Journal of Statistical Theory and Applications
Vol. 19(2), June 2020, pp. 185–195
DOI: https://doi.org/10.2991/jsta.d.200507.002; ISSN 1538-7887
https://www.atlantis-press.com/journals/jsta

Generalized Chain Exponential-Type Estimators under Stratified Two-Phase Sampling with Subsampling the Nonrespondents

Aamir Sanaullah[1], Muhammad Hanif[2,*]

[1]COMSATS University Islamabad, Lahore Campus, Islamabad, Pakistan
[2]National College of Business Administration and Economics, Lahore, Pakistan

ARTICLE INFO

Article History

Received 03 Mar 2019
Accepted 14 Nov 2019

Keywords

Auxiliary variables
Nonresponse
Stratified two-phase sampling
Exponential-type estimator
Chain estimator

ABSTRACT

In this paper some generalized exponential-type chain estimators have been proposed for the finite population mean in the presence of nonresponse under stratified two-phase sampling when mean of another auxiliary variable is readily available. The expressions for the bias and mean square error of proposed estimators have been derived. The comparisons for proposed estimators have been made in theory with Hansen-Hurwitz's, J. Am. Stat. Assoc. 41 (1946), 517–529, and Tabasum and Khan's, J. Indian Soc. Agric. Stat. 58 (2004), 300–306, two-phase ratio and product estimators modified to the stratified sampling. An empirical study has also been carried out to demonstrate the performances of the estimators.

1. INTRODUCTION

In a survey it is aimed to get hold of information on the subject of a target population. In prior of the survey the demarcation of the target population should be apparently affirmed. In an ideal world, all the selected units take part and make available, the requested information. On the other hand, the reality is not the same, notwithstanding how carefully the survey is planned and conducted, to acquire information on some of the units will not be possible, due to the variety of reasons even after callbacks, which is known as nonresponse.

In the presence of nonresponse, obtaining high response rates in the presence of nonresponse has been main aim of the survey statisticians. This growing interest is due to the significance of nonresponse bias in survey sampling. Madow *et al.* [1] discussed weighting adjustment and imputation methods to deal with 11 situations of nonresponse. Lessler and Kalsbeek [2] provided weighting adjustment and imputation procedure for 15 different situations. Little and Rubin [3] suggest to ignore the incomplete information. This may be used where nonresponse is very low otherwise by doing such a method there occur a serious bias. Reweighting does not guarantee the adjustment of nonresponse bias. It may happen but most often one only can assume if the auxiliary information correlates strongly both with response propensity and study variable(s). If both of those conditions are satisfied the variance and mean square error (MSE) are reduced. A successful method of adjusting nonresponse bias is to use strongly correlated auxiliary information. In result of this, nonresponse and variance both may reduce Djerf [4,5], and Horngren [6]. Kalton ([7], p. 63) states "among the potential variables for use in forming weighting classes, the ones that are most effective in reducing nonresponse bias are those that are highly correlated both with the survey variables and the (0, 1) response variable." Two types of auxiliary variables can be used if the auxiliary variables are known for all sampled units, then the adjustment is called sample-based; if they are known for the entire population, the adjustment is population-based [8,9]. The population-based adjustment is especially useful when the population totals are known. Sample-based adjustments need data for the full sample but do not require knowing control totals for the entire population. Sample- and population-based adjustments are equally effective for dealing with nonresponse bias [10,11]. Hansen and Hurwitz [12] were the first to develop a procedure to elicit response from the subsample of nonresponse. They envisaged an estimator for the estimation of population mean in the presence of nonresponse. Variance expression along with the optimum sampling fraction among nonrespondents was also derived. The procedure presented by Hansen-Hurwitz, is the edition of two-phase sampling, proposed by Neyman [13]. The technique was illustrated under simple random sampling design and it is also equally holds good for stratified sampling design and for other sampling designs.

*Corresponding author. Email: drmianhanif@gmail.com

Following Hansen-Hurwitz [12], Cochran [14] proposed a ratio estimator in simple random sampling for dealing with nonresponse. Okafor and Lee [15] advised a ratio estimator, which was first proposed by Khare and Srivastave [16] under two-phase sampling. Tabasum and Khan [17,18] extended the work done by Okafor and Lee [15] and studied some properties of the estimator in the presence of nonresponse under two-phase sampling. Singh *et al.* [19] developed some generalized exponential-type estimators under two-phase sampling to deal with response. Ismail *et al.* [20], Gamrot [21] and Shabbir and Khan [22] have recommended some improvements for the estimation of population mean in the presence of nonresponse using single or more auxiliary variables. Sanaullah *et al.* (2014b) proposed some generalized exponential-type estimators under stratified sampling for estimating population mean in two different situations of nonresponse.

1.1. Notations and Stratified Two-Phase Sampling with Subsampling the Nonrespondents

In many situations of practical importance, the population mean of either of the auxiliary variable, e.g. $\overline{X}_h$ is not available in prior of a survey, in such a situation it is very usual to estimate it by the sample mean $\overline{x}'_h$ based on a preliminary first-phase sample of size $n'_h \left(n' = \sum_{h=1}^{L} n'_h \right)$ of which $n_h \left(n = \sum_{h=1}^{L} n_h \right)$ is a subsample, i.e. $\left(n_h \subset n'_h \right)$. At the most, we use only knowledge of the population mean of another auxiliary variable, e.g. $\overline{Z}_h$, which is closely related to $\overline{X}_h$ but remotely correlated to the main variable. That is if $\overline{Z}_h$ is known to us, then it is advisable to estimate $\overline{X}_h$ by $\hat{\overline{X}}_h = \overline{x}'_h \frac{\overline{Z}}{\overline{z}'_{st}}$, where h = 1, 2, ⋯, L, which would provide a better estimate of $\overline{X}_h$ than $\overline{x}'_h$ (Sanaullah et al., 2014a). Let us assume that at the first phase, all the n'_h units provide information on auxiliary characteristics. At the second phase from the sample n_h, let $n_{h(1)}$ units provide the response for the requested information and $n_{h(2)}$ units do not. Following Hansen-Hurwitz [12] sub-sampling, a sub-sample of size r_h from $n_{h(2)}$ non-respondents is selected at random and is approached for their direct interview such that $r_h = n_{h(2)}/\mathrm{k_h}$, $\mathrm{k_h} > 1$. Here it is assumed that all the r_h units provide the requested information.

