Songklanakarin J. Sci. Technol.
35 (6), 749-760, Nov. - Dec. 2013

SJST SONGKLANAKARIN JOURNAL OF SCIENCE AND TECHNOLOGY
http://www.sjst.psu.ac.th

*Original Article*

# Estimators in simple random sampling: Searls approach

Jirawan Jitthavech* and Vichit Lorchirachoonkul

*School of Applied Statistics, National Institute of Development Administration, Bang Kapi, Bangkok, 10240 Thailand.*

Received 12 March 2013; Accepted 14 August 2013

**Abstract**

This paper investigates four new estimators in simple random sampling, biased sample mean, ratio estimator, and two linear regression estimators, using known coefficients of variation of the study variable and auxiliary variable. The properties of the new estimators are obtained. Comparisons among the new estimators, three traditional estimators, Searls sample mean, Koyuncu and Kadilar (KK) estimators, Sisodia and Dwivedi (SD) estimator, and Shabbir and Gupta (SG) estimator are undertaken. The analysis shows that the two proposed linear regression estimators are more efficient than the new biased sample mean and at least as efficient as the three traditional estimators and SD estimator. At least one of the proposed linear regression estimators is always more efficient than the new ratio estimator and Searls sample mean. From the numerical results using two data sets published in the literature, the proposed linear regression estimators are clearly more efficient than all seven existing estimators.

**Keywords:** estimator selection criteria, linear regression estimator, bias, MSE, relative efficiency, simple random sampling

## 1. Introduction

In simple random sampling where the population size is *N*, and the sample size is *n*, Searls (1964) proposed a biased sample mean with a smaller MSE than the traditional sample mean $\bar{y}$, obtained by utilizing a known coefficient of variation, $C_y$, without the finite population correction factor $\gamma = 1/n - 1/N$. However, in this paper, Searls sample mean using the finite population correction factor is defined as

$$\bar{y}'_{wy} = \frac{\bar{y}}{1+\gamma C_y^2}, \tag{1}$$

The bias and MSE of Searls sample mean, respectively, are shown to be

$$\text{Bias}(\bar{y}'_{wy}) = -\frac{\gamma C_y^2 \bar{Y}}{1+\gamma C_y^2}, \tag{2}$$

$$\text{MSE}(\bar{y}'_{wy}) = \frac{\gamma C_y^2 \bar{Y}^2}{1+\gamma C_y^2}, \tag{3}$$

It is obvious from Equation 3 that the relative efficiency of Searls sample mean $\bar{y}'_{wy}$ with respect to the traditional sample mean $\bar{y}$ is equal to $(1+\gamma C_y^2) > 1$. Without the finite population correction, the results in Equation 1-3 become the same results as in the original Searls work (Searls, 1964). However, Searls sample mean should be used under the condition $\gamma C_y^2 << 1$ to avoid a large unacceptable bias in the estimate.

When we can find an auxiliary variable *X* which is highly correlated with the study variable *Y*, the use of known auxiliary information in the ratio estimator can reduce the MSE of the sample mean. The traditional combined ratio estimator is defined as

$$\bar{y}_r = \frac{\bar{X}}{\bar{x}}\bar{y}. \tag{4}$$

The bias and MSE of $\bar{y}_r$, to first order of approximation, respectively, are given by Cochran (1977)

* Corresponding author.
Email address: jirawan@as.nida.ac.th

$$\text{Bias}(\bar{y}_r) \approx \gamma \bar{Y}\left(C_x^2 - \rho C_y C_x\right), \quad (5)$$

$$\text{MSE}(\bar{y}_r) \approx \gamma \bar{Y}^2\left(C_y^2 + C_x^2 - 2\rho C_y C_x\right), \quad (6)$$

which is less than $\text{MSE}(\bar{y})$ if $\rho > C_x/2C_y$. Motivated by Searls (1964), several researchers proposed estimators using the known coefficient of variation of the auxiliary variable, $C_x$, for estimating the population mean when the study variable is highly correlated with the auxiliary variable, for example, the Sisodia and Dwivedi (SD) ratio estimator (Sisodia and Dwivedi, 1981), which is defined as:

$$\bar{y}_{SD} = \left(\frac{\bar{X} + C_x}{\bar{x} + C_x}\right)\bar{y}. \quad (7)$$

The bias and MSE of $\bar{y}_{SD}$, to first order of approximation, respectively, are given by

$$\text{Bias}(\bar{y}_{SD}) \approx \gamma \bar{Y}\frac{\bar{X}}{\bar{X} + C_x}\left(\frac{\bar{X}C_x^2}{\bar{X} + C_x} - \rho C_y C_x\right), \quad (8)$$

$$\text{MSE}(\bar{y}_{SD}) \approx \gamma \bar{Y}^2\left[C_y^2 + \left(\frac{\bar{X}}{\bar{X} + C_x}\right)\left(\frac{\bar{X}C_x^2}{\bar{X} + C_x} - 2\rho C_y C_x\right)\right]. \quad (9)$$

By Equation 6 and 9, the relative efficiency of $\bar{y}_{SD}$ with respect to $\bar{y}_r$ is greater than unity when

$$\rho < \frac{C_x}{2C_y}\left(1 + \frac{\bar{X}}{\bar{X} + C_x}\right). \quad (10)$$

With respect to $\bar{y}$, the relative efficiency of $\bar{y}_{SD}$ is greater than unity when

$$\rho > \frac{C_x}{2C_y}\frac{\bar{X}}{\bar{X} + C_x}. \quad (11)$$

Gupta and Shabbir (2008) improved the estimation of population mean by proposing the ratio estimator

$$\bar{y}_p = \left[w_1\bar{y} + w_2(\bar{X} - \bar{x})\right]\left(\frac{\eta\bar{X} + \lambda}{\eta\bar{x} + \lambda}\right), \quad (12)$$

where $\eta$ and $\lambda$ are either constants or functions of the known parameters of auxiliary variable, $w_1$ and $w_2$ are constants. The bias and MSE of $\bar{y}_p$, to first order of approximation, respectively, are given by Koyuncu and Kadilar (2010) (KK estimator)

$$\text{Bias}(\bar{y}_p) \approx (w_1 - 1)\bar{Y} + \gamma\left[w_1\bar{Y}(\tau^2 C_x^2 - \tau\rho C_y C_x) + w_2\bar{X}\tau C_x^2\right], \quad (13)$$

$$\text{MSE}(\bar{y}_p) \approx \frac{(1 - \gamma C_x^2\tau^2)\,\text{MSE}(\bar{y}_{lr})}{(1 - \gamma C_x^2\tau^2) + \text{MSE}(\bar{y}_{lr})/\bar{Y}^2}, \quad (14)$$

where $\text{MSE}(\bar{y}_{lr}) \approx \gamma\bar{Y}^2 C_y^2(1 - \rho^2)$ (Cochran, 1977), and $\tau = \frac{\eta\bar{X}}{\eta\bar{X} + \lambda} < 1$. The values of $w_1$ and $w_2$, which minimize $\text{MSE}(\bar{y}_p)$ are given by

$$w_1 = \frac{1 - \gamma C_x^2\tau^2}{1 + \gamma[C_y^2(1 - \rho^2) - \tau^2 C_x^2]}, \quad (15)$$

$$w_2 = \frac{\bar{Y}}{\bar{X}}\left\{\tau + \frac{(1 - \gamma C_x^2\tau^2)(\rho C_y - 2\tau C_x)}{C_x[1 + \gamma C_y^2(1 - \rho^2) - \gamma C_x^2\tau^2]}\right\}. \quad (16)$$

It is clear from Equation 14 that $\text{MSE}(\bar{y}_p)$ is less than $\text{MSE}(\bar{y}_{lr})$. Table 1 gives six values of $\eta$ and $\lambda$ as suggested by Koyuncu and Kadilar (2010), where $\beta_2(x)$ is the kurtosis of the auxiliary variable.

