1\times m)} \sum_{y_{(m\times m)}}$$

where $S_{(1\times m)} = [s^2_{y_1} \ldots s^2_{y_j} \ldots s^2_{y_m}]$ and $\gamma_2 = 1/n_2$. (1)

We develop Isaki's (1983) uni-variate regression estimator into a multivariate regression estimator under two- phase i.e.

$$t_{reg} = [\, t_{reg1} \quad t_{reg2} \quad \ldots \quad t_{regj} \quad \ldots \quad t_{regm} \,],$$

where,

$$t_{regj} = s^2_{y_{j(2)}} + \sum_{k=1}^{n} \alpha_k (S^2_{x_k} - s^2_{x_{(1)k}}), \qquad j = 1,2,\ldots,m \tag{2}$$

The variance covariance matrix of t_{reg} is,

$$\sum_{treg_{(m\times m)}} = S'S\left[\gamma_2 \sum_{y_{(m\times m)}} - \gamma_1 \sum_{yx_{(m\times n)}} \sum_{x_{(n\times n)}}^{-1} \sum_{xy_{(n\times m)}}\right]. \quad \text{where}$$

$S_{(1xm)} = S$ and $\gamma_1 = 1/n_1$. (3)

We modify Shabbir and Gupta's (2015) estimator into a multivariate regression-type estimator for estimating a vector of population variance using populations mean of auxiliary information under two- phase as,

$$t_{rg} = [\, t_{rg1} \quad t_{rg2} \quad \ldots \quad t_{rgj} \quad \ldots \quad t_{rgm} \,],$$

where,

$$t_{rgj} = s^2_{y_{j(2)}} + \sum_{k=1}^{n} \beta_k (\bar{V}_k - \bar{v}_{(1)k}), \qquad j = 1,2,\ldots,m \tag{4}$$

The variance covariance matrix of t_{rg} is,

$$\sum_{trg_{(m\times m)}} = S'S\left[\gamma_2 \sum_{y_{(m\times m)}} - \gamma_1 \sum_{yv_{(m\times n)}} \sum_{v_{(n\times n)}}^{-1} \sum_{vy_{(n\times m)}}\right]. \tag{5}$$

We modify Asghar et al. (2014) into a multivariate exponential ratio type estimator for estimating for estimating a vector of population variances using multi-auxiliary variables as,

$$t_a = [\, t_{a1} \quad t_{a2} \quad \ldots \quad t_{aj} \quad \ldots \quad t_{am} \,],$$

where,

$$t_{aj} = s^2_{y_{j(2)}} \exp \sum_{k=1}^{n} \left[\frac{\bar{V}_k - \bar{v}_{(1)k}}{\bar{V}_k + (a_k - 1)\bar{v}_{(1)k}}\right], \qquad j = 1,2,\ldots,m \tag{6}$$

The variance covariance matrix of t_a is,

$$\sum_{t_{a(m\times m)}} = S'S\left[\gamma_2 \sum_{y_{(m\times m)}} - \gamma_1 \sum_{yv_{(m\times n)}} \sum_{v_{(n\times n)}}^{-1} \sum_{vy_{(n\times m)}}\right]. \tag{7}$$

The presentation of the paper is as follows; Section 2 is based on some useful results for multivariate under two-phase sampling design. Section 3 is based on the derivation of our proposed estimator. Whereas in Sections 4 and 5 numerical study with real data and simulated study with simulated data are discussed respectively. Finally, conclusion and discussions are presented in Section 6.

2. Some useful results under two-phase sampling

We consider a finite population with U (<∞) identifiable units. Let Y be the variable under study taking value y_j where $j = 1,2,\ldots,m$. Let $\bar{X}_1, \bar{X}_2, \ldots, \bar{X}_n$ be the population means of auxiliary variables and $S^2_{yj(2)}$ be the population variance of study variable. Further let S^2_x an S^2_y be the population variances and $C^2_{x_i}$, C^2_y be the coefficient of variation and ρ_{yx} denotes the correlation coefficient between study and auxiliary variables.