When there occurs nonresponse on study variable as well as on the auxiliary variable, the usual two-phase ratio and product estimators for population mean are defined in stratified sampling respectively as

$$t_1 = \overline{y}^*_{st} \overline{x}'_{st} / \overline{x}^*_{st}, \text{ (Ratio estimator)} \tag{1}$$

$$t_2 = \overline{y}^*_{st} \overline{x}^*_{st} / \overline{x}'_{st}, \text{ (Product estimator)} \tag{2}$$

where $\overline{y}^*_{st}$ and $\overline{x}^*_{st}$ are Hansen-Hurwitz estimators modified to the stratified sampling for population means $\overline{X}$ and $\overline{Y}$ respectively and these are defined as $\overline{y}^*_{st} = \sum_{h=1}^{L} P_h \left(n_{h(1)} \overline{y}_{h(1)} + n_{h(2)} \overline{y}_{h(2)r} \right) / n_h$, and $\overline{x}^*_{st} = \sum_{h=1}^{L} P_h \left(n_{h(1)} \overline{x}_{h(1)} + n_{h(2)} \overline{x}_{h(2)r} \right) / n_h$ with $P_h = N/N_h$, $\left(\overline{y}_{h(1)}, \overline{x}_{h(1)} \right)$, and $\left(\overline{y}_{h(2)r}, \overline{x}_{h(2)r} \right)$ are the sample means for hth stratum based on $n_{h(1)}$ and $n_{h(2)r}$ units respectively, and $\overline{x}'_{st} = \sum_{h=1}^{L} P_h \overline{x}'_h$ is the sample mean based on $n'_h = \sum_{h=1}^{L} n_h$. It is to be pointed out that usual two-phase ratio estimator was Tabasum and Khan [17] in simple random sampling and t_1 is modified form of Tabasum and Khan [17] to two-phase the stratified sampling. The *MSEs* for the ratio estimator t_1 and product estimator t_2 are given respectively as

$$MSE(t_1) = \sum_{h=1}^{L} P_h^2 \left(\lambda'_h S^2_{yh} + \lambda_h \left(S^2_{yh} + R^2_h S^2_{xh} - 2 R_h S_{xyh} \right) + \lambda^*_h \left(S^2_{yh(2)} + R^2_h S^2_{xh(2)} - 2 R_h S_{xyh(2)} \right) \right) \tag{3}$$

and

$$MSE(t_2) = \sum_{h=1}^{L} P_h^2 \left(\lambda'_h S^2_{yh} + \lambda_h \left(S^2_{yh} + R^2_h S^2_{xh} + 2 R_h S_{xyh} \right) + \lambda^*_h \left(S^2_{yh(2)} + R^2_h S^2_{xh(2)} + 2 R_h S_{xyh(2)} \right) \right) \tag{4}$$

where $\left(S^2_{yh}, S^2_{xh} \right)$, and $\left(S^2_{yh(2)}, S^2_{xh(2)} \right)$ are the variances from respondents and nonrespondents respectively with $R_h = \overline{Y}_h / \overline{X}_h$, $\lambda_h = \left(\frac{1}{n_h} - \frac{1}{N_h} \right)$, $\lambda'_h = \left(\frac{1}{n'_h} - \frac{1}{N_h} \right)$, $\lambda^*_h = \left(\frac{k_h - 1}{n_h} \right) W_{h(2)}$ and $W_{h(2)} = N_{2h}/N$.

When population means of the two auxiliary variables are available, Sanaullah *et al.* [23] proposed some exponential-type ratio-cum-ratio estimators for stratified two-phase sampling in the presence of nonresponse as

$$t_3 = \sum_{h=1}^{l} P_h \bar{y}_h^* \exp\left(\frac{\sum_{h=1}^{l} P_h\left(\bar{X}_h - \bar{x}_h'\right)}{\sum_{h=1}^{l} P_h\left(\bar{X}_h + (a-1)\bar{x}_h'\right)}\right)\exp\left(\frac{\sum_{h=1}^{l} P_h\left(\bar{Z}_h - \bar{z}_h^*\right)}{\sum_{h=1}^{l} P_h\left(\bar{Z}_h + (b-1)\bar{z}_h^*\right)}\right)\left(\text{Exponential-type ratio-cum-ratio estimator}\right) \tag{5}$$

where (a, b) are suitably chosen constants to be determined such that MSE of t_3 is minimum.

The MSE of t_3 is as

$$MSE\left(t_3\right) \approx \bar{Y}^2 \sum_{h=1}^{l} P_h^2 \left[\begin{array}{l} \frac{1}{\bar{Y}^2}\left(\lambda_h S_{yh}^2 + \lambda_h^* S_{yh(2)}^2\right) + \frac{1}{a^2\bar{X}^2}\lambda_h' S_{xh}^2 + \frac{1}{b^2\bar{Z}^2}\left(\lambda_h S_{zh}^2 + \lambda_h^* S_{zh(2)}^2\right) \\ -2\left(\frac{1}{a\bar{Y}\bar{X}}\lambda_h' S_{yxh} + \frac{1}{b\bar{Z}\bar{Y}}\left(\lambda_h S_{yzh} + \lambda_h^* S_{yzh2}\right) - \frac{\lambda_h' S_{xzh}}{ab\bar{X}\bar{Z}}\right) \end{array}\right] \tag{6}$$

Sanaullah *et al.* [24] envisaged an exponential-type chain ratio estimator under stratified two-phase sampling and population mean of an auxiliary variable x in not known but population mean of another variable z is on hand.

$$t_4 = \sum_{h=1}^{L} P_h \bar{y}_h \exp\left[\frac{\sum_{h=1}^{L} P_h\left(\bar{x}_h' \frac{\bar{Z}}{\sum_{h=1}^{L} P_h \bar{z}_h'} - \bar{x}_h\right)}{\sum_{h=1}^{L} P_h\left(\bar{x}_h' \frac{\bar{Z}}{\sum_{h=1}^{L} P_h \bar{z}_h'} + \bar{x}_h\right)}\right]\left(\text{Exponential-type chain ratio}\right) \tag{7}$$

The MSE of t_4 is as

$$MSE(t_4) \approx \bar{Y}^2 \sum_{h=1}^{L} P_h^2 \left[\begin{array}{l} \frac{1}{\bar{Y}^2}\lambda_h S_{yh}^2 + \frac{1}{4}\left(\frac{1}{\bar{X}^2}\left(\lambda_h S_{xh}^2 - \lambda_h' S_{xh}^2\right) + \frac{1}{\bar{Z}^2}\lambda_h' S_{zh}^2\right) \\ -\left(\frac{1}{\bar{Y}\bar{X}}\left(\lambda_h S_{xyh} - \lambda_h' S_{xyh}\right) + \frac{1}{\bar{Y}\bar{Z}}\lambda_h' S_{yzh}\right) \end{array}\right] \tag{8}$$

In this study, an attempt has been made for the development of generalized exponential-type chain ratio and product estimators using two auxiliary variables under stratified two-phase random sampling. The estimators have been proposed for the case when there occurs nonresponse on all the variables in second phase.

2. PROPOSED GENERALIZED EXPONENTIAL-TYPE CHAIN RATIO-CUM-RATIO AND PRODUCT-CUM-PRODUCT ESTIMATORS

Now it is assumed that information on a secondary auxiliary variable z is to be had. Then taking motivation from Sanaullah *et al.* [23,24], the inspiration of exponential-type chain ratio and exponential ratio-cum-ratio estimators have been combined together under stratified two-phase sampling design when there are auxiliary variables x and z which are correlated with study variable y in case of nonresponse. By following the same lines, another estimator (exponential-type chain product-cum-product estimator) has been proposed with its properties in the presence of nonresponse.

