

11-1-2010

Effect of Measurement Errors on the Separate and Combined Ratio and Product Estimators in Stratified Random Sampling

Housila P. Singh

Vikram University, Ujjain, India, hpsujn@rediffmail.com

Namrata Karpe

Vikram University, namratakarpe@yahoo.com

Follow this and additional works at: <http://digitalcommons.wayne.edu/jmasm>



Part of the [Applied Statistics Commons](#), [Social and Behavioral Sciences Commons](#), and the [Statistical Theory Commons](#)

Recommended Citation

Singh, Housila P. and Karpe, Namrata (2010) "Effect of Measurement Errors on the Separate and Combined Ratio and Product Estimators in Stratified Random Sampling," *Journal of Modern Applied Statistical Methods*: Vol. 9 : Iss. 2 , Article 8.

DOI: 10.22237/jmasm/1288584420

Available at: <http://digitalcommons.wayne.edu/jmasm/vol9/iss2/8>

This Regular Article is brought to you for free and open access by the Open Access Journals at DigitalCommons@WayneState. It has been accepted for inclusion in Journal of Modern Applied Statistical Methods by an authorized editor of DigitalCommons@WayneState.

Effect of Measurement Errors on the Separate and Combined Ratio and Product Estimators in Stratified Random Sampling

Housila P. Singh Namrata Karpe
Vikram University,
Ujjain India

Separate and combined ratio, product and difference estimators are introduced for population mean μ_Y of a study variable Y using auxiliary variable X in stratified sampling when the observations are contaminated with measurement errors. The bias and mean squared error of the proposed estimators have been derived under large sample approximation and their properties are analyzed. Generalized versions of these estimators are given along with their properties.

Key words: Auxiliary variate, bias, mean squared error, measurement error, study variate.

Introduction

Statistical procedures for the analysis of data presume that observations are correct measurements for the characteristics being studied. When applied to a real world data set, it is assumed it is possible to take measurements without error on the theoretical construct of the variables. This is untenable in many applied situations when observation errors are a rule rather than an exception.

Hence, an auxiliary variable is commonly used in survey sampling to improve the precision of estimates. When auxiliary variable information is available researchers are able to utilize it in methods of estimation to obtain the most efficient estimator. Examples are ratio, product and regression estimation methods. Using auxiliary information at the estimation stage, a large number of estimation procedures for approximating the population mean μ_Y of a study variable Y have been proposed and their properties studied based on data originating under various kinds of sampling

schemes and under the supposition that observations have been recorded without error. Such an assumption may not be tenable in actual practice and data may contain observational or measurement errors due to various reasons (Cochran, 1968; Sukhatme, 1984).

Chandhok and Han (1990) have studied the properties of a ratio estimator under two sampling schemes; simple random sampling without replacement and the Mizuno scheme when measurement errors are present. Shalabh (1997) studied the properties of the classical ratio estimator in simple random sampling when the data on both the characteristics Y (study variable) and X (auxiliary variable) are subject to measurement errors. Manisha and Singh (2001), Maneesha and Singh (2002) and Singh and Karpe (2008a) have also considered the problem of estimating the population mean using auxiliary information in the presence of measurement errors. Later Singh and Karpe (2008b, 2009a, 2009c) studied the effect of measurement errors on the classes of estimators proposed for population variance and coefficient of variation. This article discusses the properties of separate and combined ratio and product estimators in stratified random sampling when the data are subject to measurement errors on both the characteristics Y and X .

Housila P. Singh is a Professor in the School of Studies in Statistics. Email: hpsujn@rediffmail.com. Namrata Karpe is a Research Scholar in the School of Studies in Statistics. Email: namratarupe@yahoo.com.

Suggested Estimators

Separate Ratio Estimator in Stratified Random Sampling in the Presence of Measurement Errors