Under two-phase sampling design, $x_{(1)1}, x_{(1)2}, \ldots, x_{(1)n}$ is observed by n_1 and $y_{(2)1}, y_{(2)2}, \ldots, y_{(2)m}$ is observed by n_2.

In order to derive the expressions of mean square error, let $\gamma_2 = \frac{1}{n_2}$ and $\gamma_1 = \frac{1}{n_1}$ be the sampling fractions and $s^2_{y_{j(2)}} = S^2_{y_{j(2)}}(1 + \varepsilon_{y_{j(2)}})$, $s^2_{x(1)k} = S^2_{x(1)k}(1 + \varepsilon_{x_{(1)k}})$, $\bar{v}_{(1)k} = \bar{V}_k(1 + \bar{e}_{v_{(1)k}})$, where $\varepsilon_{y_{j(2)}}$, $\varepsilon_{x_{(1)k}}$ and $\bar{e}_{x_{(1)k}}$ be the sampling errors and $E(\varepsilon_{y_{j(2)}}) = E(\varepsilon_{x_{(1)k}}) = E(\bar{e}_{v_{(1)k}}) = 0$.

For multivariate estimators, we use following expectation results for the derivation of variance covariance matrix expressions,

Let

$\Delta_y = [\, \varepsilon_{y_{(2)1}} \quad \varepsilon_{y_{(2)2}} \quad . \quad . \quad \varepsilon_{y_{(2)m}} \,]$, $\Delta_x = [\, \varepsilon_{x_{(1)1}} \quad \varepsilon_{x_{(x)2}} \quad . \quad . \quad \varepsilon_{x_{(1)n}} \,]$,

$\bar{D}_v = [\, e_{v_{(1)1}} \quad e_{v_{(1)2}} \quad . \quad . \quad e_{v(1)n} \,]$

$$E_1E_{2/1}(\Delta'_x\Delta_x) = \gamma_1 \sum_{x_{(n\times n)}}, E_1E_{2/1}(\Delta'_y\Delta_y) = \gamma_2 \sum_{y_{(m\times m)}},$$

$$E_1E_{2/1}(\Delta'_x\Delta_y) = \gamma_1 \sum_{xy_{(n\times m)}}, E_1E_{2/1}(\bar{D}'_v\bar{D}_v) = \gamma_1 \sum_{v_{(n\times n)}}$$

$$E_1E_{2/1}(\Delta'_y\bar{D}_v) = \gamma_1 \sum_{yv_{(m\times n)}}, E_1E_{2/1}(\bar{D}'_v\Delta_y) = \gamma_1 \sum_{vy_{(n\times m)}}$$

3. Proposed generliazed regression-cum-exponential estimator

A multivariate regression-cum-exponential type estimator is proposed for full information case using multi-auxiliary variable in two-phase sampling design. We propose following multivariate estimator for population variance vector,

$$t_s = [t_{sj}]_{(1xm)},$$

$$\text{where}, t_{sj} = \left(s^2_{y_{l(2)}} + \sum_{k=1}^{n} \omega_{kj}(S^2_{x_k} - s^2_{x_{(1)k}})\right) \exp \sum_{k=1}^{n} \left[\frac{S^2_{x_k} - s^2_{x_{(1)k}}}{S^2_{x_k} + s^2_{x_{(1)k}}}\right];$$

$$j = 1,2,\ldots,m \tag{8}$$

3.1. Derivation of the mean square error

The Mean Square Error (MSE) is derived by using $s^2_{y_{j(2)}} = S^2_{y_j}(1 + \varepsilon_{y_{j(2)}})$, $s^2_{x_{k_{(1)}}} = S^2_{x_k}(1 + \varepsilon_{k(1)})$, where $\varepsilon_{y_{(2)j}}$ and $\varepsilon_{x_{(1)k}}$ are the sampling errors,

$$t_s = [t_{sj}]_{(1\times m)}, t_{sj} = \left(s^2_{y_{j(2)}} + \sum_{k=1}^{n} \omega_{kj}(S^2_{x_k} - s^2_{x_{(1)k}})\right) \exp \sum_{k=1}^{n} \left[\frac{S^2_{x_k} - s^2_{x_{(1)k}}}{S^2_{x_k} + s^2_{x_{(1)k}}}\right], \tag{9}$$

