Estimation for finite population variance in double sampling

G. Diana, C. Tommasi

2001.22

Dipartimento di Scienze Statistiche
Università degli Studi
Via C. Battisti 241-243
35121 Padova

Dicembre 2001

Estimation for finite population variance in double sampling

Giancarlo Diana   Chiara Tommasi

Abstract

Using two-phase sampling scheme, we propose general classes of estimators for finite population variance. These classes differ for the involved sample means and variances of two auxiliary variables. For each class we provide the minimum MSE and show that there exists a chain regression type estimator which reaches this minimum. This estimator, in the widest class, is better than estimators proposed by Singh (1991) and Gupta *et al.* (1992), which are optimal in some sub-classes.

keywords: two-phase sampling – auxiliary variable – regression type estimator.

1 Introduction

In this paper we propose general classes of estimators for the finite population variance in double sampling.
If X and Z are two auxiliary variables, then the possible auxiliary quantities that can be computed in a two-phase sampling are, the first and the second phase means and variances of X and Z.
A specific class of estimators is identified by the involved auxiliary quantities. Thus, we may have as many classes as the possible choices of auxiliary quantities. In each class we can obtain an estimator which minimizes the MSE (up to terms of order n^{-1}). This estimator has a good interpretation: it is a chain regression type estimator. It will be denoted as the "best" estimator in the class. Actually, any other estimator which is equivalent, at the first orded of approximation, to the best estimator is optimum as well.
In the second section we define the most general class of estimators, where all the auxiliary quantities are used. Furthermore we provide the best estimator expression and the minimum MSE.

Other authors dealt with this problem. Singh (1991) proposed a quite general estimator whose structure is only partially defined. This estimator is an unknown function of the auxiliary sample quantities for X only. On the other hand, Gupta *et al.* (1992) proposed three estimators whose structures are well defined. These estimators are in fact completely known except for the value of some parameters. But they also depend on a subset of the auxiliary quantities only. In Section 3 we show that Singh's estimator is equivalent to the best estimator in a specific class depending on X only. In the same way (*cf.* Section 4), the estimators proposed by Gupta *et al.* are equivalent to the best estimators in some well specified classes.
In Section 5 we compare the different classes of estimators and give some guidelines for choosing a specific class and its best estimator. Finally, we provide an appendix concerning the mathematical details to prove some results reported in Sections 2 and 5.

2 A general class of estimators

Let $\mathcal{U} = \{1, \ldots, i, \ldots, N\}$ be a finite population and let Y be the the study variable and X and Z two auxiliary variables taking values Y_i, X_i and Z_i for the i-th population unit. We are interested in estimating the population variance of Y, defined as

$$S_Y^2 = \frac{\sum_{i=1}^{N}(Y_i - \overline{Y})}{N-1},$$

where $\overline{Y}$ is the population mean of Y. When X and Z are related with Y but no information is available about the population mean and variance of X, the estimation of S_Y^2 can be based on double sampling. In this scheme, we assume that a preliminary large sample of n' ($n' < N$) units is drawn by a simple random sample without replacement (SRSWOR). At this phase only X and Z are measured. In a second phase a smaller sample of size n ($n < n'$) is drawn by a SRSWOR as well. At this phase all the variables Y, X and Z are observed. Let $\overline{x}'$, $\overline{z}'$ and $s_X'^2$, $s_Z'^2$ be the sample means and variances, respectively, of X and Z at the first phase. On the other hand, $\overline{y}$, $\overline{x}$ and $\overline{z}$ and s_Y^2, s_X^2 and s_Z^2 denote the sample means and variances of Y, X and Z at the second phase. The proposed class of estimators is defined as

$$\hat{S}_Y^2 = g(s_Y^2, t), \tag{1}$$

where $t = (s_X^2, \overline{x}, s_Z^2, \overline{z}, s_Z'^2, \overline{z}', s_X'^2, \overline{x}')$ and g is a function such that

a. $g : \mathcal{S} \to \mathbb{R}$ where $\mathcal{S} \in \mathbb{R}^9$ is a convex and bounded set which contains the point (S_Y^2, T), where $T = \mathrm{E}(t) = \left(S_X^2, \overline{X}, S_Z^2, \overline{Z}, S_Z^2, \overline{Z}, S_X^2, \overline{X}\right)$.

b. It is a continuous and bounded function in $\mathcal{S}$.

c. Its first and second partial derivatives are continuous and bounded in $\mathcal{S}$.

d. $g(s_Y^2, T) = s_Y^2$.