Recently, Shabbir and Gupta (2011) proposed an exponential ratio type estimator (SG estimator)

$$\bar{y}_{SG} = k_1\bar{y} + k_2(\bar{X} - \bar{x})\exp\left[\frac{\bar{X} - \bar{x}}{(2N + 1)\bar{X} + \bar{x}}\right], \quad (17)$$

where

$$k_1 = \frac{1 - \frac{\gamma C_x^2}{8(1 + N)^2}}{1 + \gamma C_y^2(1 - \rho^2)} \quad \text{and}$$

$$k_2 = \frac{\bar{Y}}{\bar{X}}\left[\frac{1}{2(1 + N)} - k_1\left(\frac{1}{1 + N} - \rho\frac{C_y}{C_x}\right)\right].$$

The bias and MSE of $\bar{y}_{SG}$, to first order of approximation, respectively. are given by

$$\text{Bias}(\bar{y}_{SG}) \approx \bar{Y}\left\{k_1 - 1 + k_1\gamma\left[\frac{3C_x^2}{8(1 + N)^2} - \frac{\rho C_y C_x}{2(1 + N)}\right]\right\} + \frac{\gamma C_x^2 k_2\bar{X}}{2(1 + N)^2}, \quad (18)$$

Table 1. Some members of Koyuncu and Kadilar (KK) estimators in simple random sampling.

| KK Estimator | $\bar{y}_{p(0)}$ | $\bar{y}_{p(1)}$ | $\bar{y}_{p(2)}$ | $\bar{y}_{p(3)}$ | $\bar{y}_{p(4)}$ | $\bar{y}_{p(5)}$ |
|---|---|---|---|---|---|---|
| $\eta$ | 0 | 1 | 1 | 1 | $\beta_2(x)$ | $C_x$ |
| $\lambda$ | 1 | $\rho$ | $C_x$ | $\beta_2(x)$ | $C_x$ | $\beta_2(x)$ |

$$\text{MSE}(\bar{y}_{SG}) \approx \bar{Y}^2 \left\{ 1 - \frac{\gamma C_x^2}{4(1+N)^2} - \frac{\left[1 - \frac{\gamma C_x^2}{8(1+N)^2}\right]^2}{1+\gamma C_y^2(1-\rho^2)} \right\}. \quad (19)$$

This paper is an attempt to extend the powerful Searls approach to the traditional estimators using auxiliary variable in simple random sampling. In the next two sections, we suggest a new biased sample mean, a new ratio estimator and two new regression estimators in simple random sampling. Section 4 presents the comparison among the new estimators, the traditional estimators, and the KK estimators, $\bar{y}_{p(i)}$, $i = 0, 1, 2, \ldots, 5$ (Koyuncu and Kadilar, 2010), SD estimator and SG estimator. The numerical results for estimators are presented in Section 5 using two data sets from the published literature. Finally, the last section is a summary and conclusions.

## 2. New ratio estimator and new biased sample mean

When the study variable $Y$ is highly correlated with the auxiliary variable $X$, a ratio estimate is commonly used for estimating the population mean. A new ratio estimator using coefficients of variation of $Y$ and $X$, can be written as

$$\bar{y}'_{rwyx} = \frac{\bar{X}}{\bar{x}'_{wx}} \bar{y}'_{wy}, \quad (20)$$

where $\bar{y}'_{wy} = \bar{y}/(1+\gamma C_y^2)$, $\bar{x}'_{wx} = \bar{x}/(1+\gamma C_x^2)$, and $\bar{X}$ is the population mean of the auxiliary variable. The coefficients of variation, $C_y$ and $C_x$, are assumed to be known. The new ratio estimator $\bar{y}'_{rwyx}$ in Equation 20 can be expressed in terms of the traditional ratio estimator, $\bar{y}_r$, as

$$\bar{y}'_{rwyx} = \frac{1+\gamma C_x^2}{1+\gamma C_y^2} \bar{y}_r. \quad (21)$$

The bias and MSE of $\bar{y}'_{rwyx}$, to first degree of approximation, respectively, are given by

$$\text{Bias}(\bar{y}'_{rwyx}) = \frac{1+\gamma C_x^2}{1+\gamma C_y^2} \text{Bias}(\bar{y}_r) - \frac{\gamma(C_y^2 - C_x^2)}{1+\gamma C_y^2} \bar{Y}, \quad (22)$$

$$\text{MSE}(\bar{y}'_{rwyx}) \approx \left( \frac{1+\gamma C_x^2}{1+\gamma C_y^2} \right)^2 \text{MSE}(\bar{y}_r). \quad (23)$$

From Equation 23, it can be concluded that the relative efficiency of the new ratio estimator $\bar{y}'_{rwyx}$ with respect to the traditional ratio estimator, $\bar{y}_r$, is greater than unity if $C_x < C_y$. When only is used in the estimation, the new ratio estimator in Equation 20 becomes

$$\bar{y}'_{rwx} = \frac{\bar{X}}{\bar{x}'_{wx}} \bar{y}'_{wx}, \quad (24)$$

where a new biased sample mean $\bar{y}'_{wx}$, using the coefficient of variation of the auxiliary variable, is defined as

$$\bar{y}'_{wx} = \frac{\bar{y}}{1+\gamma C_x^2}. \quad (25)$$

It is obvious from Equation 24 and 25 that $\bar{y}'_{rwx}$ is actually the traditional ratio estimator since $\bar{y}'_{wx}$ and $\bar{x}'_{wx}$ use the same weight, $1+\gamma C_x^2$. The bias and MSE of $\bar{x}'_{wx}$ can be expressed in the similar form as Equation 2 and 3, respectively. The bias of $\bar{y}'_{wx}$ can also be written in the similar manner but the MSE of $\bar{y}'_{wx}$, to first order of approximation, is given by

$$\text{MSE}(\bar{y}'_{wx}) \approx \frac{\gamma C_y^2 \bar{Y}^2}{(1+\gamma C_x^2)^2}. \quad (26)$$

By Equation 3 and 26, it can be verified that the relative efficiency of $\bar{y}'_{wx}$ with respect to $\bar{y}'_{wy}$ is greater than unity if $C_y < C_x\sqrt{2+\gamma C_x^2}$.