Consider,

$$t_{sj} = \left[S^2_{y_{j(2)}}(1 + \varepsilon_{y_{j(2)}}) + \sum_{k=1}^{n} \omega_{kj}\{S^2_{x_k} - S^2_{x_k}(1 + \varepsilon_{x_{(1)k}})\}\right] \times \exp \sum_{k=1}^{n} \left[\frac{S^2_{x_k} - S^2_{x_k}(1 + \varepsilon_{x_{(1)k}})}{S^2_{x_k} + S^2_{x_k}(1 + \varepsilon_{x_{(1)_k}})}\right], \tag{10}$$

$$t_{sj} = \left((S^2_{y_{j(2)}} + S^2_{y_{j(2)}}\, \varepsilon_{y_{j(2)}}) - \sum_{k=1}^{n} \omega_{kj} S^2_{x_k}\, \varepsilon_{x_{(1)k}} \right) \exp \sum_{k=1}^{n} \frac{-\varepsilon_{x_{(1)k}}}{2} \left(1 + \frac{\varepsilon_{x_{(1)k}}}{2} \right)^{-1}, \quad (11)$$

or

$$t_{sj} = \left((S^2_{y_{j(2)}} + S^2_{y_{j(2)}}\, \varepsilon_{y_{j(2)}}) - \sum_{k=1}^{n} \omega_{kj} S^2_{x_k}\, \varepsilon_{x_{(1)k}} \right) \times \left[1 + \sum_{k=1}^{n} \left\{ \frac{-\varepsilon_{x_{(1)k}}}{2} \left(1 - \frac{\varepsilon_{x_{(1)k}}}{2} + \frac{\varepsilon^2_{x_{(1)k}}}{4} \pm \ldots \right) \right\} \right], \quad (12)$$

Now retaining terms up to order one, we have,

$$t_{sj} = \left((S^2_{y_{j(2)}} + S^2_{y_{j(2)}}\, \varepsilon_{y_{j(2)}}) - \sum_{k=1}^{n} \omega_{kj} S^2_{x_k}\, \varepsilon_{x_{(1)k}} \right) \left(1 - \sum_{k=1}^{n} \frac{\varepsilon_{x_{(1)k}}}{2} \right) \quad (13)$$

After simplification we have,

$$t_{sj} = \left(S^2_{y_{j(2)}} + S^2_{y_{j(2)}}\, \varepsilon_{y_{(2)l}} - \sum_{k=1}^{n} \omega_{kj} S^2_{x_k}\, \varepsilon_{x_{(1)k}} - S^2_{y_{j(2)}} \sum_{k=1}^{n} \frac{\varepsilon_{x_{(1)k}}}{2} \right), \quad (14)$$

or

$$t_{sj} = \left[S^2_{y_{j(2)}} + S^2_{y_{j(2)}}\, \varepsilon_{y_{j(2)}} - \frac{1}{2} \sum_{k=1}^{n} \left(2\omega_{kj} \frac{S^2_{x_k}}{S^2_{y_{j(2)}}} + 1 \right) \varepsilon_{x_{(1)k}} \right], \quad (15)$$

or

$$t_{sj} - S^2_{y_{j(2)}} = S^2_{y_{j(2)}} \left(\varepsilon_{y_{j(2)}} - \frac{1}{2} \sum_{k=1}^{n} (2\omega_{kj}\phi_{kj} + 1) \varepsilon_{x_{(1)k}} \right) \text{ where } \phi_{kj} = \frac{S^2_{x_k}}{S^2_{y_{j(2)}}}. \quad (16)$$

$$t_s = [t_{sj}]_{(1\times m)} = \left[S^2_{y_{j(2)}} \left(\varepsilon_{y_{j(2)}} - \frac{1}{2} \sum_{k=1}^{n} (2\omega_{kj}\phi_{kj} + 1) \varepsilon_{x_{(1)k}} \right) \right]_{(1\times m)}$$