Let

$$g_0 = \left.\frac{\partial g(s_Y^2, t)}{\partial s_Y^2}\right|_{(s_Y^2, t) = (S_Y^2, T)} \quad \text{and} \quad g_i = \left.\frac{\partial g(s_Y^2, t)}{\partial t_i}\right|_{(s_Y^2, t) = (S_Y^2, T)} \qquad i = 1, \ldots, 8$$

be the partial derivatives of g with respect to the first component, s_Y^2, and other components, t_i, $i = 1, \ldots, 8$, respectively. From point (d), we have that $g\left(S_Y^2, T\right) = S_Y^2$ and $g_0 = 1$.

Expanding $\hat{S}_Y^2$ at the point (S_Y^2, T) in a second order Taylor's series we have

$$\begin{aligned} \hat{S}_Y^2 &= g\left(S_Y^2, T\right) + \left(s_Y^2 - S_Y^2\right) g_0 + \sum_{i=1}^{8} (t_i - T_i)\, g_i \\ &= s_Y^2 + \left(s_X^2 - S_X^2\right) g_1 + \left(\overline{x} - \overline{X}\right) g_2 + \left(s_Z^2 - S_Z^2\right) g_3 + \left(\overline{z} - \overline{Z}\right) g_4 \\ &+ \left(s_Z'^2 - S_Z^2\right) g_5 + \left(\overline{z}' - \overline{Z}\right) g_6 + \left(s_X'^2 - S_X^2\right) g_7 + \left(\overline{x}' - \overline{X}\right) g_8. \end{aligned} \tag{2}$$

Since the population mean and variance of X are unknown, we have to impose the following constraints

$$g_7 = -g_1 \quad \text{and} \quad g_8 = -g_2.$$

and equation (2) becomes

$$\begin{aligned} \hat{S}_Y^2 &= s_Y^2 + \left(s_X^2 - s_X'^2\right) g_1 + (\overline{x} - \overline{x}')\, g_2 + \left(s_Z^2 - S_Z^2\right) g_3 + \left(\overline{z} - \overline{Z}\right) g_4 \\ &+ \left(s_Z'^2 - S_Z^2\right) g_5 + \left(\overline{z}' - \overline{Z}\right) g_6. \end{aligned} \tag{3}$$

Any estimator which minimizes the first order approximation for MSE($\hat{S}_Y^2$) is an "optimum" estimator in the class. Thus, the "best" values, g_i^*'s, for g_i's, $i = 1, \ldots 6$, are obtained minimizing MSE($\hat{S}_Y^2$) with respect to g_i. In Appendix we provide some mathematical details to get the expression for MSE($\hat{S}_Y^2$) up to terms of $O(n^{-1})$.
Let $\delta_Y = \left(Y - \overline{Y}\right)^2$, $\delta_X = \left(X - \overline{X}\right)^2$ and $\delta_Z = \left(Z - \overline{Z}\right)^2$ and let us define the random vector variables $\tilde{X} = (\delta_X, X)'$, $\tilde{Z} = (\delta_Z, Z)'$ and $\tilde{V} = (\tilde{X}', \tilde{Z}')'$. We have that

$$\begin{bmatrix} g_1^* \\ g_2^* \\ g_3^* \\ g_4^* \end{bmatrix} = - \begin{bmatrix} \beta_{\delta_Y,\delta_X|X,\delta_Z,Z} \\ \beta_{\delta_Y,X|\delta_X,\delta_Z,Z} \\ \beta_{\delta_Y,\delta_Z|\delta_X,X,Z} \\ \beta_{\delta_Y,Z|\delta_X,X,\delta_Z} \end{bmatrix} = -S_{\tilde{V},\tilde{V}}^{-1} S_{\delta_Y,\tilde{V}},$$

where

$$S_{\delta_Y,\tilde{V}} = \begin{bmatrix} \mathrm{Cov}(\delta_Y, \delta_X) \\ \mathrm{Cov}(\delta_Y, X) \\ \mathrm{Cov}(\delta_Y, \delta_Z) \\ \mathrm{Cov}(\delta_Y, Z) \end{bmatrix} \quad \text{and} \quad S_{\tilde{V},\tilde{V}} = \begin{bmatrix} S_{\tilde{X},\tilde{X}} & S_{\tilde{X},\tilde{Z}} \\ S_{\tilde{Z},\tilde{X}} & S_{\tilde{Z},\tilde{Z}} \end{bmatrix}$$

with

$$S_{\tilde{X},\tilde{X}} = \begin{bmatrix} \mathrm{V}(\delta_X) & \mathrm{Cov}(\delta_X, X) \\ \mathrm{Cov}(\delta_X, X) & \mathrm{V}(X) \end{bmatrix}, \quad S_{\tilde{Z},\tilde{Z}} = \begin{bmatrix} \mathrm{V}(\delta_Z) & \mathrm{Cov}(\delta_Z, Z) \\ \mathrm{Cov}(\delta_Z, Z) & \mathrm{V}(Z) \end{bmatrix}$$

and

$$S_{\tilde{X},\tilde{Z}} = \begin{bmatrix} \mathrm{Cov}(\delta_X, \delta_Z) & \mathrm{Cov}(\delta_X, Z) \\ \mathrm{Cov}(X, \delta_Z) & \mathrm{Cov}(X, Z), \end{bmatrix}, \quad S_{\tilde{Z},\tilde{X}} = S'_{\tilde{X},\tilde{Z}}.$$