Consider a finite population $U = (u_1, u_2, \dots, u_N)$ of size N and let Y and X respectively be the study and auxiliary variables associated with each unit $u_j = (j = 1, 2, \dots, N)$ of the population. Let the population of size N be stratified into L strata with the h^{th} stratum containing N_h units, where $h = 1, 2, \dots, L$ such that $\sum_{h=1}^L N_h = N$. A simple random sample size n_h is drawn without replacement from the h^{th} stratum such that $\sum_{h=1}^L n_h = n$. Let (y_{hi}, x_{hi}) be the observed pair values instead of true pair values (Y_{hi}, X_{hi}) of two characteristics (Y, X) on i^{th} unit of the h^{th} stratum, where $i = 1, 2, \dots, N_h$ and $h = 1, 2, \dots, L$. In addition, let:

$$\left(\bar{y}_h = \frac{1}{n_h} \sum_{i=1}^{n_h} y_{hi}, \bar{x}_h = \frac{1}{n_h} \sum_{i=1}^{n_h} x_{hi} \right),$$

$$\left(\bar{y}_{st} = \sum_{h=1}^L W_h \bar{y}_h, \bar{x}_{st} = \sum_{h=1}^L W_h \bar{x}_h \right),$$

$$\left(\mu_{Yh} = \frac{1}{N_h} \sum_{i=1}^{N_h} y_{hi}, \mu_{Xh} = \frac{1}{N_h} \sum_{i=1}^{N_h} x_{hi} \right),$$

and

$$\left(\mu_Y = \sum_{h=1}^L W_h \mu_{Yh}, \mu_X = \sum_{h=1}^L W_h \mu_{Xh} \right)$$

be the samples means and population means of (Y, X) respectively, where $W_h = \frac{N_h}{N}$ is the stratum weight. Let the observational or measurement errors be

$$u_{hi} = y_{hi} - Y_{hi} \quad (2.1)$$

and

$$v_{hi} = x_{hi} - X_{hi} \quad (2.2)$$

which are stochastic in nature and are uncorrelated with mean zero and variances σ_{Uh}^2 and σ_{Vh}^2 respectively. Further let ρ_h be the population correlation coefficient between Y and X in the h^{th} stratum.

For simplicity in exposition, assume that u_{hi} 's and v_{hi} 's are uncorrelated although (Y_{hi}, X_{hi}) are correlated; such a specification can be relaxed at the cost of some algebraic complexity. It is also assumed that the finite population correction terms $(1 - f_h)$ and $(1 - f)$ can be ignored where $f_h = \frac{n_h}{N_h}$ and

$$f = \frac{n}{N}.$$

To estimate the population mean μ_Y , the traditional unbiased estimator (i.e., stratified sample mean) \bar{y}_{st} is used, but it does not utilize the sample information on auxiliary characteristic X . Assuming that μ_{Xh} is known and is different from zero, this method yields a separate ratio estimator of the population mean μ_Y :

$$t_{RS} = \sum_{h=1}^L W_h \bar{y}_h \frac{\mu_{Xh}}{\bar{x}_h}. \quad (2.3)$$

To obtain the bias and mean squared error of $t_{RS} : \bar{y}_h = \mu_{Yh} (1 + \delta_{\bar{y}_h})$, and $\bar{x}_h = \mu_{Xh} (1 + \delta_{\bar{x}_h})$, such that

$$E(\delta_{\bar{y}_h}) = E(\delta_{\bar{x}_h}) = 0$$

$$E(\delta_{\bar{y}_h}^2) = \frac{C_{Yh}^2}{n_h} \left(1 + \frac{\sigma_{Uh}^2}{\sigma_{Yh}^2} \right) = \frac{C_{Yh}^2}{n_h \theta_{Yh}},$$

$$E(\delta_{\bar{x}_h}^2) = \frac{C_{Xh}^2}{n_h} \left(1 + \frac{\sigma_{Vh}^2}{\sigma_{Xh}^2} \right) = \frac{C_{Xh}^2}{n_h \theta_{Xh}},$$

EFFECT OF MEASUREMENT ERRORS ON PRODUCT ESTIMATORS

$$E(\delta_{\bar{y}_h} \delta_{\bar{x}_h}) = \frac{1}{n_h} \rho_h C_{Yh} C_{Xh},$$

where

$$C_{Yh} = \frac{\sigma_{Yh}}{\mu_{Yh}},$$

$$C_{Xh} = \frac{\sigma_{Xh}}{\mu_{Xh}},$$

$$\theta_{Yh} = \frac{\sigma_{Uh}^2}{\sigma_{Uh}^2 + \sigma_{Yh}^2},$$

and

$$\theta_{Xh} = \frac{\sigma_{Vh}^2}{\sigma_{Vh}^2 + \sigma_{Xh}^2}.$$

Expressing (2.3) in terms of δ 's as

$$t_{RS} = \sum_{h=1}^L W_h \mu_{Yh} (1 + \delta_{\bar{y}_h}) (1 + \delta_{\bar{x}_h})^{-1} \quad (2.4)$$