$$j = 1, 2, \ldots, m \quad (17)$$

For variance covariance matrix we proceed as,

$$\sum_{t_{s(m\times m)}} = E_1 E_{2/1} (\mathbf{t}_{s_{(1\times m)}} - S_{(1\times m)})' (\mathbf{t}_{s_{(1\times m)}} - S_{(1\times m)})$$

$$= S^4_y E_1 E_{2/1} \left(\varepsilon_{y_{(1\times m)}} - \frac{1}{2} \varepsilon_{x_{(1\times n)}} \Omega_{(n\times m)} \right)' \left(\varepsilon_{y_{(1\times m)}} - \frac{1}{2} \varepsilon_{x_{(1\times n)}} \Omega_{(n\times m)} \right), \quad (18)$$

where,

$$\Omega_{(n\times m)} = (2\omega_{kj}\phi_{kj} + 1)_{(n\times m)}. \quad (19)$$

Using the results given in Section 2, we have,

$$\sum_{t_{s(m\times m)}} = S'S \left(\gamma_2 \sum_{y_{(m\times m)}} - \frac{1}{2}\gamma_1 \sum_{yx_{(m\times n)}} \Omega_{(n\times m)} - \frac{1}{2}\gamma_1 \Omega'_{(m\times n)} \sum_{yx_{(n\times m)}} + \frac{1}{4}\gamma_1 \Omega'_{(m\times n)} \sum_{x_{(n\times n)}} \Omega_{(n\times m)} \right). \quad (20)$$

We differentiate the above expression with respect to Ω and get the optimum value of Ω as,

$$\Omega_{opt(n\times m)} = 2 \sum_{x_{(n\times n)}}^{-1} \sum_{yx_{(n\times m)}}.$$

On using the optimum value of Ω in (20) and we get the minimum value of variance covariance matrix of t_s as,

$$\min \sum_{t_{s(m\times m)}} = S'S \left(\gamma_2 \sum_{y_{(m\times m)}} - \gamma_1 \sum_{yx_{(m\times n)}} \sum_{x_{(n\times n)}}^{-1} \sum_{yx_{(n\times m)}} \right). \quad (21)$$

and

$$\sum_{ts_{(m\times m)}} = [\text{cov}(t_j, t_k)]_{(m\times m)};\ j, k = 1, 2, \ldots, m \text{ for } j = k,$$

$$\text{cov}(t_j, t_k) = \text{var}(t_j)$$

Now in following Remark 1, we are discussing some multivariate estimators as special cases which can be obtained directly from the above results such as t_{s1}, t_{s2}, and t_{sn} along with their variance covariance matrices $\sum_{(t_{s_1})}$, $\sum_{(t_{s2})}$, and $\sum_{(t_{sn})}$ using multi-auxiliary variables.

3.1.1. Remark 1

It is noted that we may get the different multivariate estimators for any number of auxiliary variables assigning different values to m and n into (8). For example taking $m = 3$ & $n = 3$ into (8) one may get a trivariate estimator for variance as,

$$t_{s1} = [t_{s1j}]_{(1\times 3)},$$

where

$$t_{s1j} = \left(s^2_{y_{j(2)}} + \sum_{k=1}^{3} \omega_{kj} (S^2_{x_k} - s^2_{x_{(1)k}}) \right) \exp \sum_{k=1}^{3} \left[\frac{S^2_{x_k} - s^2_{x_{(1)k}}}{S^2_{x_k} + s^2_{x_{(1)k}}} \right]$$

$$k = 1, 2, 3 \ \&\ j = 1, 2, 3 \quad (22)$$

and the variance covariance matrix may be obtained from (20) as,

$$\sum_{t_{s1(3\times 3)}} = S'S \left(\gamma_2 \sum_{y_{(3\times 3)}} - \frac{1}{2}\gamma_1 \sum_{yx_{(3\times 3)}} \Omega_{(3\times 3)} - \frac{1}{2}\gamma_1 \Omega'_{(3\times 3)} \sum_{yx_{(3\times 3)}} + \frac{1}{4}\gamma_1 \Omega'_{(3\times 3)} \sum_{x_{(3\times 3)}} \Omega_{(3\times 3)} \right). \quad (23)$$

Similarly, one may get different bivariate and univariate estimators taking $m = 2$ & 1 respectively into (8) for any number of auxiliary variables, and also one may get variance covariance matrices directly from (20).