## 3. New linear regression estimators

By following the definition of the traditional linear regression estimator (Cochran, 1977), the first new linear regression estimator is defined as

$$\bar{y}'_{lrwyx} = \bar{y}'_{wy} + b_{wyx}(\bar{X} - \bar{x}'_{wx}). \quad (27)$$

The value of $b_{wyx}$ that minimizes the variance of $\bar{y}'_{lrwyx}$ can be shown to be

$$b_{wyx} = \frac{1+\gamma C_x^2}{1+\gamma C_y^2} b, \quad (28)$$

where $b = \rho S_y / S_x$ (Cochran, 1977). By substituting $b_{wyx}$ from Equation 28 into Equation 27, the bias of $\bar{y}'_{lrwyx}$ can be written as

$$\text{Bias}(\bar{y}'_{lrwyx}) = \text{Bias}(\bar{y}'_{wy})(1 - \rho \frac{C_x}{C_y}). \quad (29)$$

The MSE of $\bar{y}'_{lrwyx}$, to first order of approximation, is given by

$$\text{MSE}(\bar{y}'_{lrwyx}) \approx \frac{\text{MSE}(\bar{y}_{lr})}{(1+\gamma C_y^2)^2}. \quad (30)$$

Again, when only $C_x$ is used in the estimation, the new linear regression estimator in Equation 27 becomes

$$\bar{y}'_{lrwx} = \bar{y}'_{wx} + b_{wx}(\bar{X} - \bar{x}'_{wx}). \quad (31)$$

Similarly, it can be shown that $b_{wx} = b$,

$$\text{Bias}(\bar{y}'_{lrwx}) = \text{Bias}(\bar{y}'_{wx})(1 - \rho \frac{C_y}{C_x}), \quad (32)$$

$$\mathrm{MSE}(\bar{y}'_{lrwx}) \approx \frac{\mathrm{MSE}(\bar{y}_{lr})}{(1+\gamma C_x^2)^2}. \quad (33)$$

From Equation 30 and 33 it can be concluded that both new linear regression estimators, $\bar{y}'_{lrwyx}$ and $\bar{y}'_{lrwx}$ are always more efficient than the traditional regression estimator, $\bar{y}_{lr}$.

## 4. Estimator comparisons

Four new estimators, $\bar{y}'_{wx}$, $\bar{y}'_{rwyx}$, $\bar{y}'_{lrwyx}$, and $\bar{y}'_{lrwx}$, have been presented in the above sections. The first part is limited to the relative efficiency comparison among four new estimators. The selected new estimators are compared with the traditional estimators, Searls sample mean, KK estimators, SD estimator, and SG estimator in the second part. The relative efficiencies of $\bar{y}'_{wx}$, $\bar{y}'_{rwyx}$, $\bar{y}'_{lrwyx}$, and $\bar{y}'_{lrwx}$ with respect to and are shown in Table 2. The comparison conditions in Table 2 can be derived directly from the MSE formulae in Section 2 and 3. The estimator $\bar{y}'_{wx}$ is eliminated from the future comparison since the relative efficiency of $\bar{y}'_{wx}$ with respect to $\bar{y}'_{lrwx}$ is always less than unity.

**Lemma 1:** If $(1+\gamma C_x^2)^2 \ge (1+\gamma C_y^2)$, then $\bar{y}'_{lrwx}$ is more efficient than $\bar{y}'_{rwyx}$ unless $\rho = C_x(1+\gamma C_x^2)^4 / C_y(1+\gamma C_y^2)^2$. Otherwise, $\bar{y}'_{lrwx}$ is more efficient than $\bar{y}'_{rwyx}$ providing that

$$\left[\rho - \frac{C_x(1+\gamma C_x^2)^4}{C_y(1+\gamma C_y^2)^2}\right]^2 > \left[\frac{(1+\gamma C_x^2)^4}{(1+\gamma C_y^2)^2} - 1\right]\left[\frac{C_x^2(1+\gamma C_x^2)^4}{C_y^2(1+\gamma C_y^2)^2} - 1\right],$$

and as efficient as $\bar{y}'_{rwyx}$ providing that $\rho$ takes on a particular value such as

$$\left[\rho - \frac{C_x(1+\gamma C_x^2)^4}{C_y(1+\gamma C_y^2)^2}\right]^2 = \left[\frac{(1+\gamma C_x^2)^4}{(1+\gamma C_y^2)^2} - 1\right]\left[\frac{C_x^2(1+\gamma C_x^2)^4}{C_y^2(1+\gamma C_y^2)^2} - 1\right],$$

which is subject to $|\rho| \le 1$.

Lemma 2: If $C_x(1+\gamma C_x^2) \ne C_y$, then $\bar{y}'_{lrwyx}$ is more efficient than $\bar{y}'_{rwyx}$ for $0 < |\rho| < 1$; otherwise, $\bar{y}'_{lrwyx}$ is as efficient as $\bar{y}'_{rwyx}$ when $\rho = \frac{C_x}{C_y}(1+\gamma C_x^2)^2$.

The proofs of Lemma 1 and 2 are shown in Appendix. From Lemma 2, it can be concluded that the relative efficiency of the linear regression estimator $\bar{y}'_{lrwyx}$ with

Table 2. Relative efficiency comparisons among the proposed estimators: .

| Estimators | $\bar{y}'_{wx}$ | $\bar{y}'_{rwyx}$ | $\bar{y}'_{lrwyx}$ |
|---|---|---|---|
| $\bar{y}'_{wx}$ | 1 | | |
| $\bar{y}'_{rwyx}$ | $\ge 1$ if $\rho \ge \frac{C_y}{2C_x}\left[1+\left(\frac{C_x}{C_y}\right)^2 - \frac{(1+\gamma C_y^2)^2}{(1+\gamma C_x^2)^4}\right]$ | 1 | |
| $\bar{y}'_{lrwyx}$ | $\ge 1$ if $\rho^2 \ge 1-\left(\frac{(1+\gamma C_y^2)}{(1+\gamma C_x^2)}\right)^2$ | $\ge 1$ if $\left[\rho - \frac{C_x}{C_y}\left(1+\gamma C_x^2\right)\right]^2 \ge \left[\left(1+\gamma C_x^2\right)^2 - 1\right]\left[\frac{C_x^2}{C_y^2}\left(1+yC_x^2\right)^2 - 1\right]$ | 1 |
| $\bar{y}'_{lrwx}$ | $>1$ for $0 < \lvert\rho\rvert \le 1$ | $\ge 1$ if $\left[\rho - \frac{C_x\left(1+\gamma C_x^2\right)^4}{C_y\left(1+\gamma C_y^2\right)}\right]^2 \ge \left[1-\frac{\left(1+\gamma C_x^2\right)^4}{\left(1+\gamma C_y^2\right)^2}\right]\left[1-\frac{C_x^2\left(1+\gamma C_x^2\right)^4}{C_y^2\left(1+\gamma C_y^2\right)^2}\right]$ | 1 |

respect to the ratio estimator $\overline{y}'_{rwyx}$ is not less than unity. Therefore, in this study, only two new estimators, $\overline{y}'_{lrwyx}$ and $\overline{y}'_{lrwx}$, are proposed. However, since these two proposed estimators have been shown to be biased, $\overline{y}'_{lrwyx}$ and $\overline{y}'_{lrwx}$, should be used under $\gamma C_y^2 < 1$ and $\gamma C_x^2 < 1$ respectively to avoid the unacceptable large bias.

Next, the two proposed estimators, $\overline{y}'_{lrwyx}$ and $\overline{y}'_{lrwx}$, are compared with seven existing estimators consisting of three traditional estimators, $\overline{y}, \overline{y}_r, \overline{y}_{lr}$, Searls sample mean $\overline{y}'_{wy}$, KK estimators $\overline{y}_p(i)$, SD estimator $\overline{y}_{SD}$, and SG estimator $\overline{y}_{SG}$. It can be concluded from Table 3 that the proposed linear regression estimators, $\overline{y}'_{lrwyx}$ and $\overline{y}'_{lrwx}$ are always more efficient than the traditional estimators, $\overline{y}$ and $\overline{y}_{lr}$. Again, as shown in Table 3, $\overline{y}'_{lrwyx}$ is always more efficient than Searls sample mean, $\overline{y}'_{wy}$.