Assuming $|\delta_{\bar{x}_h}| < 1$, the right hand side of (2.4) is expanded as

$$t_{RS} = \sum_{h=1}^L W_h \mu_{Yh} (1 + \delta_{\bar{y}_h}) \begin{pmatrix} 1 - \delta_{\bar{x}_h} + \delta_{\bar{x}_h}^2 \\ -\delta_{\bar{x}_h}^3 + \dots \end{pmatrix}$$

$$= \sum_{h=1}^L W_h \mu_{Yh} \left\{ \begin{matrix} 1 + \delta_{\bar{y}_h} - \delta_{\bar{x}_h} - \delta_{\bar{y}_h} \delta_{\bar{x}_h} \\ + \delta_{\bar{x}_h}^2 + \delta_{\bar{y}_h} \delta_{\bar{x}_h}^2 - \delta_{\bar{x}_h}^3 + \dots \end{matrix} \right\}$$

Neglecting terms of δ 's having power greater than two, results in

$$t_{RS} = \sum_{h=1}^L W_h \mu_{Yh} \left\{ \begin{matrix} 1 + \delta_{\bar{y}_h} - \delta_{\bar{x}_h} - \delta_{\bar{y}_h} \delta_{\bar{x}_h} \\ + \delta_{\bar{x}_h}^2 + \delta_{\bar{y}_h} \delta_{\bar{x}_h}^2 \end{matrix} \right\},$$

$$t_{RS} = \mu_Y + \sum_{h=1}^L W_h \mu_{Yh} \left\{ \begin{matrix} \delta_{\bar{y}_h} - \delta_{\bar{x}_h} - \delta_{\bar{y}_h} \delta_{\bar{x}_h} \\ + \delta_{\bar{x}_h}^2 + \delta_{\bar{y}_h} \delta_{\bar{x}_h}^2 \end{matrix} \right\},$$

$$(t_{RS} - \mu_Y) = \sum_{h=1}^L W_h \mu_{Yh} \left\{ \begin{matrix} \delta_{\bar{y}_h} - \delta_{\bar{x}_h} - \delta_{\bar{y}_h} \delta_{\bar{x}_h} \\ + \delta_{\bar{x}_h}^2 + \delta_{\bar{y}_h} \delta_{\bar{x}_h}^2 \end{matrix} \right\}, \quad (2.5)$$

Taking the expectation of both sides of (2.5) results in the bias of t_{RS} to the first degree of approximation,

$$B(t_{RS}) = \sum_{h=1}^L W_h \mu_{Yh} \left(\frac{C_{Xh}^2}{n_h \theta_{Xh}} \right) (1 - \theta_{Xh} K_h) \quad (2.6)$$

where

$$K_h = \rho_h \left(\frac{C_{Yh}}{C_{Xh}} \right).$$

Squaring both sides of (2.5), neglecting terms of δ 's having power greater than two and then taking the expectation of both sides gives the mean squared error of t_{RS} to the first degree of approximation as

$$MSE(t_{RS}) = \sum_{h=1}^L W_h^2 \left(\frac{\mu_{Yh}^2}{n_h} \right) \left[\frac{C_{Yh}^2}{\theta_{Yh}} + \frac{C_{Xh}^2}{\theta_{Xh}} (1 - 2K_h \theta_{Xh}) \right] \quad (2.7)$$

The variance of \bar{y}_{st} is:

$$Var(\bar{y}_{st}) = \sum_{h=1}^L W_h^2 \frac{\mu_{Yh}^2 C_{Yh}^2}{n_h \theta_{Yh}} \quad (2.8)$$

and, from (2.7) and (2.8),

$$MSE(t_{RS}) - Var(\bar{y}_{st}) = \sum_{h=1}^L W_h^2 \left(\frac{\mu_{Yh}^2}{n_h} \right) \frac{C_{Xh}^2}{\theta_{Xh}} (1 - 2K_h \theta_{Xh})$$

which is less than zero if

$$\sum_{h=1}^L W_h^2 \left(\frac{\mu_{Yh}^2}{n_h} \right) \frac{C_{Xh}^2}{\theta_{Xh}} (1 - 2K_h \theta_{Xh}) < 0$$

or if $(1 - 2K_h \theta_{Xh}) < 0$

or if $\frac{1}{2\theta_{Xh}} < K_h$ (2.9)

or if $K_h > \frac{1}{2\theta_{Xh}}$

Thus, the proposed separate ratio estimator t_{RS} is more efficient than the usual unbiased estimator \bar{y}_{st} if condition (2.9) holds.