A bivariate estimator based on three auxiliary variables can be obtained from (8) taking $m = 2$ and $n = 3$ as,

$$t_{s2} = [t_{s2j}]_{(1\times 2)}, \text{ where}$$

$$t_{s2j} = \left(s^2_{y_{l(2)}} + \sum_{k=1}^{3} \omega_{kj} (S^2_{x_k} - s^2_{x_{(1)k}}) \right) \exp \left[\sum_{k=1}^{3} \frac{S^2_{x_k} - s^2_{x_{(1)k}}}{S^2_{x_k} + s^2_{x_{(1)k}}} \right]; \quad (24)$$

The variance covariance matrix is obtained as,

$$\sum_{t_{s2(2\times 2)}} = S'S \left(\gamma_2 \sum_{y_{(2\times 2)}} - \frac{1}{2}\gamma_1 \sum_{yx_{(2\times 3)}} \Omega_{(3\times 2)} - \frac{1}{2}\gamma_1 \Omega'_{(2\times 3)} \sum_{yx_{(3\times 2)}}' + \frac{1}{4}\gamma_1 \Omega'_{(2\times 3)} \sum_{x_{(3\times 3)}} \Omega_{(3\times 2)} \right). \quad (25)$$

Similarly a univariate estimator based on three auxiliary variables can be obtained from (8) taking $m = 1$ and $n = 3$ as,

$$t_{s3} = [t_{s3j}]_{(1\times 1)}, \text{ where}$$

$$t_{s3j} = \left(s^2_{y_{(2)}} + \sum_{k=1}^{3} \omega_k (S^2_{x_k} - s^2_{x_{(1)k}}) \right) \exp \sum_{k=1}^{3} \left[\frac{S^2_{x_k} - s^2_{x_{(1)k}}}{S^2_{x_k} + s^2_{x_{(1)k}}} \right], \quad (26)$$

and the variance covariance matrix for the estimator in (26) may be obtained as,

Table 1
Relative efficiencies of the Proposed Estimators and Existing estimators.

Estimators	Trivariate case			Bivariate case			Univariate case		
	Single auxiliary	Two auxiliary	Three auxiliary	Single auxiliary	Two auxiliary	Three auxiliary	Single auxiliary	Two auxiliary	Three auxiliary
t_0	100	100	100	100	100	100	100	100	100
t_s	126.719	146.147	337.245	120.601	137.821	234.98	110.234	114.909	125.554
t_a	120.962	125.345	135.619	111.019	114.888	119.34	110.158	110.164	110.594
t_{rg}	120.962	125.345	135.619	111.019	114.888	119.34	110.158	110.164	110.594

$$\sum_{t_{s3(1\times1)}} = S'S\left(\gamma_2 \sum_{y_{(1\times1)}} - \frac{1}{2}\gamma_1 \sum_{yx_{(1\times3)}} \Omega_{(3\times1)} - \frac{1}{2}\gamma_1 \Omega'_{(1\times3)} \sum_{yx_{(3\times1)}} + \frac{1}{4}\gamma_1 \Omega'_{(1\times3)} \sum_{x_{(3\times3)}} \Omega_{(3\times1)}\right). \quad (27)$$

Hence one may obtain different estimators along with their MSE's under univariate case for multi-auxiliary variables.