2.1. Generalized Exponential-Type Chain Ratio-Cum-Ratio Estimator

Motivated from Sanaullah *et al.* [23,24], we consider a form of an exponential-type chain ratio-cum-ratio estimator for stratified two-phase sampling in the presence of non-response as

$$t^{1}_{r(2.2)} = \sum_{h=1}^{L} P_h \bar{y}^*_h \exp\left(1 - \frac{2\sum_{h=1}^{L} P_h \bar{x}^*_h}{\sum_{h=1}^{L} P_h \left(\bar{x}'_h \frac{\bar{Z}}{\sum_{h=1}^{L} P_h \bar{z}'_h} + \bar{x}^*_h\right)}\right) \exp\left(1 - \frac{2\sum_{h=1}^{L} P_h \bar{z}^*_h}{\sum_{h=1}^{L} P_h \left(\bar{Z}_h + \bar{z}^*_h\right)}\right), \tag{9}$$

The estimator $t^{1}_{r(2.2)}$ in (9) leads to the form of generalized exponential-type chain ratio-cum-ratio estimator for population mean under stratified two-phase sampling in case of nonresponse as

$$t^{g}_{r(a,b)} = \sum_{h=1}^{L} P_h \bar{y}^*_h \exp\left(1 - \frac{a\sum_{h=1}^{L} P_h \bar{x}^*_h}{\sum_{h=1}^{L} P_h \bar{x}'_h \frac{\bar{Z}}{\sum_{h=1}^{L} P_h \bar{z}'_h} + (a-1)\sum_{h=1}^{L} P_h \bar{x}^*_h}\right) \exp\left(1 - \frac{b\sum_{h=1}^{L} P_h \bar{z}^*_h}{\sum_{h=1}^{L} P_h \bar{Z}_h + (b-1)\sum_{h=1}^{L} P_h \bar{z}^*_h}\right) \tag{10}$$

where $a\,(\neq 0)$ and $b\,(\neq 0)$ are assumed unknown constants to be determined in such way whose values make the *MSE* of $t^{g}_{r(a,b)}$ minimum.

It is observed that for various values of a and b in (10), we get various exponential-type chain ratio-cum-ratio estimators as deduced. From this class some examples are presented in Table 1 as follows:

In order to obtain the bias and mean square of the estimators, let us define

$$\left.\begin{array}{l} \bar{y}^*_h = \bar{Y}_h(1+e^*_{0h}), \bar{x}'_h = \bar{X}_h(1+e'_{1h}), \bar{x}^*_h = \bar{X}_h(1+e^*_{1h}), \bar{z}'_h = \bar{Z}_h(1+e'_{2h}), \bar{z}^*_h = \bar{Z}_h(1+e^*_{2h}) \\ \vartheta^*_{200} = V^*_{200} - V'_{200}, \vartheta^*_{110} = V^*_{110} - V'_{110}, S^2_{yh} = \sum_{i=1}^{N_h} \frac{\left(y_i - \bar{Y}\right)^2}{N_h - 1}, S^2_{yh(2)} = \sum_{i=1}^{N_{h(2)}} \frac{\left(y_i - \bar{Y}_{h(2)}\right)^2}{N_{h(2)} - 1} \end{array}\right\}, \tag{11}$$

where e^*_{ih} shows sampling error at second phase sampling in the presence of non-response, and e'_{ih} shows sampling error at first phase sampling without nonresponse and we consider that $\mathrm{E}\left(e^*_{ih}\right) = \mathrm{E}\left(e'_{ih}\right) = 0$ where $i = 0, 1, 2$.

Let $V_{r,\,s,\,t} = \sum_{h=1}^{L} P_h^{r+s+t} E\left(\left(\frac{\bar{x}_h - \bar{X}_h}{\bar{X}}\right)^r \left(\frac{\bar{y}_h - \bar{Y}_h}{\bar{Y}}\right)^s \left(\frac{\bar{z}_h - \bar{Z}_h}{\bar{Z}}\right)^t\right)$ where $(r, s, t) = 0, 1, 2$, and using (11), expectations are defined as

$$\left.\begin{array}{ll} E\left(e^*_0\right)^2 = \frac{1}{\bar{Y}^2} \sum_{h=1}^{l} P_h^2 \left(\lambda_h S^2_{yh} + \lambda^*_h S^2_{yh2}\right) = V^*_{020} & E\left(e'_1\right)^2 = \frac{1}{\bar{X}^2} \sum_{h=1}^{l} P_h^2 \lambda'_h S^2_{xh} = V'_{200} \\ E\left(e^*_1\right)^2 = \frac{1}{\bar{X}^2} \sum_{h=1}^{l} P_h^2 \left(\lambda_h S^2_{xh} + \lambda^*_h S^2_{xh2}\right) = V^*_{200} & E\left(e^*_2\right)^2 = \frac{1}{\bar{Z}^2} \sum_{h=1}^{l} P_h^2 \left(\lambda_h S^2_{zh} + \lambda^*_h S^2_{zh2}\right) = V^*_{002} \\ E\left(e^*_0.e^*_2\right) = \frac{1}{\bar{Y}\bar{Z}} \sum_{h=1}^{l} P_h^2 \left(\lambda_h S_{yzh} + \lambda^*_h S_{yzh2}\right) = V^*_{011} & E\left(e^*_0.e^*_1\right) = \frac{1}{\bar{Y}\bar{X}} \sum_{h=1}^{l} P_h^2 \left(\lambda_h S_{xyh} + \lambda^*_h S_{xyh2}\right) = V^*_{110} \\ E\left(e^*_0.e'_1\right) = \frac{1}{\bar{Y}\bar{X}} \sum_{h=1}^{l} P_h^2 \lambda'_h S_{yxh} = V'_{110} & E\left(e'_1.e'_2\right) = \frac{1}{\bar{Z}\bar{X}} \sum_{h=1}^{l} P_h^2 \lambda'_h S_{xzh} = V'_{101} \end{array}\right\}, \tag{12}$$

Table 1 Some members of the class of the estimator $t^g_{r(a,b)}$.

Exponential-Type Chain Ratio-Cum-Ratio Estimators $t^g_{r(a,b)}$	a	b
$t^1_{r(2,2)} = \sum_{h=1}^{L} P_h \overline{y}^*_h \exp\left(1 - \frac{2\sum_{h=1}^{L} P_h \overline{x}^*_h}{\sum_{h=1}^{L} P_h \left(\overline{x}'_h \frac{\overline{Z}}{\sum_{h=1}^{L} P_h \overline{z}'_h} + \overline{x}^*_h\right)}\right) \exp\left(1 - \frac{2\sum_{h=1}^{L} P_h \overline{z}^*_h}{\sum_{h=1}^{L} P_h \left(\overline{Z}_h + \overline{z}^*_h\right)}\right)$	2	2
$t^2_{r(2,1)} = \sum_{h=1}^{L} P_h \overline{y}^*_h \exp\left(1 - \frac{2\sum_{h=1}^{L} P_h \overline{x}^*_h}{\sum_{h=1}^{L} P_h \left(\overline{x}'_h \frac{\overline{Z}}{\sum_{h=1}^{L} P_h \overline{z}'_h} + \overline{x}^*_h\right)}\right) \exp\left(1 - \frac{\sum_{h=1}^{L} P_h \overline{z}^*_h}{\sum_{h=1}^{L} P_h \overline{Z}_h}\right)$	2	1
$t^3_{r(1,2)} = \sum_{h=1}^{L} P_h \overline{y}^*_h \exp\left(1 - \frac{\sum_{h=1}^{L} P_h \overline{x}^*_h}{\sum_{h=1}^{L} P_h \overline{x}'_h \frac{\overline{Z}}{\sum_{h=1}^{L} P_h \overline{z}'_h}}\right) \exp\left(1 - \frac{2\sum_{h=1}^{L} P_h \overline{z}^*_h}{\sum_{h=1}^{L} P_h \left(\overline{Z}_h + \overline{z}^*_h\right)}\right)$	1	2