If the observations on both the variables X and Y are recorded without error, then the MSE of t_{RS} at (2.7) reduces to:

$$MSE(t_{RS})_t = \sum_{h=1}^L W_h^2 \left(\frac{\mu_{Yh}^2}{n_h} \right) [C_{Yh}^2 + C_{Xh}^2 (1 - 2K_h)]$$

(2.10)

Expression (2.10) can be obtained from (2.7) by setting $\theta_{Xh} = \theta_{Yh} = 1$. From (2.7) and (2.10):

$$MSE(t_{RS}) - MSE(t_{RS})_t = \sum_{h=1}^L W_h^2 \left(\frac{\mu_{Yh}^2}{n_h} \right) \left[\left(\frac{1 - \theta_{Yh}}{\theta_{Yh}} \right) C_{Yh}^2 + \left(\frac{1 - \theta_{Xh}}{\theta_{Xh}} \right) C_{Xh}^2 \right]$$

(2.11)

which is always positive. Thus, it follows from (2.11) that the presence of measurement errors associated with both variables are accountable for increasing the mean squared error of the separate ratio-estimator t_{RS} .

Separate Product Estimator in Stratified Random Sampling in the presence of Measurement Errors

Next, define the product estimators in stratified random sampling in the presence of measurement errors of the population mean μ_Y as

$$t_{PS} = \sum_{h=1}^L W_h \bar{y}_h \frac{\bar{x}_h}{\mu_{Xh}} \quad (2.12)$$

Express (2.12) in terms of δ' s as

$$t_{PS} = \sum_{h=1}^L W_h \mu_{Yh} (1 + \delta_{\bar{y}_h}) (1 + \delta_{\bar{x}_h}) \quad (2.13)$$

Assuming that $|\delta_{\bar{x}_h}| < 1$, the right hand side of (2.13) is expanded as

$$t_{PS} = \sum_{h=1}^L W_h \mu_{Yh} \{1 + \delta_{\bar{y}_h} + \delta_{\bar{x}_h} + \delta_{\bar{y}_h} \delta_{\bar{x}_h}\},$$

$$t_{PS} = \mu_Y + \sum_{h=1}^L W_h \mu_{Yh} \{\delta_{\bar{y}_h} + \delta_{\bar{x}_h} + \delta_{\bar{y}_h} \delta_{\bar{x}_h}\},$$

or

$$(t_{PS} - \mu_Y) = \sum_{h=1}^L W_h \mu_{Yh} \{\delta_{\bar{y}_h} + \delta_{\bar{x}_h} + \delta_{\bar{y}_h} \delta_{\bar{x}_h}\}, \quad (2.14)$$

and taking the expectation of both sides of (2.14) results in the bias of t_{PS} to the first degree of approximation,

$$B(t_{PS}) = \sum_{h=1}^L W_h \mu_{Yh} C_{Xh}^2 K_h \quad (2.15)$$

Squaring both sides of (2.14) and neglecting terms of δ' s having power greater than two and taking expectations of both sides, provides the mean squared error of t_{PS} to the first degree of approximation as

$$MSE(t_{PS}) = \sum_{h=1}^L W_h^2 \left(\frac{\mu_{Yh}^2}{n_h} \right) \left[\frac{C_{Yh}^2}{\theta_{Yh}} + \frac{C_{Xh}^2}{\theta_{Xh}} (1 + 2K_h \theta_{Xh}) \right] \quad (2.16)$$

From (2.16) and (2.8)

$$MSEP(t_{RS}) - \text{Var}(\bar{y}_{st}) = \sum_{h=1}^L W_h^2 \left(\frac{\mu_{Yh}^2}{n_h} \right) \frac{C_{Xh}^2}{\theta_{Xh}} (1 + 2K_h \theta_{Xh})$$

which is less than zero if

$$\sum_{h=1}^L W_h^2 \left(\frac{\mu_{Yh}^2}{n_h} \right) \frac{C_{Xh}^2}{\theta_{Xh}} (1 + 2K_h \theta_{Xh}) < 0$$

or if $(1 + 2K_h \theta_{Xh}) < 0$ (2.17)

or if $K_h < -\frac{1}{2\theta_{Xh}}$

Thus, the proposed separate product estimator t_{PS} is more efficient than the usual unbiased estimator \bar{y}_{st} if condition (2.17) holds.