4. Numerical illustration

We conducted an empirical study to demonstrate the efficiency of the proposed class of Multivariate (trivariate and bivariate) estimator using one, two, and three auxiliary variables. Similarly a univariate version from the proposed multivariate estimator has also been presented for its numerical efficiency. This empirical study has been constructed using a real population data and the population detail is given in Appendix. We consider the three study variables (Y_1, Y_2, Y_3) for trivariate estimator, two study variables (Y_1, Y_2) for bivariate estimator and one study variable $Y1$ for univariate estimator and further in each case used three auxiliary variables, two auxiliaries and single auxiliary variable were used. Variances and co-variances are presented in Appendix and expressions in (23), (25) and (27) were used to calculate the MSE values respectively for the trivariate, bivariate and univariate estimators. In case of multivariate estimators traces were compared with. We have computed the percent relative efficiencies (PREs) of our regression-cum-exponential estimator t_s with respect to t_0 following,

$$PREs(t_*, t_0) = \frac{Var(t_0)}{MSE(t_*)} \times 100$$

Table 1 demonstrates the relative efficiency of each estimator. It is observed that by increasing number of the auxiliary variables we get more efficient results in all three cases i.e. multivariate, bivariate and univariate.

5. Simulation study

However to assess the performance of our proposed estimators we have also computed the results by simulation study. We have taken a model for our regression-cum-exponential estimator as,

$$\text{Observed Model } Y_{ij} = \sum X + \varepsilon_i \qquad \varepsilon_i \sim N(0, 1)$$

The three study variables with three auxiliary variables under Model are distinct by the equation as,

$$Y_1 = 0.2X_1 + 0.3X_2 + 0.9X_3 + \varepsilon,$$

$$Y_2 = 0.6X_1 + 0.5X_2 + 0.3X_3 + \varepsilon,$$

$$Y_3 = 0.5X_1 + 0.7X_2 + 0.4X_3 + \varepsilon,$$

On the first-phase we selected a sample of size $n_1 = 0.5N$ units by SRSWOR using R function. In second-phase, we again selected a sample of size $n_2 = 0.4n_1$ units from the samples selected at first phase by SRSWOR using R function when $N = 3000$. This procedure is repeated 5000 times to calculate the several values of t. At last we have calculated the variance covariance matrices as,

Table 2
Var-Cov matrices of the estimators under the observed model.

	t_{s1}	t_{s2}	t_{s3}
Variance Covariance Matrix of Proposed Estimator t_s			
t_{s1}	0.51432340	0.11534450	0.19009570
t_{s2}	0.11534450	0.07561682	0.10896780
t_{s3}	0.19009570	0.10896780	0.16854950
	t_{01}	t_{02}	t_{03}
Variance Covariance Matrix of t_0 Estimator			
t_{01}	0.5147962	0.11543237	0.1902855
t_{02}	0.1154324	0.07574655	0.1091623
t_{03}	0.1902855	0.10916228	0.1688723
	t_{a1}	t_{a2}	t_{a3}
Variance Covariance Matrix of t_a Estimator			
t_{a1}	0.5113866	0.11627633	0.1917395
t_{a2}	0.1162763	0.07741426	0.1116057
t_{a3}	0.1917395	0.11160569	0.1727231
	t_{reg1}	t_{reg2}	t_{reg3}
Variance Covariance Matrix of modified t_{reg} Estimator			
t_{reg1}	0.5159262	0.11570484	0.1906460
t_{reg2}	0.1157048	0.07580751	0.1092363
t_{reg3}	0.1906460	0.10923628	0.1689522
	t_{rg1}	t_{rg2}	t_{rg3}
Variance Covariance Matrix of t_{rg} Estimator			
t_{rg1}	0.5147965	0.11543233	0.1902858
t_{rg2}	0.1154323	0.07574648	0.1091623
t_{rg3}	0.1902858	0.10916227	0.1688726

Table 3
Determinants of variance covariance matrices.

Estimators	Determinants $\lvert\sum\rvert$
$\mathbf{t}_s$	0.0002516782
$\mathbf{t}_0$	0.0002531813
$\mathbf{t}_a$	0.0002632589
$\mathbf{t}_{reg}$	0.0002536346
$\mathbf{t}_{rg}$	0.0002531837

Table 4
MSE values of the univariate estimators.