Moreover, if we define the following partial regression coefficients,

$$(\beta_{\delta_X,\delta_Z|Z}, \beta_{\delta_X,Z|\delta_Z})' = -S_{\tilde{Z},\tilde{Z}}^{-1} S_{\delta_X,\tilde{Z}}$$

and

$$(\beta_{X,\delta_Z|Z}, \beta_{X,Z|\delta_Z})' = -S_{\tilde{Z},\tilde{Z}}^{-1} S_{X,\tilde{Z}},$$

where

$$S_{\delta_X,\tilde{Z}} = \begin{bmatrix} \mathrm{Cov}(\delta_X, \delta_Z) \\ \mathrm{Cov}(\delta_X, Z) \end{bmatrix} \quad \text{and} \quad S_{X,\tilde{Z}} = \begin{bmatrix} \mathrm{Cov}(X, \delta_Z) \\ \mathrm{Cov}(X, Z) \end{bmatrix},$$

then

$$\begin{bmatrix} g_5^* \\ g_6^* \end{bmatrix} = \begin{bmatrix} -\beta_{\delta_Y,\delta_X|X,\delta_Z,Z}\, \beta_{\delta_X,\delta_Z|Z} - \beta_{\delta_Y,X|\delta_X,\delta_Z,Z}\, \beta_{X,\delta_Z|Z} \\ -\beta_{\delta_Y,\delta_X|X,\delta_Z,Z}\, \beta_{\delta_X,Z|\delta_Z} - \beta_{\delta_Y,X|\delta_X,\delta_Z,Z}\, \beta_{X,Z|\delta_Z} \end{bmatrix}.$$

Replacing g_i^*'s, $i = 1, \ldots, 6$, in (3) we get the following chain regression type estimator,

$$\begin{aligned} \hat{S}^2_{Y,reg} &= s_Y^2 - \beta_{\delta_Y,\delta_X|X,\delta_Z,Z}\Big\{s_X^2 - \Big[s_X'^2 - \beta_{\delta_X,\delta_Z|Z}\left(s_Z'^2 - S_Z^2\right) - \beta_{\delta_X,Z|\delta_Z}\left(\overline{z}' - \overline{Z}\right)\Big]\Big\} \\ &- \beta_{\delta_Y,X|\delta_X,\delta_Z,Z}\Big\{\overline{x} - \Big[\overline{x}' - \beta_{X,\delta_Z|Z}\left(s_Z'^2 - S_Z^2\right) - \beta_{X,Z|\delta_Z}\left(\overline{z}' - \overline{Z}\right)\Big]\Big\} \\ &- \beta_{\delta_Y,\delta_Z|\delta_X,X,Z}\left(s_Z^2 - S_Z^2\right) - \beta_{\delta_Y,Z|\delta_X,X,\delta_Z}\left(\overline{z} - \overline{Z}\right), \end{aligned}$$

whose bias is of order n^{-1}. We call $\hat{S}^2_{Y,reg}$ the "best" estimator in the class g but any other estimator which is at the first order equivalent to $\hat{S}^2_{Y,reg}$ is optimum as well.
Let us denote the minimum first order approximation of $\mathrm{MSE}(\cdot)$ as $\mathrm{MSE}^*(\cdot)$. In class g this minimum MSE (except terms of $O(n^{-2})$) has a simple form,

$$\mathrm{MSE}^*\left(\hat{S}_Y^2\right) = \mathrm{V}(\delta_Y)\left[\left(\frac{1}{n'} - \frac{1}{N}\right)\left(1 - \rho^2_{\delta_Y,\tilde{Z}}\right) + \left(\frac{1}{n} - \frac{1}{n'}\right)\left(1 - \rho^2_{\delta_Y,\tilde{V}}\right)\right] \tag{4}$$

where

$$\rho^2_{\delta_Y,\tilde{Z}} = \frac{S'_{\delta_Y,\tilde{Z}}\, S^{-1}_{\tilde{Z},\tilde{Z}}\, S_{\delta_Y,\tilde{Z}}}{\mathrm{V}(\delta_Y)} \quad \text{with} \quad S_{\delta_Y,\tilde{Z}} = \begin{bmatrix} \mathrm{Cov}(\delta_Y, \delta_Z) \\ \mathrm{Cov}(\delta_Y, Z) \end{bmatrix}$$

and

$$\rho^2_{\delta_Y,\tilde{V}} = \frac{S'_{\delta_Y,\tilde{V}}\, S^{-1}_{\tilde{V},\tilde{V}}\, S_{\delta_Y,\tilde{V}}}{\mathrm{V}(\delta_Y)}.$$

From now on $\rho_{\cdot,\cdot}$ will denote the correlation coefficient between the specified variables.