If the observations on both the variables X and Y are recorded without error, then the MSE of the separate product estimator t_{PS} is given by

$$MSE(t_{RS})_t = \sum_{h=1}^L W_h^2 \left(\frac{\mu_{Yh}^2}{n_h} \right) [C_{Yh}^2 + C_{Xh}^2 (1 + 2K_h)]$$

(2.18)

Expression (2.18) can be obtained from (2.16) by setting $\theta_{Xh} = \theta_{Yh} = 1$. From (2.16) and (2.18):

$$MSE(t_{PS}) - MSE(t_{PS})_t = \sum_{h=1}^L W_h^2 \left(\frac{\mu_{Yh}^2}{n_h} \right) \left[\left(\frac{1 - \theta_{Yh}}{\theta_{Yh}} \right) C_{Yh}^2 + \left(\frac{1 - \theta_{Xh}}{\theta_{Xh}} \right) C_{Xh}^2 \right]$$

(2.19)

which is always positive. Thus, it follows from (2.19) that the presence of measurement errors associated with both variables are accountable for increasing the mean squared error of the separate product-estimator t_{PS} .

Separate Difference Estimator in Stratified Random Sampling in the presence of Measurement Errors

A separate difference estimator is defined in stratified random sampling in the presence of measurement errors for population mean μ_Y , as

$$t_{dS} = \sum_{h=1}^L W_h \{ \bar{y}_h + d_h (\mu_{Xh} - \bar{x}_h) \}$$

(2.20)

where d_h 's are suitably chosen constants.

It can be observed that the estimator t_{dS} is an unbiased estimator for the population mean μ_Y , and the variance of t_{dS} is given by

$$Var(t_{dS}) = \sum_{h=1}^L W_h^2 \left\{ Var(\bar{y}_h) + d_h^2 Var(\bar{x}_h) - 2Cov(\bar{y}_h, \bar{x}_h) \right\}$$

(2.21)

where

$$Var(\bar{y}_h) = \frac{\sigma_{Yh}^2 + \sigma_{Uh}^2}{n_h} = \frac{\sigma_{Yh}^2}{n_h \theta_{Yh}}$$

$$Var(\bar{x}_h) = \frac{\sigma_{Xh}^2 + \sigma_{Vh}^2}{n_h} = \frac{\sigma_{Xh}^2}{n_h \theta_{Xh}}$$

$$Cov(\bar{y}_h, \bar{x}_h) = \frac{\sigma_{YXh}}{n_h} = \frac{\sigma_{Xh}^2}{n_h} \beta_{YXh}$$

$$\sigma_{YXh} = Cov(y_{hi}, x_{hi}) = E\{ (y_{hi} - E(y_{hi})) (x_{hi} - E(x_{hi})) \}$$

and

$$\beta_{YXh} = \frac{\sigma_{YXh}}{\sigma_{Xh}^2}$$

Thus,

$$Var(t_{dS}) = \sum_{h=1}^L \frac{W_h^2}{n_h} \left\{ \frac{\sigma_{Yh}^2}{\theta_{Yh}} + d_h^2 \frac{\sigma_{Xh}^2}{\theta_{Xh}} - 2d_h \sigma_{XY}^2 \beta_{YXh} \right\}$$

$$= \sum_{h=1}^L \frac{W_h^2}{n_h} \left\{ \frac{\sigma_{Yh}^2}{\theta_{Yh}} + d_h \frac{\sigma_{Xh}^2}{\theta_{Xh}} (d_h - 2\beta_{YXh} \theta_{Xh}) \right\}$$

(2.22)

which is minimized for

$$d_h = \beta_{YXh} \theta_{Xh}$$

(2.23)

where $\beta_{YXh} = \frac{\sigma_{YXh}}{\sigma_{Xh}^2}$ is the population regression coefficient of Y on X in the h^{th} stratum. Thus, the resulting (minimum) variance of t_{dS} is given by

$$\min.Var(t_{dS}) = \sum_{h=1}^L \frac{W_h^2}{n_h} \left\{ \frac{\sigma_{Yh}^2}{\theta_{Yh}} - \beta_{YXh}^2 \theta_{Xh} \sigma_{Xh}^2 \right\}$$

or

$$\min.Var(t_{dS}) = \sum_{h=1}^L \frac{W_h^2 \sigma_{Yh}^2}{n_h \theta_{Yh}} (1 - \rho_h^2 \theta_{Xh} \theta_{Yh}) \tag{2.24}$$