Estimators	MSE's
$\mathbf{t}_s$	0.07561682
$\mathbf{t}_0$	0.07574655
$\mathbf{t}_a$	0.07741426
$\mathbf{t}_{reg}$	0.07580751
$\mathbf{t}_{rg}$	0.07574648

Variance Vector $= S = [S_{y1}^2 \quad S_{y2}^2 \quad S_{y3}^2]_{(1\times3)}$ and

$$\sum_{(3\times3)} = \begin{bmatrix} \mathrm{var}(t_1) & \mathrm{cov}(t_1t_2) & \mathrm{cov}(t_1t_3) \\ \mathrm{cov}(t_2t_1) & \mathrm{var}(t_2) & \mathrm{cov}(t_2t_3) \\ \mathrm{cov}(t_1t_3) & \mathrm{cov}(t_2t_3) & \mathrm{var}(t_3) \end{bmatrix}$$

where

$$\mathrm{var}(t_i) = \frac{1}{q}\sum_{i=1}^{q}(t_i - T)^2, T = \frac{1}{q}\sum_{i=1}^{q} t_i \text{ and}$$

$$\mathrm{cov}(t_i) = \frac{1}{q}\sum_{i=1}^{q}(t_i - T)(t_j - T)$$

5.1. Simulated population

$$X_1 \sim N(12, 2),\ X_2 \sim N(15, 3),\ X_3 \sim N(18, 4),$$

In Table 3 variance-covariance matrices for each multivariate estimator are given. Table 3 shows the results of the determinants for the variance-covariance matrices (given in Table 2). From Table 3 it is found that the determinant for the proposed multivariate estimator is less than the determinant of any other existing multivariate estimator which is providing evidence that the proposed estimator is more efficient. Table 4 shows the values of the MSE's of the univariate estimator obtained through the simulation study.

We performed simulation study at various sample sizes, but in order to avoid length of the paper and also complexion of the simulation results it is difficult to show the results of all possible simulations. We presented simulation results only for one sample size (i.e. $n_1 = 0.5N$) and taking iterations of $k = 5000$ samples and results are shown in Table 2. In simulation study it is confirmed that the proposed estimator is asymptotically normal and also the MSE's remain less than existing estimators as the sample size increases which reveals that proposed estimator is more consistent and further it is noted that MSE of the proposed estimators gradually decreases by increasing the sample size from 1% to 50% (as 1%, 2.5%, 5%, 10%, 20%, 50%) of population size (N) which reflected the consistency of the proposed estimator and this gain in consistency is more than the other existing estimators.

6. Concluding remarks

From empirical study given in Table 1 it is concluded that our proposed estimator (trivariate/bivariate/univariate) $\mathbf{t}_s$ is more efficient than $t_0,\ t_a,\ t_{rg}$ estimators. From Table 1, it is also observe that by increasing the auxiliary information, our proposed estimator gives more efficient result. From results of the simulation study presented in Tables 2 and 3 it is concluded that proposed multivariate estimator are more efficient as they have less MSE values as well as values of the determinants are minimum than the determinant values of the existing estimators. Table 4 reveals that univariate version of the proposed estimator has minimum value of the MSE than the MSE's of the existing estimator and therefore it is concluded that our proposed regression-cum-exponential estimator performs more efficiently. In simulation it is confirmed that proposed estimator is asymptotically normal and more consistent the existent estimators.

Appendix A

Source of Population: Gujarati (2004), pg. 385 $N = 35$, $n_1 = 18$ & $n_2 = 9$ (see Tables A1–A3).

Table A1
Details of variables for Population.

Population	Y_1	Y_2	Y_3	X_1	X_2	X_3
1	Hours (average hours worked during the year)	Assets (average family assets holdings)	ERSP (Average yearly earnings of spouse)	Rate (average hourly wage)	School (average highest grade of school computed)	NINE (average yearly non-earned income)

Table A2
Covariance and correlation covariance matrix.

Cov	Y_1	Y_2	Y_3	X_1	X_2	X_3
Y_1	4110.787	132319.6	2041.883	17.43723	51.01613	5953.167
Y_2	132319.584	8229167.3	201899.782	985.23859	2113.56387	387062.227
Y_3	2041.883	201899.8	65934.879	64.51825	164.66193	8224.772
X_1	17.43723	985.23859	64.51825	0.2116820	0.4508139	43.31764
X_2	51.01613	2113.56387	164.66193	0.4508139	1.3638151	86.03345
X_3	5953.167	387062.227	8224.772	43.3176387	86.0334454	18669.06387

Table A3
Correlation matrix.