Using (11), the estimator in (10) can be expressed in the form of e's as

$$t^g_{r(a,b)} = \sum_{h=1}^{L} P_h \overline{Y}(1+e_{0h}) \exp\left(\frac{\sum_{h=1}^{L} P_h \left(\frac{\overline{X}_h\left(1+e'_{1h}\right)}{\sum_{h=1}^{L} P_h \overline{Z}_h\left(1+e'_{2h}\right)}\overline{Z} - \overline{X}_h\left(1+e^*_{1h}\right)\right)}{\sum_{h=1}^{L} P_h \left(\frac{\overline{X}_h\left(1+e'_{1h}\right)}{\sum_{h=1}^{L} P_h \overline{Z}_h\left(1+e'_{2h}\right)}\overline{Z} + (a-1)\overline{X}_h\left(1+e^*_{1h}\right)\right)}\right) \exp\left(\frac{\sum_{h=1}^{L} P_h \left(\overline{Z}_h - \overline{Z}_h\left(1+e^*_{2h}\right)\right)}{\sum_{h=1}^{L} P_h \left(\overline{Z}_h + (b-1)\overline{Z}_h\left(1+e^*_{2h}\right)\right)}\right) \tag{13}$$

We expand the right-hand side of (13) and neglect the terms in e_i higher than two. After some simplification we will have,

$$t^g_{r(a,b)} - \overline{Y} \approx \overline{Y}\left[e^*_0 + \left(\begin{array}{l} \frac{e'_1 - e^*_1 - e'_2}{a} - \frac{e^*_2}{b} + \left(\frac{e'^2_2}{2a} + \frac{e^{*2}_1}{a} + \frac{(b-1)e^{*2}_2}{b}\right) - \frac{2}{a^2}\left(e^{*2}_1 + e'^2_1 + e'^2_2\right) \\ -\left(\frac{b-1}{b}\right)^2 e^{*2}_2 + \frac{2(b-1)}{b^2}e^{*2}_2 - \frac{1}{a}\left(e^*_1e'_1 - e^*_1e'_2 + e'_1e'_2 - e^*_0e'_1 + e^*_0e^*_1 + e^*_0e'_2\right) \\ +\frac{4}{a^2}\left(e^*_1e'_1 - e^*_1e'_2 + e'_1e'_2\right) - \frac{e^*_0e^*_2}{b} - \frac{1}{ab}\left(e^*_0e^*_1 - e^*_0e'_1 + e^*_0e'_2\right) \end{array}\right)\right] \tag{14}$$

Using (14), the expressions for the bias and MSE of $t^{g}_{r(a,b)}$ are obtained respectively as

$$\text{Bias}\left(t^{g}_{r(a,b)}\right) \approx \overline{Y}\left(\begin{array}{l}\frac{1}{a}\left(\vartheta^{*}_{200} + \frac{V'_{002}}{2} + \vartheta^{*}_{110} + V'_{011}\right) + \frac{3\,(b-1)}{b^{2}}V^{*}_{002} \\ -\frac{2}{a^{2}}\left(\vartheta^{*}_{200} + V'_{002}\right) - \frac{V^{*}_{011}}{b} - \frac{1}{ab}\left(\vartheta^{*}_{110} + V'_{011}\right)\end{array}\right) \tag{15}$$

and

$$MSE\left(t^{g}_{r(a,b)}\right) \approx \overline{Y}^{2}\left(V^{*}_{020} + \frac{1}{a^{2}}\left(\vartheta^{*}_{200} + V'_{002}\right) + \frac{V^{*}_{002}}{b^{2}} - \frac{2}{a}\left(\vartheta^{*}_{110} + V'_{011}\right) - 2\frac{V^{*}_{011}}{b} + \frac{2}{ab}\left(\vartheta^{*}_{101} + V'_{002}\right)\right) \tag{16}$$

The optimal values of a and b for which the MSE $\left(t^{g}_{r(a,b)}\right)$ is minimum, are obtained as

$$\left.\begin{array}{l} a_{opt} = \dfrac{\left(BV^{*}_{002} - A^{2}\right)}{\left(CV^{*}_{002} - AV^{*}_{011}\right)} \quad \text{and} \quad b_{opt} = \dfrac{\left(BV^{*}_{002} - A^{2}\right)}{\left(BV^{*}_{011} - AC\right)} \\ \text{where} \\ A = \vartheta^{*}_{101} + V'_{002},\ B = \vartheta^{*}_{200} + V'_{002}\ C = \vartheta^{*}_{110} + V'_{011} \end{array}\right\} \tag{17}$$

The minimum value of $MSE\left(t^{g}_{r(a,b)}\right)$ as

$$\min .MSE\left(t^{g}_{r(a_{opt},b_{opt})}\right) \approx \overline{Y}^{2}\left(V^{*}_{020} - \frac{\left(BV^{*2}_{011} + C^{2}V^{*}_{002} - 2ACV^{*}_{011}\right)}{BV^{*}_{002} - A^{2}}\right) \tag{18}$$

The bias and MSE expressions for the class of estimators presented in Table 1, can be obtained by putting different values of a and b into (15) and (16) respectively, such as

For $a = 2$and $b = 2$, the bias and MSE of $t^{1}_{r(2,2)}$ is obtained as

$$\text{Bias}\left(t^{1}_{r(2,2)}\right) \approx \frac{\overline{Y}}{2}\left(\vartheta^{*}_{002} + \frac{V'_{002}}{2} - V'_{200} - \frac{3}{2}\left(\vartheta^{*}_{110} + V'_{011}\right) - V^{*}_{011}\right) \tag{19}$$

and

$$MSE\left(t^{1}_{r(2,2)}\right) \approx \overline{Y}^{2}\left(V^{*}_{020} + \frac{1}{4}\left(\vartheta^{*}_{200} + V'_{002}\right) + \frac{1}{4}V^{*}_{002} - \left(\vartheta^{*}_{110} + V'_{011}\right) - V^{*}_{011} + \frac{1}{2}\left(\vartheta^{*}_{101} + V'_{002}\right)\right) \tag{20}$$