**Lemma 3:** If $C_x(1+\gamma C_x^2) \neq C_y$, then $\overline{y}'_{lrwyx}$ is more efficient than $\overline{y}_r$ for $0 < |\rho| < 1$; otherwise, $\overline{y}'_{lrwyx}$ is as efficient as $\overline{y}_r$ when $\rho = \frac{C_x}{C_y}(1+\gamma C_x^2)^2$.

**Lemma 4:** If $C_x \neq \frac{C_y}{(1+\gamma C_y^2)}$, then $\overline{y}'_{lrwyx}$ is more efficient than $\overline{y}_r$ for $0 < |\rho| < 1$; otherwise, $\overline{y}'_{lrwyx}$ is as efficient as $\overline{y}_r$ when $\rho = \frac{C_x}{C_y}(1+\gamma C_y^2)^2$.

Table 3. Relative efficiencies of proposed estimators with respect to existing estimators.

| Estimators Existing | Proposed Estimators | |
|---|---|---|
| | $\overline{y}'_{lrwx}$ | $\overline{y}'_{lrwyx}$ |
| $\overline{y}$ | $>1$ for $0 < \|\rho\| \leq 1$ | $>1$ for $0 < \|\rho\| \leq 1$ |
| $\overline{y}_r$ | $\geq 1$ if $\left[\rho - \frac{C_x}{C_y}(1+\gamma C_x^2)^2\right]^2 \geq \left[(1+\gamma C_x^2)^2 - 1\right]\left[\frac{C_x^2}{C_y^2}(1+\gamma C_x^2)^2 - 1\right]$ | $\geq 1$ if $\left[\rho - \frac{C_x}{C_y}(1+\gamma C_y^2)^2\right]^2 \geq \left[(1+\gamma C_y^2)^2 - 1\right]\left[\frac{C_x^2}{C_y^2}(1+\gamma C_y^2)^2 - 1\right]$ |
| $\overline{y}_{lr}$ | $>1$ for $0 < \|\rho\| \leq 1$ | $>1$ for $0 < \|\rho\| \leq 1$ |
| $\overline{y}'_{wy}$ | $\geq 1$ if $\rho^2 \geq 1 - \frac{(1+\gamma C_x^2)^2}{1+\gamma C_y^2}$ | $>1$ for $0 < \|\rho\| \leq 1$ |
| $\overline{y}_p(i)$ | $\geq 1$ if $\rho^2 \geq 1 - \frac{C_x^2}{C_y^2}(1-\gamma C_x^2\tau^2)(2+\gamma C_x^2)$ | $\geq 1$ if $\rho^2 \geq 1-(1-\gamma C_x^2\tau^2)(2+\gamma C_y^2)$ |
| $\overline{y}_{SD}$ | $\geq 1$ if $\left[\rho - \frac{C_x}{C_y}\frac{\overline{X}}{\overline{X}+C_x}(1+\gamma C_x^2)^2\right]^2 \geq \left[(1+\gamma C_x^2)^2 - 1\right]\left[\frac{C_x^2}{C_y^2}\left(\frac{\overline{X}}{\overline{X}+C_x}\right)^2(1+\gamma C_x^2)^2 - 1\right]$ | $\geq 1$ if $\left[\rho - \frac{C_x}{C_y}\frac{\overline{X}}{\overline{X}+C_x}(1+\gamma C_y^2)^2\right]^2 \geq \left[(1+\gamma C_y^2)^2 - 1\right]\left[\frac{C_x^2}{C_y^2}\left(\frac{\overline{X}}{\overline{X}+C_x}\right)^2(1+\gamma C_y^2)^2 - 1\right]$ |
| $\overline{y}_{SG}$ | $\geq 1$ if $(1-\rho^2)\left\{\rho^2 + \frac{C_x^2}{C_y^2}\left[(2+\gamma C_x^2) - \frac{(1+\gamma C_x^2)^2}{4(1+N)^2}\right] - 1\right\} \geq \frac{C_x^4(1+\gamma C_x^2)^2}{64C_y^4(1+N)^4}$ | $\geq 1$ if $(1-\rho^2)\left[1+\rho^2+\gamma C_y^2 - \frac{C_x^2(1+\gamma C_y^2)^2}{4C_y^2(1+N)^2}\right] \geq \frac{C_x^4(1+\gamma C_y^2)^2}{64C_y^4(1+N)^4}$ |

**Lemma 5:** If $\gamma C_x^2\tau^2 \le \dfrac{1+\gamma C_y^2}{2+\gamma C_y^2}$, then $\bar{y}'_{lrwyx}$ is more efficient than KK estimators $\bar{y}_p(i)$, $i = 0,1,\ldots,5$, for $0<|\rho|<1$. Otherwise, $\bar{y}'_{lrwyx}$ is more efficient than KK estimators when $\rho^2 > 1-(1-\gamma C_x^2\tau^2)(2+\gamma C_y^2)$ and as efficient as KK estimators when $\rho^2 = 1-(1-\gamma C_x^2\tau^2)(2+\gamma C_y^2)$.

**Lemma 6:** If $C_y \le C_x\sqrt{(1-\gamma C_x^2\tau^2)(2+\gamma C_x^2)}$, then $\bar{y}'_{lrwx}$ is more efficient than KK estimators $\bar{y}_p(i)$, $i = 0,1,\ldots,5$, for $0<|\rho|<1$. Otherwise, $\bar{y}'_{lrwx}$ is more efficient than KK estimators when $\rho^2 > 1-\dfrac{C_x^2}{C_y^2}(1-\gamma C_x^2\tau^2)(2+\gamma C_x^2)$, and as efficient as KK estimators when $\rho^2 = 1-\dfrac{C_x^2}{C_y^2}(1-\gamma C_x^2\tau^2)(2+\gamma C_x^2)$ subject to $|\rho|\le 1$.

**Lemma 6:** If $C_y \le C_x\sqrt{(1-\gamma C_x^2\tau^2)(2+\gamma C_x^2)}$, then $\bar{y}'_{lrwx}$ is more efficient than KK estimators $\bar{y}_p(i)$, $i = 0,1,\ldots,5$, for $0<|\rho|<1$. Otherwise, $\bar{y}'_{lrwx}$ is more efficient than KK estimators when $\rho^2 > 1-\dfrac{C_x^2}{C_y^2}(1-\gamma C_x^2\tau^2)(2+\gamma C_x^2)$, and as efficient as KK estimators when $\rho^2 = 1-\dfrac{C_x^2}{C_y^2}(1-\gamma C_x^2\tau^2)(2+\gamma C_x^2)$ subject to $|\rho|\le 1$.

**Lemma 7:** If $C_x\left(\dfrac{\bar{X}}{\bar{X}+C_x}\right)(1+\gamma C_x^2) \ne C_y$, then $\bar{y}'_{lrwx}$ is more efficient than $\bar{y}_{SD}$ for $0<|\rho|<1$; otherwise, $\bar{y}'_{lrwx}$ is as efficient as $\bar{y}_{SD}$ when $\rho = \dfrac{C_x}{C_y}\left(\dfrac{\bar{X}}{\bar{X}+C_x}\right)(1+\gamma C_x^2)^2$.

**Lemma 8:** If $C_x\left(\dfrac{\bar{X}}{\bar{X}+C_x}\right) \le C_y$ and $C_x\left(\dfrac{\bar{X}}{\bar{X}+C_x}\right) \ne \dfrac{C_y}{(1+\gamma C_y^2)}$, then $\bar{y}'_{lrwyx}$ is more efficient than $\bar{y}_{SD}$ for $0<|\rho|<1$; otherwise, $\bar{y}'_{lrwyx}$ is as efficient as $\bar{y}_{SD}$ when $\rho = \dfrac{C_x}{C_y}\left(\dfrac{\bar{X}}{\bar{X}+C_x}\right)(1+\gamma C_y^2)^2$.