Population	Y_1	Y_2	Y_3	X_1	X_2	X_3
Y_1	1.0000000	0.7194220	0.1240255	0.5911166	0.6813454	0.6795546
Y_2	0.7194220	1.0000000	0.2740945	0.7464858	0.6308988	0.9875102
Y_3	0.12402545	0.2740945	1.0000000	0.54611367	0.5491081	0.2344256
X_1	0.5911166	0.7464858	0.5461137	1.0000000	0.8390302	0.6890672
X_2	0.6813454	0.6308988	0.5491081	0.8390302	1.0000000	0.5391732
X_3	0.6795546	0.9875102	0.2344256	0.6890672	0.5391732	1.0000000

References

Ahmad, Z., Hanif, M., 2010. Generalized multi-phase multivariate regression estimator for partial information case using multi-auxiliary variables. World Appl. Sci. J. 110 (3), 370–379.

Abu-Dayyeh, W., Ahmed, M.S., 2005. Ratio and regression estimators for the variance under two-phase sampling. Int. J. Stat. Sci. 4, 49–56.

Ahmed, M.S., Raman, M.S., Hossain, M.I., 2000. Some competitive estimators of finite population variance Multivariate Auxiliary Information. Inf. Manage. Sci. 11 (1), 49–54.

Asghar, A., Sanaullah, A., Hanif, M., 2014. Generalized exponential type estimator for population variance in survey sampling. Rev. Colomb. Estadistica 37 (1), 211–222.

Ahmad, Z., Hussain, I., Hanif, M., 2016. Estimation of finite population variance in successive sampling using multi-auxiliary variables. Commun. Stat. Theory Methods 45 (3), 553–565.

Cebrian, A.A., Garcia, M.R., 1997. Variance estimation using auxiliary information: an almost unbiased multivariate ratio estimator. Metrika 45 (1), 171–178.

Graunt, J., 1662. Natural and Political Observations upon the Bill of Mortality. John Martyn, London.

Gujarati, D.N., 2004. Basic Econometrics. The McGraw-Hill Companies.

Isaki, C., 1983. Variance estimation using auxiliary information. J. Am. Stat. Assoc. 78, 117–123.

John, S., 1969. On multivariate ratio and product estimators. Biometrika 56, 533–536.

Kadilar, C., Cingi, H., 2006. Improvement in variance estimation using auxiliary information. Hacettepe J. Math. Stat. 35 (1), 111–115.

Olkin, I., 1958. Multivariate ratio estimation for finite populations. Biometrika 45, 154–165.

Singh, S., Singh, H.P., Tailor, R., Allen, J., Kozak, M., 2009. Estimation of ratio of two finite population means in the presence of nonresponse. Commun. Stat. Theory Methods 38 (19), 3608–3621.

Singh, H.P., Pal, S.K., 2016. Improved estimation of finite population variance using auxiliary information in presence of measurement error. Rev. Investigacion Operacional 37 (2), 147–162.

Singh, H.P., Solanki, R.S., 2013. A new procedure for variance estimation in simple random sampling using auxiliary information. Stat. Pap. 54 (2), 479–497.

Shabbir, J., Gupta, S., 2015. A note on generalized exponential type estimator for population variance in survey sampling. Rev. Colomb. Estadistica 38 (2), 385–397.

Upadhyaya, L.N., Singh, H.P., 2006. Almost unbiased ratio and product-type estimators of finite population variance in sample surveys. Stat. Transition 7 (5), 1087–1096.

Yadav, S.K., Kadilar, C., 2013. Improved exponential type ratio estimator of population variance. Rev. Colomb. Estadística 36 (1). 145–15.

Further reading

Hanif, M., Ahmed, Z., Ahmad, M., 2009. Generalized multivariate ratio estimator using multi-auxiliary variables for Multi-Phase sampling. Pak. J. Stat. 25 (4), 615–629.