For $a = 2$and $b = 1$, the bias and MSE of $t^{2}_{r(2,1)}$ is obtained as

$$\text{Bias}\left(t^{2}_{r(2,1)}\right) \approx -\overline{Y}\left(\frac{V'_{002}}{4} + V^{*}_{011}\right) \tag{21}$$

and

$$MSE\left(t^{2}_{r(2,1)}\right) \approx \overline{Y}^{2}\left(V^{*}_{020} + \frac{1}{4}\left(\vartheta^{*}_{200} + V'_{002}\right) + V^{*}_{002} - \left(\vartheta^{*}_{110} + V'_{011}\right) - 2V^{*}_{011} + \left(\vartheta^{*}_{101} + V'_{002}\right)\right) \tag{22}$$

For $a = 1$and $b = 2$, the bias and MSE of $t^{3}_{r(1,2)}$ is obtained as

$$\text{Bias}\left(t^{3}_{r(1,2)}\right) \approx \overline{Y}\left(\left(\vartheta^{*}_{200} + \frac{V'_{002}}{2} + \vartheta^{*}_{110} + V'_{011}\right) + \frac{3}{4}V^{*}_{002} - 2\left(\vartheta^{*}_{200} + V'_{002}\right) - \frac{V^{*}_{011}}{2} - \frac{1}{2}\left(\vartheta^{*}_{110} + V'_{011}\right)\right) \tag{23}$$

and

$$MSE\left(t^{3}_{r(1,2)}\right) \approx \overline{Y}^{2}\left(V^{*}_{020} + \left(\vartheta^{*}_{200} + V'_{002}\right) + \frac{V^{*}_{002}}{4} - 2\left(\vartheta^{*}_{110} + V'_{011}\right) - V^{*}_{011} + \left(\vartheta^{*}_{101} + V'_{002}\right)\right) \tag{24}$$

2.2. Generalized Exponential-Type Chain Product-cum-Product Estimator

Motivated from Sanaullah *et al.* (2014a, 2014b), we consider a form of an exponential-type chain product-cum-product estimator for stratified two-phase sampling in the presence of non-response as

$$t^{g}_{p(c,d)} = \sum_{h=1}^{L} P_h \bar{y}^*_h \exp\left(\frac{c\sum_{h=1}^{L} P_h \bar{x}^*_h}{\sum_{h=1}^{L} P_h \bar{x}'_h \frac{\bar{Z}}{\sum_{h=1}^{L} P_h \bar{z}'_h} + (c-1)\sum_{h=1}^{L} P_h \bar{x}^*_h} - 1 \right) \exp\left(\frac{d\sum_{h=1}^{L} P_h \bar{z}^*_h}{\sum_{h=1}^{L} P_h \bar{Z}_h + (d-1)\sum_{h=1}^{L} P_h \bar{z}^*_h} - 1 \right), \tag{25}$$

where $c\,(\neq 0)$ and $d\,(\neq 0)$ are assumed unknown constants to be determined in such way whose values make the *MSE* of $t^{g}_{p(c,d)}$ minimum.

It is observed that for various values of c and d in (25), we get various exponential chain product-type estimators as deduced class of $t^{g}_{p(c,d)}$. From this class some examples can be considered in Table 2 as follows:

Table 2 Some members of the class of the estimator $t^{g}_{p(c,d)}$.

Exponential-Type Chain Product-Cum-Product Estimators $t^{g}_{p(c,d)}$	***c***	***d***
$t^{1}_{p(2,2)} = \sum_{h=1}^{L} P_h \bar{y}^*_h \exp\left(\frac{2\sum_{h=1}^{L} P_h \bar{x}^*_h}{\sum_{h=1}^{L} P_h \left(\bar{x}^*_h - \bar{x}'_h \frac{\bar{Z}}{\sum_{h=1}^{L} P_h \bar{z}'_h} \right)} - 1 \right) \exp\left(\frac{\sum_{h=1}^{L} P_h \bar{z}^*_h}{\sum_{h=1}^{L} P_h \left(\bar{z}^*_h - \bar{Z}_h \right)} - 1 \right)$	2	2
$t^{2}_{p(2,1)} = \sum_{h=1}^{L} P_h \bar{y}^*_h \exp\left(\frac{\sum_{h=1}^{L} P_h \bar{x}^*_h}{\sum_{h=1}^{L} P_h \left(\bar{x}^*_h + \bar{x}'_h \frac{\bar{Z}}{\sum_{h=1}^{L} P_h \bar{z}'_h} \right)} - 1 \right) \exp\left(\frac{\sum_{h=1}^{L} P_h \bar{z}^*_h}{\sum_{h=1}^{L} P_h \bar{Z}_h} - 1 \right)$	2	1
$t^{3}_{p(1,2)} = \sum_{h=1}^{L} P_h \bar{y}^*_h \exp\left(\frac{\sum_{h=1}^{L} P_h \bar{x}^*_h}{\sum_{h=1}^{L} P_h \left(\bar{x}'_h \frac{\bar{Z}}{\sum_{h=1}^{L} P_h \bar{z}'_h} \right)} - 1 \right) \exp\left(\frac{\sum_{h=1}^{L} P_h \bar{z}^*_h}{\sum_{h=1}^{L} P_h \left(\bar{z}^*_h + \bar{Z}_h \right)} - 1 \right)$	1	2

We adapt the procedure (11)–(18), expressions for the bias and MSE of $t^{g}_{p(c,d)}$ are obtained respectively as follows:

$$Bias\left(t^{g}_{p(c,d)}\right) = -\overline{Y}\left(\begin{array}{l}\frac{1}{c}\left(\vartheta^{*}_{200} + \frac{V'_{002}}{2} + \vartheta^{*}_{110} + V'_{011}\right) + \frac{d-1}{d}\left(1 - \frac{d-1}{d} + \frac{2}{d}\right)V^{*}_{002} \\ -\frac{2}{c^{2}}\left(\vartheta^{*}_{200} + V'_{002}\right) - \frac{V^{*}_{011}}{d} - \frac{1}{cd}\left(\vartheta^{*}_{110} + V'_{011}\right)\end{array}\right) \tag{26}$$

and

$$MSE\left(t^{g}_{p(c,d)}\right) \approx \overline{Y}^{2}\left(V^{*}_{020} + \frac{1}{c^{2}}\left(\vartheta^{*}_{200} + \vartheta'_{002}\right) + \frac{V^{*}_{002}}{d^{2}} + \frac{2}{c}\left(\vartheta^{*}_{110} + \vartheta'_{011}\right) + 2\frac{V^{*}_{011}}{d} + \frac{2}{cd}\left(\vartheta^{*}_{101} + \vartheta'_{002}\right)\right) \tag{27}$$

The optimum values of c and d are obtained as

$$c_{opt} = -\frac{\left(BV^{*}_{002} - A^{2}\right)}{\left(CV^{*}_{002} - AV^{*}_{011}\right)} \quad \text{and} \quad d_{opt} = -\frac{\left(BV^{*}_{002} - A^{2}\right)}{\left(BV^{*}_{011} - AC\right)} \tag{28}$$

The minimum value of $MSE(t^{g}_{p(c,d)})$ is obtained as

$$\min .MSE\left(t^{g}_{p(c_{opt},d_{opt})}\right) \approx \overline{Y}^{2}\left(V^{*}_{020} - \frac{\left(BV^{*2}_{011} + C^{2}V^{*}_{002} - 2ACV^{*}_{011}\right)}{BCV^{*}_{002} - A^{2}}\right) \tag{29}$$

The bias and MSE expressions for the estimators presented in Table 2 can be obtained directly from (26) and (27) respectively by putting different values of $c$and d.