3 Using one auxiliary variable

Let us assume now that only the auxiliary variable X can be observed, as in Singh (1991). In this case, vector t in class (1) becomes $t = (s_X^2, \overline{x}, s_X'^2, \overline{x}')$ and $T = \left(S_X^2, \overline{X}, S_X^2, \overline{X}\right)$. With similar constraints as in previous section, approximation (2) is

$$\hat{S}^2_{Y0} = s_Y^2 + \left(s_X^2 - s_X'^2\right) g_1 + \left(\overline{x} - \overline{x}'\right) g_2$$

and

$$\begin{bmatrix} g_1^* \\ g_2^* \end{bmatrix} = -\begin{bmatrix} \beta_{\delta_Y,\delta_X|X} \\ \beta_{\delta_Y,X|\delta_X} \end{bmatrix} = -S_{\tilde{X},\tilde{X}}^{-1} S_{\delta_Y,\tilde{X}},$$

where

$$S_{\delta_Y,\tilde{X}} = \begin{bmatrix} \mathrm{Cov}(\delta_Y,\delta_X) \\ \mathrm{Cov}(\delta_Y,X) \end{bmatrix}.$$

In addition,

$$\mathrm{MSE}^*\left(\hat{S}_{Y0}^2\right) = \mathrm{V}(\delta_Y)\left[\frac{1}{n'} - \frac{1}{N} + \left(\frac{1}{n} - \frac{1}{n'}\right)\left(1 - \rho_{\delta_Y,\tilde{X}}^2\right)\right], \quad (5)$$

where $\rho^2_{\delta_Y,\tilde{X}}$ is defined as $\rho^2_{\delta_Y,\tilde{Z}}$ with $\tilde{X}$ instead of $\tilde{Z}$.
Replacing g_i^*'s in the expression of $\hat{S}_{Y0}^2$ we get the best estimator in this class,

$$\hat{S}_{Y0,reg}^2 = s_Y^2 - \beta_{\delta_Y,\delta_X|X}\left(s_X^2 - s_X'^2\right) - \beta_{\delta_Y,X|\delta_X}\left(\overline{x} - \overline{x}'\right). \quad (6)$$

In the paper quoted above, Singh proposes the following class of estimators for the population variance,

$$S_h^2 = s_Y^2\, h\left(\frac{\overline{x}}{\overline{x}'}, \frac{s_X^2}{s_X'^2}\right),$$

where h is a function such that $h(1,1) = 1$, which satisfies similar regularity conditions as g. It is easy proved that, at the first order of approximation, Singh's estimator is equivalent to (6). Thus, S_h^2 reaches the minimum MSE given in (5).

4 A partial use of the auxiliary variables

When we can observe both the auxiliary variables, X and Z, we may use only a few of the auxiliary quantities employed in Section 2.
For example, in (1), instead of $t = (s_X^2, \overline{x}, s_Z^2, \overline{z}, s_Z'^2, \overline{z}', s_X'^2, \overline{x}')$, we may use $t_1 = (\overline{x}, \overline{z}', \overline{x}')$, $t_2 = (s_X^2, s_Z'^2, s_X'^2)$, or $t_3 = (s_X^2, \overline{x}, s_Z'^2, \overline{z}', s_X'^2, \overline{x}')$, *cf.* Gupta *et al.* (1992).
Using suitable constraints on g_i's, we would have respectively,

1.

$$\hat{S}_{Y1}^2 = s_Y^2 + (\overline{x} - \overline{x}')\, g_1 + \left(\overline{z}' - \overline{Z}\right) g_2,$$

$$\begin{bmatrix} g_1^* \\ g_2^* \end{bmatrix} = \begin{bmatrix} -\beta_{\delta_Y,X} \\ -\beta_{\delta_Y,Z} \end{bmatrix} = \begin{bmatrix} -\mathrm{V}(X)^{-1}\,\mathrm{Cov}(\delta_Y,X) \\ -\mathrm{V}(Z)^{-1}\,\mathrm{Cov}(\delta_Y,Z) \end{bmatrix},$$

$$\mathrm{MSE}^*\left(\hat{S}_{Y1}^2\right) = \mathrm{V}(\delta_Y)\left[\left(\frac{1}{n'} - \frac{1}{N}\right)\left(1 - \rho_{\delta_Y,Z}^2\right) + \left(\frac{1}{n} - \frac{1}{n'}\right)\left(1 - \rho_{\delta_Y,X}^2\right)\right]; \quad (7)$$