**Lemma 9:** If $C_x \le \dfrac{2C_y\sqrt{1+N}}{1+\gamma C_y^2}$, then $\bar{y}'_{lrwyx}$ in a large population is more efficient than $\bar{y}_{SG}$ for $0<|\rho|<1$.

**Lemma 10:** If $C_x^2/64(1+N)^2 \le C_y^2 \le C_x^2(2+\gamma C_x^2)$, then $\bar{y}'_{lrwx}$ in a large population is more efficient than $\bar{y}_{SG}$ for $0<|\rho|<1$. If $C_y^2 > C_x^2(2+\gamma C_x^2)$, then $\bar{y}'_{lrwx}$ in a large population is more efficient than $\bar{y}_{SG}$ when $\rho^2 > 1-\dfrac{C_x^2(2+\gamma C_x^2)}{C_y^2}$, and as efficient as $\bar{y}_{SG}$ when $\rho^2 = 1-\dfrac{C_x^2(2+\gamma C_x^2)}{C_y^2}$.

The proofs of Lemma 3-10 are shown in Appendix. The above analysis shows that the relative efficiencies of the two proposed linear regression estimators, $\bar{y}'_{lrwx}$ and $\bar{y}'_{lrwyx}$ with respect to three traditional estimators, $\bar{y}, \bar{y}_r, \bar{y}_{lr}$ and SD estimator $\bar{y}_{SD}$, respectively, are not less than unity. The proposed linear regression estimator $\bar{y}'_{lrwyx}$ is more efficient than Searls sample mean $\bar{y}'_{wy}$ but $\bar{y}'_{lrwx}$ is more efficient than $\bar{y}'_{wy}$ if $C_y < C_x$. Lemma 5, 6, 9 and 10 give the conditions that the two proposed linear regression estimators, $\bar{y}'_{lrwx}$ and $\bar{y}'_{lrwyx}$, are more efficient than KK estimators $\bar{y}_p(i)$ and SG estimator $\bar{y}_{SG}$.

## 5. Numerical examples

For comparison of estimators, consider two data sets from published literature. Example 1 (*see* Gupta and Shabbir, 2008): The study variable $Y$ is the level of apple production (in 1000 tons) and the auxiliary variable $X$ is the number of apple trees in 104 villages in the East Anatolia Region in 1999. The required data are summarized as follows:

$N = 104$, $n = 20$, $\bar{Y} = 6.254$, $\bar{X} = 13931.683$, $C_y = 1.866$,
$C_x = 1.653$, $\beta_2(x) = 17.516$, $\rho = 0.865$.

In Example 1, $C_x < C_y$, $\gamma C_x^2 = 0.1406 < 1$, and $\gamma C_y^2 = 0.1406 < 1$,. From Table 3, it can be concluded that $\bar{y}'_{lrwx}$ and $\bar{y}'_{lrwyx}$ are always more efficient than $\bar{y}$ and that $\bar{y}'_{lrwyx}$ is always more efficient than $\bar{y}'_{wy}$. Since the conditions in the first if parts of Lemma 7, 8, 9, and 10 are obviously satisfied, $\bar{y}'_{lrwx}$ and $\bar{y}'_{lrwyx}$ are more efficient than $\bar{y}_{SD}$ and $\bar{y}_{SG}$. Similarly, $\bar{y}'_{lrwyx}$ and $\bar{y}'_{lrwx}$ are more efficient than $\bar{y}_r$ by Lemma 3 and 4, and $\bar{y}'_{lrwyx}$ is more efficient than $\bar{y}_p(i)$, $i = 0,1,\ldots,5$, by Lemma 5. Since $(1+\gamma C_x^2)^2/(1+\gamma C_y^2)$ in this example is greater than unity, it can be concluded from Table 3 that $\bar{y}'_{lrwx}$ is more efficient than $\bar{y}'_{wy}$. The values of $\tau$ and the RHS of the condition in Lemma 6 summarized in Table 4 lead to the conclusion that the condition in the first part of Lemma 6 is satisfied for all values of $\tau$ in this example. Therefore, $\bar{y}'_{lrwx}$ is more efficient than $\bar{y}_p(i)$ $i = 0,1,\ldots,5$, by Lemma 6. In summary, $\bar{y}'_{lrwyx}$ and $\bar{y}'_{lrwx}$ are more efficient than three traditional estimators, Searls sample mean, KK estimators, SD estimator, and SG estimator. In Example 1, $\bar{y}'_{lrwyx}$ is more efficient than $\bar{y}'_{lrwx}$ since $C_y > C_x$. In other words, $\bar{y}'_{lrwyx}$ is the most efficient estimator among the compared estimators.

Example 2 (see Koyuncu and Kadilar, 2009): The study variable $Y$ is the number of teachers and the auxiliary variable $X$ is the number of students in primary and secondary schools in 923 districts of Turkey in 2007. The population statistics are given by

$$N = 923, n = 180, \bar{Y} = 436.4345, \bar{X} = 11440.4984,$$
$$C_y = 1.7183, C_x = 1.8645,$$
$$\beta_2(x) = 18.7208, \rho = 0.9543.$$

In Example 2, $C_x > C_y$, $\gamma C_y^2 = 0.0132 < 1$, and $\gamma C_x^2 = 0.0155 < 1$. By the similar reasoning as in Example 1, $\bar{y}'_{lrwyx}$ and $\bar{y}'_{lrwx}$ are more efficient than the three traditional estimators, Searls sample mean, KK estimators, SD estimator, and SG estimator. But in Example 2, $\bar{y}'_{lrwx}$ is the most efficient estimator among compared estimators since $C_y < C_x$.

The MSEs, relative efficiencies, and relative absolute biases of estimators with respect to the traditional ratio estimator $\bar{y}_r$ of Example 1 and 2 are summarized in Table 5. The numerical results are consistent with the above analysis. The relative efficiency of $\bar{y}'_{lrwyx}$ in Example 1 is highest equal to 130.32% and the relative efficiency of $\bar{y}'_{lrwx}$ in Example 2 is highest equal to 122.88%. The relative efficiencies of both proposed linear regression estimators are clearly higher than the traditional estimators, Searls sample mean, KK, SD, and SG estimators. In both examples, the relative absolute bias of $\bar{y}'_{lrwx}$ is in the order as the traditional ratio estimator.