3. EFFICIENCY COMPARISONS

Now we compare the proposed generalized exponential-type chain estimators with usual Hansen and Hurwitz's [12] unbiased estimator $\overline{y}^{*}_{st}$ and Tabasum and Khan [17] estimators t_1 as

i. Exponential-Type Chain Ratio-Cum-Ratio Estimators

$$\left\langle \begin{array}{c}\min .MSE\left(t^{g}_{r(a^{opt},b^{opt})}\right) < MSE\left(\overline{y}^{*}_{st}\right) \\ \text{if } \frac{2V'_{101}V^{*}_{011}V'_{110}}{4\left(V'^{2}_{110}V^{*}_{002} + V'_{200}V^{*2}_{011}\right)} < 1\end{array}\right\rangle \text{ and } \left\langle \begin{array}{c}\min .MSE\left(t^{g}_{r(a^{opt},b^{opt})}\right) < MSE(t_{1}) \\ \text{if } \frac{\left(\vartheta^{*}_{200} - 2\vartheta^{*}_{110}\right)\left(V'^{2}_{101} - V'_{200}V^{*}_{002}\right)}{\left(V^{*}_{002}V'^{2}_{110} + V'_{200}V^{*2}_{011} - 2V'_{101}V'_{110}V^{*}_{011}\right)} < 1\end{array}\right\rangle \tag{30}$$

ii. Exponential-Type Chain Product-Cum-Product Estimators

$$\left\langle \begin{array}{c}MSE\left(t^{g}_{p(c^{opt},d^{opt})}\right) < MSE\left(\overline{y}^{*}_{st}\right) \\ \text{if } \frac{2V'_{101}V^{*}_{011}V'_{110}}{4\left(V'^{2}_{110}V^{*}_{002} + V'_{200}V^{*2}_{011}\right)} < 1\end{array}\right\rangle \text{ and } \left\langle \begin{array}{c}MSE\left(t^{g}_{p(c^{opt},d^{opt})}\right) < MSE(t_{2}) \\ \text{if } \frac{\left(\vartheta^{*}_{200} + 2\vartheta^{*}_{110}\right)\left(V'^{2}_{101} - V'_{200}V^{*}_{002}\right)}{\left(V^{*}_{002}V'^{2}_{110} + V'_{200}V^{*2}_{011} - 2V'_{101}V'_{110}V^{*}_{011}\right)} < 1\end{array}\right\rangle \tag{31}$$

The proposed estimator will perform better if the above conditions hold.

4. EMPIRICAL RESULTS AND DISCUSSION

In order to examine the performance of proposed estimators under stratified two-phase sampling, we have taken two different stratified populations as,

Population-I: (Source: Koyuncu and Kadilar [25])

We consider number of teachers as study variable (Y), number of students as auxiliary variable (X) and number of classes in primary and secondary schools as another auxiliary variable (Z) for 923 districts at six 6 regions (1: Marmara, 2: Agean, 3: Mediterranean, 4: Central Anatolia, 5: Black Sea, and 6: East and Southeast Anatolia) in Turkey in 2007.

Population-II: [Source: Detailed livelihood assessment of flood affected districts of Pakistan September 2011, Food Security Cluster, Pakistan]

We consider food expenditure as study variable (Y), household earn as auxiliary variable (X) and total expenditure in May (2011) as another auxiliary variable (Z) for (6940) male and (1678) female households in flood affected districts of Pakistan September 2011.

The summery statistics for two populations are given in Appendix Table A.3. Form Table A.3 it is clear that correlations between study variable (Y) and auxiliary variables (X) and (Z) respectively ρ_{xyh} and ρ_{yzh} are positive in each stratum for population-I and these correlations are negative for population-II. It is therefore in order to examine the efficiency of chain ratio-cum-ratio estimators, population-I will be used and population-II is suitable for chain product-cum-product estimators to test empirically for their efficiency. We used stratified sampling for the selection of sample and Neyman allocation was used for allocating the sample size to different strata.

The comparison of proposed generalized exponential-type chain ratio-cum-ratio and exponential-type chain product-cum-product estimators with respect to Hansen and Hurwitz's [12] have been made with Tabasum and Khan [17] modified to the stratified two-phase ratio and stratified two-phase product estimators respectively.

Table A.2 indicates *MSE* values of each estimator at three different nonresponse rates W_{h2}(10%, 20% and 30%), taking for each nonresponse rate four different inverse sampling rates k_h(2.0, 2.5, 3.0 and 3.50). The percent relative efficiency (*PRE*) values for each estimator are computed with respect to the modified form of Hansen-Hurwitz [12] estimator $\overline{y}_{st}^{*}$ in Table A.2 as,

$$PRE = \frac{Var\left(\overline{y}_{st}^{*}\right)}{MSE\left(t_{i(a,b)}^{g}\right)} \times 100$$

where $g = 1, 2, 3$ $i = 1, r, p$ and (a, b) = {(2, 2), (2, 1), (1, 2) (a^{opt}, b^{opt})}.

From Table A.1 it is noticed that *PRE* values for the proposed exponential-type chain ratio-cum-ratio estimators $t_{r(2,2)}^{1}$, $t_{r(2,1)}^{2}$, $t_{r(1,2)}^{3}$ and $t_{r(a^{opt},b^{opt})}^{g}$ increase as the non-response rate increases from 10% to 30%. Similarly at each nonresponse rate, these *PRE* values increase for each estimator as the inverse sampling rate increases. Further it is observed that the *PRE* values of the proposed exponential-type chain ratio-cum-ratio estimators remain higher than the *PRE* values of Tabasum and Khan [17] ratio estimator (t_1) modified to the two-phase sampling. This shows the proposed exponential-type chain ratio-cum-ratio estimators perform more efficiently. Furthermore it is scrutinized that $t_{r(a^{opt},b^{opt})}^{g}$ is the most efficient estimator and from its class of exponential-type chain ratio-cum-ratio estimators $t_{r(2,2)}^{1}$, and $t_{r(2,1)}^{2}$ are the more efficient estimators.

From Table A.1 it is observed that the empirical results can be expressed same for the proposed exponential-type chain product-cum-product estimators $t_{p(2,2)}^{1}$, $t_{p(2,1)}^{2}$, and $t_{p(c^{opt},d^{opt})}^{g}$. The only estimator $t_{p(1,2)}^{3}$ losses its *PRE* values if the nonresponse rate increases from 10% to 30% and due to the reason $t_{p(1,2)}^{3}$ remain no more efficient.

5. CONCLUSION

From the empirical results and discussion, finally it is concluded that the performance of generalized exponential-type chain ratio-cum-ratio $(t_{r(a^{opt},b^{opt})}^{g}$ $t_{r(2,2)}^{1}$, & $t_{r(2,1)}^{2})$ and chain product-cum-product estimators $(t_{p(c^{opt},d^{opt})}^{g}$, $t_{p(2,2)}^{1}$, & $t_{p(2,1)}^{2})$ is better for these populations on the basis of *PRE* values, and therefore, the class of generalized exponential-type chain estimators should be preferred for their practical applications in case of nonresponse.