2.

$$\hat{S}^2_{Y2} = s^2_Y + \left(s^2_X - s'^2_X\right) g_1 + \left(s'^2_Z - S^2_Z\right) g_2,$$

$$\begin{bmatrix} g_1^* \\ g_2^* \end{bmatrix} = \begin{bmatrix} -\beta_{\delta_Y,\delta_X} \\ -\beta_{\delta_Y,\delta_Z} \end{bmatrix} = \begin{bmatrix} -\mathrm{V}(\delta_X)^{-1}\,\mathrm{Cov}(\delta_Y,\delta_X) \\ -\mathrm{V}(\delta_Z)^{-1}\,\mathrm{Cov}(\delta_Y,\delta_Z) \end{bmatrix},$$

$$\mathrm{MSE}^*\left(\hat{S}^2_{Y2}\right) = \mathrm{V}(\delta_Y)\left[\left(\frac{1}{n'} - \frac{1}{N}\right)\left(1 - \rho^2_{\delta_Y,\delta_Z}\right) + \left(\frac{1}{n} - \frac{1}{n'}\right)\left(1 - \rho^2_{\delta_Y,\delta_X}\right)\right]; \quad (8)$$

3.

$$\hat{S}^2_{Y3} = s^2_Y + \left(s^2_X - s'^2_X\right) g_1 + (\overline{x} - \overline{x}') g_2 + \left(s'^2_Z - S^2_Z\right) g_3 + \left(\overline{z}' - \overline{Z}\right) g_4,$$

$$\begin{bmatrix} g_1^* \\ g_2^* \end{bmatrix} = \begin{bmatrix} -\beta_{\delta_Y,\delta_X|X} \\ -\beta_{\delta_Y,X|\delta_X} \end{bmatrix} = -S^{-1}_{\tilde{X},\tilde{X}}\, S_{\delta_Y,\tilde{X}}$$

and

$$\begin{bmatrix} g_3^* \\ g_4^* \end{bmatrix} = \begin{bmatrix} -\beta_{\delta_Y,\delta_Z|Z} \\ -\beta_{\delta_Y,Z|\delta_Z} \end{bmatrix} = -S^{-1}_{\tilde{Z},\tilde{Z}}\, S_{\delta_Y,\tilde{Z}},$$

$$\mathrm{MSE}^*\left(\hat{S}^2_{Y3}\right) = \mathrm{V}(\delta_Y)\left[\left(\frac{1}{n'} - \frac{1}{N}\right)\left(1 - \rho^2_{\delta_Y,\tilde{Z}}\right) + \left(\frac{1}{n} - \frac{1}{n'}\right)\left(1 - \rho^2_{\delta_Y,\tilde{X}}\right)\right]; \quad (9)$$

Replacing g_i^*'s in the expression of $\hat{S}^2_{Yk}$, $k = 1, 2, 3$, we get, in each case, the corresponding best estimator.
As previously mentioned, Gupta *et al.* (1992) use vectors t_1, t_2 and t_3 to define the following chain ratio type estimators,

$$s_1^2 = s_Y^2 \left(\frac{\overline{x}}{\overline{x}'}\right)^{\alpha} \left(\frac{\overline{z}'}{\overline{Z}}\right)^{\beta}, \quad s_2^2 = s_Y^2 \left(\frac{s_X^2}{s'^2_X}\right)^{\alpha} \left(\frac{s'^2_Z}{S^2_Z}\right)^{\beta}$$

and

$$s_3^2 = s_Y^2 \left(\frac{\overline{x}}{\overline{x}'}\right)^{\alpha} \left(\frac{s_X^2}{s'^2_X}\right)^{\beta} \left(\frac{\overline{z}'}{\overline{Z}}\right)^{\gamma} \left(\frac{s'^2_Z}{S^2_Z}\right)^{\phi}.$$