Table 4. The values of and the RHS of the condition in Lemma 6.

| | KK Estimator | $\bar{y}_{p(0)}$ | $\bar{y}_{p(1)}$ | $\bar{y}_{p(2)}$ | $\bar{y}_{p(3)}$ | $\bar{y}_{p(4)}$ | $\bar{y}_{p(5)}$ |
|---|---|---|---|---|---|---|---|
| Ex. 1 | $\tau$ | 1.0000 | 0.9999 | 0.9999 | 0.9987 | 1.0000 | 0.9992 |
| | $C_x\sqrt{(1-\gamma C_x^2\tau^2)(2+\gamma C_x^2)}$ | 2.2650 | 2.2650 | 2.2650 | 2.2653 | 2.2650 | 2.2652 |
| Ex.2 | $\tau$ | 1.0000 | 0.9999 | 0.9998 | 0.9984 | 1.0000 | 0.9991 |
| | $C_x\sqrt{(1-\gamma C_x^2\tau^2)(2+\gamma C_x^2)}$ | 2.6264 | 2.6264 | 2.6264 | 2.6264 | 2.6264 | 2.6264 |

Table 5. Relative efficiencies of estimators with respect to $\bar{y}_r$ and their relative absolute biases.

| Estimator | Example 1 | | | Example 2 | | |
|---|---|---|---|---|---|---|
| | MSE | Relative Efficiency, % | Relative Abs. Bias,% | MSE | Relative Efficiency, % | Relative Abs. Bias,% |
| $\bar{y}'_{lrwyx}$ | 1.0644 | 130.32 | 2.88 | 218.8165 | 122.32 | 0.05 |
| $\bar{y}'_{lrwx}$ | 1.1232 | 123.50 | 0.23 | 217.8082 | 122.88 | 0.18 |
| $\bar{y}'_{rwyx}$ | 1.3145 | 105.53 | 2.40 | 268.8905 | 99.54 | 0.42 |
| $\bar{y}_r$ | 1.3871 | 100.00 | 0.26 | 267.6515 | 100.00 | 0.19 |
| $\bar{y}_{lr}$ | 1.3847 | 100.17 | 0.00 | 224.6335 | 119.15 | 0.00 |
| $\bar{y}_{p(0)}$ | 1.3374 | 103.72 | 3.42 | 224.3689 | 119.29 | 0.12 |
| $\bar{y}_{p(1)}$ | 1.3317 | 104.16 | 3.40 | 224.3647 | 119.29 | 0.12 |
| $\bar{y}_{p(2)}$ | 1.3317 | 104.16 | 3.40 | 224.3647 | 119.29 | 0.12 |
| $\bar{y}_{p(3)}$ | 1.3318 | 104.16 | 3.40 | 224.3647 | 119.29 | 0.12 |
| $\bar{y}_{p(4)}$ | 1.3317 | 104.16 | 3.40 | 224.3647 | 119.29 | 0.12 |
| $\bar{y}_{p(5)}$ | 1.3318 | 104.16 | 3.40 | 224.3647 | 119.29 | 0.12 |
| $\bar{y}_{SD}$ | 1.3871 | 100.00 | 0.26 | 267.5353 | 100.04 | 0.19 |
| $\bar{y}_{SG}$ | 1.3374 | 103.72 | 3.47 | 224.3689 | 119.29 | 0.12 |

## 6. Summary and Conclusions

In this paper, the powerful concept of Searls biased sample mean has been extended to cover the sample mean, the ratio, and linear regression estimators in simple random sampling. Four new estimators using the coefficient of variation of the auxiliary variable, $\bar{y}'_{wx}$, $\bar{y}'_{rwyx}$, $\bar{y}'_{lrwyx}$ and $\bar{y}'_{lrwx}$, have been investigated and their properties are also obtained. The selection criteria based on relative efficiency among these new estimators, $\bar{y}, \bar{y}_r, \bar{y}_{lr}$, $\bar{y}_p(i)$, $i = 0,1,\ldots,5$, $\bar{y}_{SD}$, and $\bar{y}_{SG}$ are shown in Table 2 to 3 and Lemma 1 to 10. In the relative efficiency comparisons, it is found that the new linear regression estimator $\bar{y}'_{lrwx}$ is more efficient than the new sample mean $\bar{y}'_{wx}$. The relative efficiency of $\bar{y}'_{lrwyx}$ with respect to the new ratio estimator $\bar{y}'_{rwyx}$ is not less than unity. The linear regression estimator $\bar{y}'_{lrwyx}$ is more efficient than another linear regression estimator $\bar{y}'_{lrwx}$ if $C_y > C_x$. This leads us to propose only two new linear regression estimators. The relative efficiencies of the two proposed linear regression estimators with respect to the three traditional estimators, $\bar{y}$, $\bar{y}_r$ and $\bar{y}_{lr}$, and SD estimator $\bar{y}_{SD}$, respectively, are shown to be at least equal to unity. The proposed linear regression estimator $\bar{y}'_{lrwyx}$ is always more efficient than Searls sample mean $\bar{y}'_{wy}$ but another linear regression estimator $\bar{y}'_{lrwx}$ is more efficient than Searls sample mean $\bar{y}'_{wy}$ if $\rho^2 > 1-(1+\gamma C_x^2)^2/(1+\gamma C_y^2)$ or $Y$ and $X$ are sufficiently high correlated. The estimator selection criteria among the proposed estimators, KK estimators $\bar{y}_p(i)$ and SG estimator $\bar{y}_{SG}$ are given in Lemma 5, 6, 9, and 10.

In two published data sets, the relative efficiencies of $\bar{y}'_{lrwyx}$ and $\bar{y}'_{lrwx}$ are clearly higher than those of other estimators as shown in Table5. The relative biases of $\bar{y}'_{lrwyx}$ and $\bar{y}'_{lrwx}$ are 2.88% and 0.23% in Example1 and 0.05% and 0.18% in Example 2, respectively. There is not much difference between all KK estimators and SG estimator as shown in Table 5.

## References

Cochran, W.G. 1977. Sampling Techniques, John Wiley and Sons, New York, U.S.A.

Gupta, S. and Shabbir, J. 2008. On Improvement in Estimating the Population Mean in Simple Random Sampling. Journal of Applied Statistics. 35, 559-566.

Koyuncu, N. and Kadilar, C. 2009. Efficient Estimators for the Population Mean. Hacettepe Journal of Mathematics and Statistics. 38, 217-225.

Koyuncu, N. and Kadilar, C. 2010. On Improvement in Estimating Population Mean in Stratified Random Sampling. Journal of Applied Statistics. 37, 999-1013.

Searls, D.T. 1964. The Utilization of a Known Coefficient of Variation in the Estimation Procedure. Journal of the American Statistical Association. 59, 1225-1226.

Shabbir, J. and Gupta, S. 2011. On Estimating Finite Population Mean in Simple and Stratified Sampling. Communications in Statistics - Theory and Methods. 40, 199-212.

Sisodia, B.V.S., and Dwivedi, V.K. 1981. A Modified Ratio Estimator using Known Coefficient of Variation of Auxiliary Variable. Journal of the Indian Society of Agricultural Statistics, 33, 13-18.

## Appendix

**Proof of Lemma 1:**

From Table 2, $\bar{y}'_{lrwx}$ is at least as efficient as $\bar{y}'_{rwyx}$ if

$$\left[\rho - \frac{C_x(1+\gamma C_x^2)^4}{C_y(1+\gamma C_y^2)^2}\right]^2 \geq \left[\frac{(1+\gamma C_x^2)^4}{(1+\gamma C_y^2)^2} - 1\right]\left[\frac{C_x^2(1+\gamma C_x^2)^4}{C_y^2(1+\gamma C_y^2)^2} - 1\right]. \quad \text{(A-1)}$$

Consider the first case where $(1+\gamma C_x^2)^2 > (1+\gamma C_y^2)$. Under this condition, if $C_x(1+\gamma C_x^2)^2 < C_y(1+\gamma C_y^2)$, then the RHS in Equation A-1 is negative. It can be concluded that $\bar{y}'_{lrwx}$ is more efficient than $\bar{y}'_{rwyx}$ for $0 < |\rho| < 1$. If $C_x(1+\gamma C_x^2)^2 = C_y(1+\gamma C_y^2)$, then the RHS is equal to zero. From Equation A-1, $\bar{y}'_{lrwx}$ is more efficient than $\bar{y}'_{rwyx}$ unless $\rho = (1+\gamma C_x^2)^2 / (1+\gamma C_y^2)$. When $\rho = (1+\gamma C_x^2)^2 / (1+\gamma C_y^2)$, $\bar{y}'_{lrwx}$ is as efficient as $\bar{y}'_{rwyx}$. If $C_x(1+\gamma C_x^2)^2 > C_y(1+\gamma C_y^2)$, it is obvious that $C_x(1+\gamma C_x^2)^4 > C_y(1+\gamma C_y^2)^2$.