CONFLICT OF INTEREST

There is no conflict of interest involved and the research was carried out with authors's own contribution without any outside funding.

ACKNOWLEDGMANTS

The authors thank the Associate Editor and anonymous reviewers for their useful comments and suggestions on an earlier version of this manuscript which led to this improved one.

REFERENCES

1. W.G. Madow, I. Olkin, D.B. Rubin, Incomplete Data in Sample Surveys, vol. 2, Academic Press, New York, NY, USA, 1983.
2. J.T. Lessler, W.D. Kalsbeek, Nonsampling Error in Surveys, Wiley, New York, NY, USA, 1992.

3. R.J.A. Little, D.B. Rubin, Statistical Analysis with Missing Data, John Wiley, New York, NY, USA, 1987.
4. K. Djerf, J. Off. Stat. 13 (1997), 29–39.
5. K. Djerf, Properties of Some Estimators Under Unit Nonresponse, Research Report no. 231, Statistics Finland, Helsinki, Finland, 2000.
6. J. Horngren, The Use of Registers as Auxiliary Information in the Swedish Labour Force Survey, R&D Report no. 1992:13, Statistics Sweden, Stockholm, Sweden, 1992.
7. G. Kalton, Compensating for Missing Survey Data, University of Michigan Press, Ann Arbor, MI, USA, 1983.
8. G. Kalton, D. Kasprzyk, Surv. Methodol. 12 (1986), 1–16.
9. S. Lundstrom, C.E. Sarndal, J. Off. Stat. 15 (1999), 305–327.
10. C.E. Sarndal, S. Lundstrom, Estimation in Surveys with Nonresponse, Wiley, Chichester, UK, 2005.
11. J.M. Brick, M.E. Jones, Metron-Int. J. Stat. LXVI (2008), 51–73.
12. M.H. Hansen, W.N. Hurwitz, J. Am. Stat. Assoc. 41 (1946), 517–529.
13. J. Neyman, J. Amer. Statist. Assoc. 33 (1938), 101–116.
14. W.G. Cochran, Sampling Techniques, third ed., John Wiley and Sons, New York, NY, USA, 1977.
15. F.C. Okafor, H. Lee, Surv. Methodol. 26 (2000), 183–188.
16. B.B. Khare, S. Srivastava, Proc. Natl. Acad. Sci. 65 (1995), 195–203.
17. R. Tabasum, I.A. Khan, J. Indian Soc. Agric. Stat. 58 (2004), 300–306.
18. R. Tabasum, I.A. Khan, Assam Stat. Rev. 20 (2006), 73–83.
19. H.P. Singh, S. Kumar, M. Kozak, Commun. Stat. Theory Methods. 39 (2010), 791–802.
20. M. Ismail, M.Q. Shahbaz, M. Hanif, Pak. J. Stat. 27 (2011), 467–476.
21. W. Gamrot, Stat. Pap. 53 (2012), 887–894.
22. J. Shabbir, N.S. Khan, Electron. J. App. Stat. Anal. 6 (2013), 1–17.
23. A. Sanaullah, M. Noor-ul-Amin, M. Hanif, Pak. J. Stat. 31 (2015), 71–94.
24. A. Sanaullah, H.A. Ali, M. Noor ul Amin, M. Hanif, Appl. Math. Comput. 226 (2014), 541–547.
25. N. Koyuncu, C. Kadilar, J. Stat. Plan. Inference. 139 (2009), 2552–2558.

APPENDIX

Table A.1 Percent relative efficiencies (*PREs*) of estimators with respect to $\bar{y}^*_{st}$ for different values of k_h each at different rate of nonresponse under case-I using two different populations.

W_{h2}	k_h	Population No	$\bar{y}^*_{st}$	t_1	$t^1_{r(2,2)}$	$t^2_{r(2,1)}$	$t^3_{r(1,2)}$	$t^g_{r(a^{opt},b^{opt})}$	t_2	$t^1_{p(2,2)}$	$t^2_{p(2,1)}$	$t^3_{p(1,2)}$	$t^g_{p(c^{opt},d^{opt})}$
10%	2.0	1	100	316.5725	1628.54	1033.76	573.39	2702.14	—-	—-	—-	—-	—-
		2	100	—-	—-	—-	—-	—-	64.9538	143.48	104.93	50.25	187.61
	2.5	1	100	343.7379	1639.70	1092.53	596.35	2808.62	—-	—-	—-	—-	—-
		2	100	—-	—-	—-	—-	—-	61.8850	137.57	100.24	47.75	183.39
	3.0	1	100	369.9346	1649.63	1146.42	616.82	2910.62	—-	—-	—-	—-	—-
		2	100	—-	—-	—-	—-	—-	59.3792	132.68	96.39	45.72	179.94
	3.5	1	100	395.2162	1657.12	1195.26	634.99	3002.54	—-	—-	—-	—-	—-
		2	100	—-	—-	—-	—-	—-	57.2944	128.58	93.18	44.04	177.11
20%	2.0	1	100	364.5528	1782.12	1049.51	555.31	2938.03	—-	—-	—-	—-	—-
		2	100	—-	—-	—-	—-	—-	65.6052	146.53	106.00	50.27	194.31
	2.5	1	100	412.3668	1842.74	1102.43	567.12	3121.84	—-	—-	—-	—-	—-
		2	100	—-	—-	—-	—-	—-	62.9599	141.99	101.99	47.99	192.63
	3.0	1	100	457.3057	1890.61	1145.43	576.28	3278.03	—-	—-	—-	—-	—-
		2	100	—-	—-	—-	—-	—-	60.8689	138.33	98.81	46.19	191.34
	3.5	1	100	499.6209	1929.9	1181.44	583.63	3415.24	—-	—-	—-	—-	—-
		2	100	—-	—-	—-	—-	—-	59.17469	135.32	96.22	44.76	190.32
30%	2.0	1	100	382.2853	1799.28	1053.66	562.84	2958.40	—-	—-	—-	—-	—-
		2	100	—-	—-	—-	—-	—-	76.5228	155.81	117.11	58.95	195.27
	2.5	1	100	436.9706	1860.79	1103.44	576.17	3135.01	—-	—-	—-	—-	—-
		2	100	—-	—-	—-	—-	—-	77.3685	154.62	116.79	59.39	194.52
	3.0	1	100	487.7213	1908.43	1142.93	586.26	3283.78	—-	—-	—-	—-	—-
		2	100	—-	—-	—-	—-	—-	78.0201	153.72	116.54	59.73	194.17
	3.5	1	100	534.9462	1946.26	1174.94	594.19	3409.55	—-	—-	—-	—-	—-
		2	100	—-	—-	—-	—-	—-	78.5378	153.03	116.35	59.99	194.09

(—-) shows data is not applicable.

Table A.2 *MSEs* of the estimators for different values of k_h each at different rate of nonresponse under case-I using two different populations.