They compute the optimum values for the involved parameters minimizing the corresponding MSEs and give the minimum value for the MSEs of s_1^2 and s_2^2.
It can be shown that, at the first order of approximation, s_1^2, s_2^2 and s_3^2 are equivalent to the best estimators corresponding to cases 1–3 above. Therefore, MSEs of s_1^2, s_2^2 and s_3^2

reach minimum values (7), (8) and (9), respectively.
Estimators based on t_1, t_2 or t_3 are commonly used when it is assumed that Z is useful to improve the estimation of $\overline{X}$ and/or S_X^2 but not the one of S_Y^2. This happens when there is a strong correlation between X and Z but not between Y and Z. Actually, this hypothesis seems to justify the following parametric classes of estimators,

$$s_Y^2 + a_1 \left\{ \overline{x} - \left[\overline{x}' + a_2(\overline{z}' - \overline{Z}) \right] \right\},$$

$$s_Y^2 + b_1 \left\{ s_X^2 - \left[s_X'^2 + b_2(s_Z'^2 - S_Z^2) \right] \right\}$$

and

$$s_Y^2 + a_1 \left\{ \overline{x} - \left[\overline{x}' + a_2(\overline{z}' - \overline{Z}) \right] \right\} + b_1 \left\{ s_X^2 - \left[s_X'^2 + b_2(s_Z'^2 - S_Z^2) \right] \right\},$$

whose optimum values for a_i's and b_i's minimize the corresponding MSEs. Again, it is easy to prove that these optimum estimators coincide with the best estimators which correspond to cases 1–3 above.
Anyway the MSE* of these estimators is never less than (4), even if X and Z are strongly related while Y and Z are almost uncorrelated. Thus, we suggest to use this type of estimators only if the efficiency loss with respect to $\hat{S}^2_{Y,reg}$ is low. A discussion about this subject and more accurate comparisons are given in the next section.

5 Conclusions

In this paper we describe different classes of estimators which depend on different auxiliary quantities and provide the best chain regression type estimator in each class. But, which specific estimator should be used? A natural guide for the choice is the comparison of the MSEs. In this section we will show that the MSE given in (4), for general class (1), is the least as possible.
As mentioned in the previous section, usually the auxiliary quantities t_1, t_2 or t_3, and so the corresponding classes, are used when the auxiliary variable Z is closely related to X but, compared to X, remotely related to Y. Since

$$\rho^2_{\delta_Y,\tilde{X}} = \frac{\rho^2_{\delta_Y,\delta_X} + \rho^2_{\delta_Y,X} - 2\,\rho_{\delta_Y,\delta_X}\,\rho_{\delta_Y,X}\,\rho_{\delta_X,X}}{1 - \rho^2_{\delta_X,X}}$$

and

$$\rho^2_{\delta_Y,\tilde{Z}} = \frac{\rho^2_{\delta_Y,\delta_Z} + \rho^2_{\delta_Y,Z} - 2\,\rho_{\delta_Y,\delta_Z}\,\rho_{\delta_Y,Z}\,\rho_{\delta_Z,Z}}{1 - \rho^2_{\delta_Z,Z}},$$

it is easy to compare MSE*$(\hat{S}^2_{Y1})$ and MSE*$(\hat{S}^2_{Y2})$ against MSE*$(\hat{S}^2_{Y3})$ (*cf.* (7), (8) and (9)). MSE*$(\hat{S}^2_{Y3})$ is the least in both cases. It is less than (5) as well. But, if we compare

$\text{MSE}^*(\hat{S}^2_{Y3})$ with $\text{MSE}^*(\hat{S}^2_Y)$, where the whole vector t is used, we have that $\text{MSE}^*(\hat{S}^2_Y)$ is never greater than $\text{MSE}^*(\hat{S}^2_{Y3})$. In fact,

$$\rho^2_{\delta_Y,\tilde{V}} \geq \rho^2_{\delta_Y,\tilde{X}}$$

since

$$\begin{aligned} S'_{\delta_Y,\tilde{V}}\, S^{-1}_{\tilde{V},\tilde{V}}\, S_{\delta_Y,\tilde{V}} &= S'_{\delta_Y,\tilde{X}}\, S^{-1}_{\tilde{X},\tilde{X}}\, S_{\delta_Y,\tilde{X}} \\ &+ S'_{\delta_Y,\tilde{X}}\, S^{-1}_{\tilde{X},\tilde{X}}\, S_{\tilde{X},\tilde{Z}}\, S^{(22)}_{\tilde{V},\tilde{V}}\, S_{\tilde{Z},\tilde{X}}\, S^{-1}_{\tilde{X},\tilde{X}}\, S_{\delta_Y,\tilde{X}} \\ &- 2\, S'_{\delta_Y,\tilde{Z}}\, S^{(22)}_{\tilde{V},\tilde{V}}\, S_{\tilde{Z},\tilde{X}}\, S^{-1}_{\tilde{X},\tilde{X}}\, S_{\delta_Y,\tilde{X}} \\ &+ S'_{\delta_Y,\tilde{Z}}\, S^{(22)}_{\tilde{V},\tilde{V}}\, S_{\delta_Y,\tilde{Z}}, \end{aligned} \tag{10}$$