When $C_x(1+\gamma C_x^2)^4 > C_y(1+\gamma C_y^2)^2$, the minimum of the LHS in Equation A-1 for $0 < |\rho| \leq 1$ occurs at $\rho = 1$ and is equal to

$$LHS_{1_{\min}} = \left[1 - \frac{C_x(1+\gamma C_x^2)^4}{C_y(1+\gamma C_y^2)^2}\right]^2.$$

Consider the difference between $LHS_{1_{\min}}$ and the RHS in Equation A-1:

$$LHS_{1_{\min}} - \left[\frac{(1+\gamma C_x^2)^4}{(1+\gamma C_y^2)^2} - 1\right]\left[\frac{C_x^2(1+\gamma C_x^2)^4}{C_y^2(1+\gamma C_y^2)^2} - 1\right] = \frac{(1+\gamma C_x^2)^4}{(1+\gamma C_y^2)^2}\left(\frac{C_x}{C_y} - 1\right)^2 \geq 0.$$

Therefore, if $C_x(1+\gamma C_x^2)^2 > C_y(1+\gamma C_y^2)$, it can be concluded that in Equation A-1, the LHS is greater than the RHS for $0 < |\rho| < 1$. In other words, if $C_x(1+\gamma C_x^2)^2 > C_y(1+\gamma C_y^2)$, $\bar{y}'_{lrwx}$ is more efficient than $\bar{y}'_{rwyx}$ for $0 < |\rho| < 1$. Next consider the case where $(1+\gamma C_x^2)^2 = (1+\gamma C_y^2)$. Under this condition, the RHS in Equation A-1 is equal to zero. From Equation A-1, it can be concluded that $\bar{y}'_{lrwx}$ is more efficient than $\bar{y}'_{rwyx}$ unless $\rho = C_x / C_y$. When $\rho = C_x / C_y$, $\bar{y}'_{lrwx}$ is as efficient as $\bar{y}'_{rwyx}$. The final case is $(1+\gamma C_x^2)^2 < (1+\gamma C_y^2)$. Under this condition, it follows immediately that $C_y > C_x\sqrt{2+\gamma C_x^2}$. In other words, if $(1+\gamma C_x^2)^2 < (1+\gamma C_y^2)$, it implies that $C_x(1+\gamma C_x^2)^2 < C_y(1+\gamma C_y^2)$ and $C_x(1+\gamma C_x^2)^4 < C_y(1+\gamma C_y^2)^2$. Therefore, from Equation A-1, $\bar{y}'_{lrwx}$ is more efficient than $\bar{y}'_{rwyx}$ if the inequality of Equation A-1 is a strictly inequality and is as efficient as $\bar{y}'_{rwyx}$ if $\rho$ takes on a particular value such that the inequality (A-1) becomes an equality. In summary, if $(1+\gamma C_x^2)^2 \geq (1+\gamma C_y^2)$, then $\bar{y}'_{lrwx}$ is more efficient than $\bar{y}'_{rwyx}$ unless $\rho = C_x(1+\gamma C_x^2)^4 / C_y(1+\gamma C_y^2)^2$. Otherwise, $\bar{y}'_{lrwx}$ is more efficient than $\bar{y}'_{rwyx}$ provided that Equation A-1 is a strictly inequality and as efficient as $\bar{y}'_{rwyx}$ provided that $\rho$ takes on a particular value such that the inequality of Equation A-1 becomes an equality.

**Proof of Lemma 2:**

From Table 2, the condition that $\bar{y}'_{lrwyx}$ is at least as efficient as $\bar{y}'_{rwyx}$ can be written as

$$\left[\rho - \frac{C_x}{C_y}(1+\gamma C_x^2)^2\right]^2 \geq \left[(1+\gamma C_x^2)^2 - 1\right]\left[\frac{C_x^2}{C_y^2}(1+\gamma C_x^2)^2 - 1\right]. \quad \text{(A-2)}$$

The inequality of Equation A-2 is the same as the inequality of Equation A-1 if the term $(1+\gamma C_x^2)$ is replaced by the term $\frac{(1+\gamma C_x^2)^2}{(1+\gamma C_y^2)}$. The proof of Lemma 2 is much simpler than the proof of Lemma 1 since $(1+\gamma C_x^2)$ is only greater than zero. It can be shown in the similar manner as in the first part of the proof of Lemma 1 that if $C_x(1+\gamma C_x^2) \neq C_y$, $\bar{y}'_{lrwyx}$ is more efficient than $\bar{y}'_{rwyx}$ for $0 < |\rho| < 1$; otherwise, $\bar{y}'_{lrwyx}$ is as efficient as $\bar{y}'_{rwyx}$ when $\rho = \frac{C_x}{C_y}(1+\gamma C_x^2)^2$.

**Proof of Lemma 3:**

From Table 3, $\bar{y}'_{lrwx}$ is more efficient than $\bar{y}_r$ if

$$\left[\rho - \frac{C_x}{C_y}(1+\gamma C_x^2)^2\right]^2 \geq \left[(1+\gamma C_x^2)^2 - 1\right]\left[\frac{C_x^2}{C_y^2}(1+\gamma C_x^2)^2 - 1\right]. \qquad \text{(A-3)}$$

Since the inequality of Equation A-3 is exactly the same as the inequality of Equation A-2, the conclusion is the same as Lemma 2.

**Proof of Lemma 4:**

From Table 3, $\bar{y}'_{lrwyx}$ is more efficient than $\bar{y}_r$ if

$$\left[\rho - \frac{C_x}{C_y}(1+\gamma C_y^2)^2\right]^2 \geq \left[(1+\gamma C_y^2)^2 - 1\right]\left[\frac{C_x^2}{C_y^2}(1+\gamma C_y^2)^2 - 1\right]. \qquad \text{(A-4)}$$

The inequality of Equation A-4 is in the same form as the inequality of Equation A-2 if the term $1+\gamma C_y^2$ is replaced by $1+\gamma C_x^2$. Therefore, it can be concluded that if $C_x \neq \frac{C_y}{(1+\gamma C_y^2)}$, $\bar{y}'_{lrwyx}$ is more efficient than $\bar{y}_r$ for $0 < |\rho| < 1$; otherwise, $\bar{y}'_{lrwyx}$ is as $\bar{y}_r$ when $\rho = \frac{C_x}{C_y}(1+\gamma C_y^2)^2$.

**Proof of Lemma 5:**

From Table 3, $\bar{y}'_{lrwyx}$ is at least as efficient as $\bar{y}_p(i)$ if

$$\rho^2 \geq 1-(1-\gamma C_x^2\tau^2)(2+\gamma C_y^2). \qquad \text{(A-5)}$$

If $\gamma C_x^2\tau^2 \leq \frac{1+\gamma C_y^2}{2+\gamma C_y^2}$, it can be shown that the RHS in Equation A-5 is less than or equal to zero. Therefore, it can be concluded that if $\gamma C_x^2\tau^2 \leq \frac{1+\gamma C_y^2}{2+\gamma C_y^2}$, $\bar{y}'_{lrwyx}$ is more efficient than KK estimators $\bar{y}_p(i)$, $i = 0,1,\cdots,5$, for $0 < |\rho| < 1$. Otherwise, $\bar{y}'_{lrwyx}$ is more efficient than KK estimators when $\rho^2 > 1-(1-\gamma C_x^2\tau^2)(2+\gamma C_y^2)$, and as efficient as KK estimators when $\rho^2 = 1-(1-\gamma C_x^2\tau^2)(2+\gamma C_y^2)$.