W_{h2}	k_h	Population No	$\bar{y}^*_{st}$	t_1	$t^1_{r(2,2)}$	$t^2_{r(2,1)}$	$t^3_{r(1,2)}$	$t^g_{r(a^{opt},b^{opt})}$	t_2	$t^1_{p(2,2)}$	$t^2_{p(2,1)}$	$t^3_{p(1,2)}$	$t^g_{p(c^{opt},d^{opt})}$
10%	2.0	1	2144.00	677.254	131.65	207.39	373.91	79.34	—-	—-	—-	—-	—-
		2	5.09881	—-	—-	—-	—-	—-	7.8499	3.5536	4.8592	10.1473	2.7177
	2.5	1	2370.93	689.749	144.59	217.01	397.57	84.41	—-	—-	—-	—-	—-
		2	5.40353	—-	—-	—-	—-	—-	8.7316	3.9278	5.3903	11.3159	2.9464
	3.0	1	2597.86	702.248	157.47	226.60	421.16	89.25	—-	—-	—-	—-	—-
		2	5.70825	—-	—-	—-	—-	—-	9.6132	4.3023	5.9219	12.4855	3.1722
	3.5	1	2824.79	714.745	170.46	236.33	444.85	94.07	—-	—-	—-	—-	—-
		2	6.01298	—-	—-	—-	—-	—-	10.4949	4.6763	6.4527	13.6534	3.3949
20%	2.0	1	2540.35	696.839	142.54	242.05	457.45	86.46	—-	—-	—-	—-	—-
		2	5.43362	—-	—-	—-	—-	—-	8.2823	3.7081	5.1259	10.8084	2.7964
	2.5	1	2965.45	719.129	160.92	268.99	522.89	94.99	—-	—-	—-	—-	—-
		2	5.90575	—-	—-	—-	—-	—-	9.3802	4.1593	5.7902	12.3072	3.0658
	3.0	1	3390.55	741.419	179.33	296.01	588.35	103.43	—-	—-	—-	—-	—-
		2	6.37788	—-	—-	—-	—-	—-	10.4781	4.6106	6.4545	13.8060	3.3331
	3.5	1	3815.66	763.711	197.71	322.96	653.77	111.724	—-	—-	—-	—-	—-
		2	6.85001	—-	—-	—-	—-	—-	11.5759	5.0621	7.1191	15.3055	3.5992
30%	2.0	1	2703.11	707.092	150.23	256.54	480.26	91.37	—-	—-	—-	—-	—-
		2	6.62876	—-	—-	—-	—-	—-	8.6625	4.2542	5.6600	11.2445	3.3946
	2.5	1	3209.59	734.510	172.48	290.87	557.05	102.37	—-	—-	—-	—-	—-
		2	7.69847	—-	—-	—-	—-	—-	9.9504	4.9788	6.5917	12.9614	3.9575
	3.0	1	3716.08	761.927	194.719	325.13	633.85	113.16	—-	—-	—-	—-	—-
		2	8.76816	—-	—-	—-	—-	—-	11.2383	5.7040	7.5240	14.6804	4.5158
	3.5	1	4222.56	789.343	216.95	359.38	710.64	123.84	—-	—-	—-	—-	—-
		2	9.83786	—-	—-	—-	—-	—-	12.5263	6.4287	8.4558	16.3981	5.0687

(—-) shows data is not applicable.

Table A.3 Data statistics.

		Population-I						Population-II	
Stratum (h)		**1**	**2**	**3**	**4**	**5**	**6**	**1**	**2**
Stratified Mean, SDs and Correlation Coefficients	N_h	127	117	103	*170*	*205*	*201*	6940	*1678*
	n_h	31	21	29	38	22	39	750	181
	n'_h	70	50	75	95	70	90	1874	453
	S_{yh}	883.84	644.92	1033.4	810.58	403.65	711.72	21.4256	22.1319
	S_{xh}	30486.7	15180.77	27549.69	18218.93	8497.77	23094.14	16625.33	12861.40
	S_{zh}	555.58	365.46	612.95	458.03	260.85	397.05	19394.09	16143.74
	$\overline{Y}_h$	703.74	413	573.17	424.66	267.03	393.84	47.9805	48.0556
	$\overline{X}_h$	20804.59	9211.79	14309..30	9478.85	5569.95	12997.59	18746.55	14303.98
	$\overline{Z}_h$	498.28	318.33	431.36	311.32	227.20	313.71	19124.75	14742.47
	ρ_{xyh}	0.9360	0.996	0.994	0.983	0.989	0.965	−0.4777	−0.4406
	ρ_{xzh}	0.9396	0.9696	0.9770	0.9640	0.9670	0.9960	0.9138	0.8035
	ρ_{yzh}	0.9790	0.976	0.984	0.983	0.964	0.983	−0.4422	−0.3547
W_h = 10% Nonresponse	S_{yh2}	510.57	386.77	1872.88	1603.3	264.19	497.84	20.4752	21.7407
	S_{xh2}	9446.93	9198.29	52429.99	34794.9	4972.56	12485.10	18121.44	15492.72
	S_{zh2}	303.92	278.51	960.71	821.29	190.85	287.99	22010.50	20204.85
	ρ_{xy2}	0.9961	0.9975	0.9998	0.9741	0.995	0.9284	−0.4826	−0.5422
	ρ_{xz2}	0.9901	0.9895	0.9964	0.9609	0.9865	0.9752	0.8566	0.7691
	ρ_{yz2}	0.9931	0.9871	0.99716	0.9942	0.985	0.9647	−0.3922	−0.3181
W_h = 20% Nonresponse	S_{yh2}	396.77	406.15	1654.4	1333.35	335.83	903.91	20.7359	22.6272
	S_{xh2}	7439.16	8880.46	45784.78	29219.3	6540.43	28411.44	16155.37	13887.44
	S_{zh2}	244.56	274.42	965.42	680.28	214.49	469.86	19251.39	17323.10
	ρ_{xy2}	0.9954	0.9931	0.996	0.9761	0.9966	0.9869	−0.4870	−0.4880
	ρ_{xz2}	0.9897	0.9884	0.9789	0.9629	0.982	0.9825	0.8845	0.8399
	ρ_{yz2}	0.9898	0.9798	0.9846	0.994	0.9818	0.9874	−0.4293	−0.3304
W_h = 30% Nonresponse	S_{yh2}	500.26	356.95	1383.7	1193.47	289.41	825.24	21.4660	22.4381
	S_{xh2}	14017.994	7812.00	38379.77	26090.6	5611.32	24571.95	16877.33	12852.95
	S_{zh2}	284.4409	247.6279	811.21	631.28	188.30	437.90	19985.52	16007.36
	ρ_{xy2}	0.9639	0.9919	0.9955	0.9801	0.9961	0.9746	−0.4808	−0.4395
	ρ_{xz2}	0.9107	0.9848	0.9771	0.9650	0.9794	0.9642	0.8939	0.8298
	ρ_{yz2}	0.9739	0.9793	0.9839	0.9904	0.9799	0.9829	−0.4347	−0.2823