where $S^{(22)}_{\tilde{V},\tilde{V}}$ is the lower 2×2 sub-matrix of $S^{-1}_{\tilde{V},\tilde{V}}$ on the right and the last three terms of (10) give a quadratic form which is never less than zero.
The details to get this result are given in Appendix.
If ρ_{δ_X,δ_Z}, $\rho_{\delta_X,Z}$, ρ_{X,δ_Z} and $\rho_{X,Z}$ have the same sign, then whenever $S_{\delta_Y,\tilde{Z}} \cong 0$, the greater the absolute values of $S_{\tilde{X},\tilde{Z}}$ elements the more $\rho^2_{\delta_Y,\tilde{V}}$ is greater than $\rho^2_{\delta_Y,\tilde{X}}$. This is inconsistent with the common suggestion for using t_1, t_2 or t_3. Thus, the only reason for using the estimators described in this section is their simplicity with respect to class (1). Anyway when a simpler estimator is chosen, we recomend to compare its MSE with respect to expression (4), in order to check the efficiency loss.

6 Appendix

In this appendix we provide some technical details to get two of the previous results.

First order approximation for $\text{MSE}(\hat{S}^2_Y)$.
Let

$$\theta_1 = \frac{1}{n'} - \frac{1}{N}, \quad \theta_2 = \frac{1}{n} - \frac{1}{N}, \quad \text{and} \quad \theta = \frac{1}{n} - \frac{1}{n'}$$

and

$$\mu_{rsv} = \sum_{i=1}^{N} \frac{(Y_i - \overline{Y})^r\,(X_i - \overline{X})^s\,(Z_i - \overline{Z})^v}{N-1}, \qquad r, s, v = 0, 1, \ldots$$

In addition, let us assume N large enough so that $(N-1)/N \cong 1$.
We can derive the expression of $\text{MSE}(\hat{S}^2_Y)$ using the following expected values, which are valid up to terms of order $O(n^{-2})$,

$$\text{E}\left[(\overline{x} - \overline{x}')^2\right] \cong \theta\, \mu_{020} \cong \theta\, \text{V}(X),$$

$$E\left[\left(s_X^2 - s_X'^2\right)^2\right] \cong \theta\,(\mu_{040} - \mu_{020}^2) \cong \theta\,V(\delta_X),$$

$$E\left[\left(s_Y^2 - S_Y^2\right)(\overline{x} - \overline{x}')\right] \cong \theta\,\mu_{210} \cong \theta\,\mathrm{Cov}(\delta_Y, X),$$

$$E\left[\left(s_Y^2 - S_Y^2\right)\left(s_X^2 - s_X'^2\right)\right] \cong \theta\,(\mu_{220} - \mu_{200}\mu_{020}) \cong \theta\,\mathrm{Cov}(\delta_Y, \delta_X),$$

$$E\left[\left(s_Y^2 - S_Y^2\right)\left(\overline{z}' - \overline{Z}\right)\right] \cong \theta_1\,\mu_{201} \cong \theta_1\,\mathrm{Cov}(\delta_Y, Z),$$

$$E\left[\left(s_X^2 - s_X'^2\right)(\overline{x} - \overline{x}')\right] \cong \theta\,\mu_{030} \cong \theta\,\mathrm{Cov}(\delta_X, X),$$

$$E\left[\left(s_Z'^2 - S_Z^2\right)^2\right] \cong \theta_1\,(\mu_{004} - \mu_{002}^2) \cong \theta_1 V(\delta_Z)$$

$$E\left[\left(\overline{z}' - \overline{Z}\right)^2\right] \cong \theta_1\,\mu_{002} \cong \theta_1\,V(Z)$$

$$E\left[\left(s_Y^2 - S_Y^2\right)\left(s_Z^2 - S_Z^2\right)\right] \cong \theta_2\,(\mu_{202} - \mu_{200}\mu_{002}) \cong \theta_2\,\mathrm{Cov}(\delta_Y, \delta_Z),$$

$$E\left[\left(s_X^2 - s_X'^2\right)\left(\overline{z} - \overline{Z}\right)\right] \cong \theta\,\mu_{021} \cong \theta\,\mathrm{Cov}(\delta_X, Z),$$

$$E\left[\left(s_Y^2 - S_Y^2\right)\left(s_Z'^2 - S_Z^2\right)\right] \cong \theta_1\,(\mu_{202} - \mu_{200}\mu_{002}) \cong \theta_1\,\mathrm{Cov}(\delta_Y, \delta_Z),$$

$$E\left[\left(s_Y^2 - S_Y^2\right)\left(\overline{z} - \overline{Z}\right)\right] \cong \theta_2\,\mu_{201} \cong \theta_2\,\mathrm{Cov}(\delta_Y, Z),$$