**Proof of Lemma 6:**

From Table 3, the condition that $\bar{y}'_{lrwx}$ is at least as efficient as $\bar{y}_p(i)$ is

$$\rho^2 \geq 1 - \frac{C_x^2}{C_y^2}(1-\gamma C_x^2\tau^2)(2+\gamma C_x^2). \qquad \text{(A-6)}$$

If $C_y \le C_x\sqrt{(1-\gamma C_x^2\tau^2)(2+\gamma C_x^2)}$, the RHS in Equation A-6 is either negative or zero. Therefore, $\bar{y}'_{lrwx}$ is more efficient than KK estimators $\bar{y}_p(i)$, $i = 0,1,\cdots,5$, for $0 < |\rho| < 1$. Otherwise, $\bar{y}'_{lrwx}$ is more efficient than KK estimators when the inequality of Equation A-6 is a strictly inequality and as efficient as KK estimators when the inequality of Equation A-6 becomes an equality.

**Proof of Lemma 7:**

From Table 3, the condition that $\bar{y}'_{lrwx}$ is at least as efficient as $\bar{y}_{SD}$ is

$$\left[\rho - \frac{C_x}{C_y}\frac{\bar{X}}{\bar{X}+C_x}(1+\gamma C_x^2)^2\right]^2 \ge \left[(1+\gamma C_x^2)^2 - 1\right]\left[\frac{C_x^2}{C_y^2}\left(\frac{\bar{X}}{\bar{X}+C_x}\right)^2(1+\gamma C_x^2)^2 - 1\right]. \quad \text{(A-7)}$$

The inequality of Equation A-7 is exactly the same as the inequality of Equation A-2 if the term $C_x\left(\frac{\bar{X}}{\bar{X}+C_x}\right)$ is replaced by the term $C_x$. It can be shown in the similar manner that if $C_x\left(\frac{\bar{X}}{\bar{X}+C_x}\right)(1+\gamma C_x^2) \ne C_y$ $\bar{y}'_{lrwx}$ is more efficient than $\bar{y}_{SD}$ for $0 < |\rho| < 1$; otherwise, $\bar{y}'_{lrwx}$ is as efficient as $\bar{y}_{SD}$ when $\rho = \frac{C_x}{C_y}\left(\frac{\bar{X}}{\bar{X}+C_x}\right)(1+\gamma C_x^2)^2$.

**Proof of Lemma 8:**

Again, from Table 3, the condition that $\bar{y}'_{lrwyx}$ is at least as efficient as $\bar{y}_{SD}$ is

$$\left[\rho - \frac{C_x}{C_y}\frac{\bar{X}}{\bar{X}+C_x}(1+\gamma C_y^2)^2\right]^2 > \left[(1+\gamma C_y^2)^2 - 1\right]\left[\frac{C_x^2}{C_y^2}\left(\frac{\bar{X}}{\bar{X}+C_x}\right)^2(1+\gamma C_y^2)^2 - 1\right]. \quad \text{(A-8)}$$

The inequality of Equation A-8 is in the same form as the inequality of Equation A-7 if the term $1+\gamma C_y^2$ is replaced by $1+\gamma C_x^2$. Therefore, it can be concluded that if $C_x\left(\frac{\bar{X}}{\bar{X}+C_x}\right) \ne \frac{C_y}{(1+\gamma C_y^2)}$, $\bar{y}'_{lrwyx}$ is more efficient than $\bar{y}_{SD}$ for $0 < |\rho| < 1$; otherwise, $\bar{y}'_{lrwyx}$ is as efficient as $\bar{y}_{SD}$ when $\rho = \frac{C_x}{C_y}\left(\frac{\bar{X}}{\bar{X}+C_x}\right)(1+\gamma C_y^2)^2$.

**Proof of Lemma 9:**

From Table 3, the condition that $\bar{y}'_{lrwyx}$ is at least as efficient as $\bar{y}_{SG}$ is

$$(1-\rho^2)\left[\rho^2 + 1 + \gamma C_y^2 - \frac{C_x^2}{C_y^2}\frac{(1+\gamma C_y^2)^2}{4(1+N)^2}\right] \ge \frac{C_x^4(1+\gamma C_y^2)^2}{64C_y^4(1+N)^4}. \quad \text{(A-9)}$$

If $C_x \le \frac{2C_y\sqrt{1+N}}{1+\gamma C_y^2}$, the terms, $\frac{C_x^2(1+\gamma C_y^2)^2}{4C_y^2(1+N)^2}$ and $\frac{C_x^2(1+\gamma C_y^2)^2}{64C_y^2(1+N)^4}$ can be approximately equal to zero for large *N*.

Under such condition and large *N*, it can be concluded that the LHS in Equation A-9 is greater than zero for $0 < |\rho| < 1$. Therefore, $\bar{y}'_{lrwyx}$ in a large population is more efficient than $\bar{y}_{SG}$ for $0 < |\rho| < 1$ if $C_x \le \dfrac{2C_y\sqrt{1+N}}{1+\gamma C_y^2}$.

**Proof of Lemma 10:**

From Table 3, $\bar{y}'_{lrwx}$ is at least as efficient as $\bar{y}_{SG}$ if

$$(1-\rho^2)\left[\rho^2 + \frac{C_x^2}{C_y^2}\left[(2+\gamma C_x^2) - \frac{(1+\gamma C_x^2)^2}{4(1+N)^2}\right] - 1\right] \ge \frac{C_x^4(1+\gamma C_x^2)^2}{64C_y^4(1+N)^4}. \qquad \text{(A-10)}$$

When *N* is large, the term $(1+\gamma C_x^2)^2 / 4(1+N)^2$ is insignificant when compared with 2. Under this condition, the LHS in Equation A-10 can be approximately reduced to

$$(1-\rho^2)\left[\rho^2 + \frac{C_x^2}{C_y^2}(2+\gamma C_x^2) - 1\right], \qquad \text{(A–11)}$$

which is greater than zero for $0 < |\rho| < 1$ if $C_y^2 \le C_x^2(2+\gamma C_x^2)$. When *N* is large, the RHS in Equation A-10 is approximately equal to zero if $C_x \le 8C_y(1+N)$. Therefore, if $C_x^2 / 64(1+N)^2 \le C_y^2 \le C_x^2(2+\gamma C_x^2)$, then $\bar{y}'_{lrwx}$ in a large population is more efficient than $\bar{y}_{SG}$ for $0 < |\rho| < 1$. When $C_y^2 > C_x^2(2+\gamma C_x^2)$, the condition $C_y^2 \ge C_x^2 / 64(1+N)^2$ is also satisfied for large *N*. Therefore, if $C_y^2 > C_x^2(2+\gamma C_x^2)$, $\bar{y}'_{lrwx}$ in a large population is more efficient than $\bar{y}_{SG}$ when $\rho^2 > 1 - \dfrac{C_x^2(2+\gamma C_x^2)}{C_y^2}$, and as efficient as SG estimator $\bar{y}_{SG}$ when $\rho^2 = 1 - \dfrac{C_x^2(2+\gamma C_x^2)}{C_y^2}$.