$$E\left[(\overline{x} - \overline{x}')\left(\overline{z} - \overline{Z}\right)\right] \cong \theta\,\mu_{011} \cong \theta\,\mathrm{Cov}(X, Z),$$

$$E\left[\left(\overline{z} - \overline{Z}\right)\left(\overline{z}' - \overline{Z}\right)\right] \cong \theta_1\,\mu_{002} \cong \theta_1\,V(Z),$$

$$E\left[\left(\overline{z} - \overline{Z}\right)\left(s_Z'^2 - S_Z^2\right)\right] \cong \theta_1\,\mu_{003} \cong \theta_1\,\mathrm{Cov}(\delta_Z, Z),$$

$$E\left[\left(\overline{z}' - \overline{Z}\right)\left(s_Z'^2 - S_Z^2\right)\right] \cong \theta_1\,\mu_{003} \cong \theta_1\,\mathrm{Cov}(\delta_Z, Z).$$

$$\begin{aligned}
\mathrm{MSE}(\hat{S}_Y^2) \;\cong\;& \theta_2\,V(\delta_Y) + \theta\,V(\delta_X) g_1^2 + \theta\,V(X)\,g_2^2 + \theta_2\,V(\delta_Z) g_3^2 \\
+\;& \theta_2\,V(Z)\,g_4^2 + \theta_1\,V(\delta_Z)\,g_5^2 + \theta_1\,V(Z)\,g_6^2 \\
+\;& 2\,[\theta\,\mathrm{Cov}(\delta_X, X)\,g_1\,g_2 + \theta\,\mathrm{Cov}(\delta_X, \delta_Z)\,g_1\,g_3 + \theta\,\mathrm{Cov}(\delta_X, Z)\,g_1\,g_4 \\
+\;& \theta\,\mathrm{Cov}(\delta_Z, X)\,g_2\,g_3 + \theta\,\mathrm{Cov}(X, Z)\,g_2\,g_4 + \theta_2\,\mathrm{Cov}(\delta_Z, Z)\,g_3\,g_4 \\
+\;& \theta_1\,V(\delta_Z)\,g_3\,g_5 + \theta_1\,\mathrm{Cov}(\delta_Z, Z)\,g_3\,g_6 + \theta_1\,\mathrm{Cov}(\delta_Z, Z)\,g_4\,g_5 \\
+\;& \theta_1\,V(Z)\,g_4\,g_6 + \theta_1\,\mathrm{Cov}(\delta_Z, Z)\,g_5\,g_6]\,.
\end{aligned}$$

Equality (10).

In order to get the right side of (10) we have used the rule to invert a partitioned matrix,

$$S^{-1}_{\tilde{V},\tilde{V}} = \begin{bmatrix} S^{(11)}_{\tilde{V},\tilde{V}} & S^{(12)}_{\tilde{V},\tilde{V}} \\ S^{(21)}_{\tilde{V},\tilde{V}} & S^{(22)}_{\tilde{V},\tilde{V}} \end{bmatrix},$$

where

$$S^{(11)}_{\tilde{V},\tilde{V}} = \left(S_{\tilde{X},\tilde{X}} - S_{\tilde{X},\tilde{Z}}\, S^{-1}_{\tilde{Z},\tilde{Z}}\, S_{\tilde{Z},\tilde{X}}\right)^{-1}, \quad S^{(22)}_{\tilde{V},\tilde{V}} = \left(S_{\tilde{Z},\tilde{Z}} - S_{\tilde{Z},\tilde{X}}\, S^{-1}_{\tilde{X},\tilde{X}}\, S_{\tilde{X},\tilde{Z}}\right)^{-1},$$

$$S^{(12)}_{\tilde{V},\tilde{V}} = -S^{-1}_{\tilde{X},\tilde{X}}\, S_{\tilde{X},\tilde{Z}}\, S^{(22)}_{\tilde{V},\tilde{V}} \quad \text{and} \quad S^{(21)}_{\tilde{V},\tilde{V}} = S'^{(12)}_{\tilde{V},\tilde{V}}$$

and the following equality

$$S^{(11)}_{\tilde{V},\tilde{V}} = S^{-1}_{\tilde{X},\tilde{X}} + S^{-1}_{\tilde{X},\tilde{X}}\, S_{\tilde{X},\tilde{Z}}\, S^{(22)}_{\tilde{V},\tilde{V}}\, S_{\tilde{Z},\tilde{X}}\, S^{-1}_{\tilde{X},\tilde{X}}.$$

References

[1] Gupta RK, Singh S and Mangat NS (1992) Some chain ratio type estimators for estimating finite population variance. Aligarh J. Statist. 12-13: 65-69

[2] Singh S (1991) Estimation of finite population variance using double sampling. Aligarh J. Statist. 11: 53-56