When data are recorded without error for the variables X and Y , the variance of t_{dS} at (2.22) reduces to:

$$Var(t_{dS})_t = \sum_{h=1}^L \frac{W_h^2}{n_h} \left\{ \sigma_{Yh}^2 + d_h \sigma_{Xh}^2 (d_h - 2\beta_{YXh}) \right\} \tag{2.25}$$

From (2.22) and (2.25):

$$\begin{aligned} Var(t_{dS}) - Var(t_{dS})_t = \\ \sum_{h=1}^L \frac{W_h^2}{n_h} \left\{ \left(\frac{1 - \theta_{Yh}}{\theta_{Yh}} \right) \sigma_{Yh}^2 + d_h^2 \left(\frac{1 - \theta_{Xh}}{\theta_{Xh}} \right) \right\} \geq 0. \end{aligned} \tag{2.26}$$

Observe from (2.26) that the presence of measurement error in both the variables X and Y inflates the variance of t_{dS} . The expression (2.25) is minimized for:

$$d_h = \beta_{YXh} \tag{2.27}$$

and the resulting (minimum) variance of t_{dS} in the absence of measurement errors is given by

$$\min.Var(t_{dS})_t = \sum_{h=1}^L \frac{W_h^2}{n_h} \sigma_{Yh}^2 (1 - \rho_h^2) \tag{2.28}$$

From (2.24) and (2.28):

$$\begin{aligned} \min.Var(t_{dS}) - \min.Var(t_{dS})_t = \\ \sum_{h=1}^L \frac{W_h^2 \sigma_{Yh}^2}{n_h} \left\{ \left(\frac{1 - \theta_{Yh}}{\theta_{Yh}} \right) + \rho_h^2 (1 - \theta_{Xh}) \right\} \end{aligned} \tag{2.29}$$

It is observed from (2.29) that the difference [$\min.Var(t_{dS}) - \min.Var(t_{dS})_t$], is always positive. Thus, the presence of measurement errors in both variables X and Y inflates the variance of t_{dS} at optimum condition, which disturbs the optimal properties of the difference estimator t_{dS} .

A Separate Class of Estimators in Stratified Random Sampling in the Presence of Measurement Errors

Whatever the sample chosen, let (\bar{y}_h, \bar{x}_h) assume values in a bounded, closed subset, P_h , of the two-dimensional real space containing the point (μ_{Yh}, μ_{Xh}) . Following an approach similar to that adopted by Srivastava (1971, 1980) for defining a class of estimators of the population μ_{Yh} , consider the class of estimators of the population μ_Y , defined by

$$t_S = \sum_{h=1}^L W_h t_h(\bar{y}_h, \bar{x}_h), \tag{2.30}$$

where $t_h(\bar{y}_h, \bar{x}_h)$ is a function of (\bar{y}_h, \bar{x}_h)

$$t_h(\mu_{Yh}, \mu_{Xh}) = \mu_{Yh} \tag{2.31}$$

$$\begin{aligned} \Rightarrow t_{h1}(\mu_{Yh}, \mu_{Xh}) &= \left. \frac{\partial t_h(\mu_{Yh}, \mu_{Xh})}{\partial \bar{y}_h} \right|_{(\bar{y}_h, \bar{x}_h)} \\ &= (\mu_{Yh}, \mu_{Xh}) \\ &= 1 \end{aligned}$$

such that it satisfies the following conditions:

- i. The function $t_h(\bar{y}_h, \bar{x}_h)$ is continuous and bounded in P_h ; and
- ii. The first, second and third order partial derivatives of $t_h(\bar{y}_h, \bar{x}_h)$ exist and are continuous and bounded in P_h .

Expanding the function $t_h(\bar{y}_h, \bar{x}_h)$ about the point $(\bar{y}_h, \bar{x}_h) = (\mu_{Yh}, \mu_{Xh})$ in a third-order partial derivative, results in

$$t_S = \sum_{h=1}^L W_h \left[\begin{aligned} & \mu_{Yh} + (\bar{y}_h - \mu_{Yh}) t_{h1}(\mu_{Yh}, \mu_{Xh}) \\ & + (\bar{x}_h - \mu_{Xh}) t_{h2}(\mu_{Yh}, \mu_{Xh}) \\ & + \frac{1}{2} \left\{ \begin{aligned} & (\bar{y}_h - \mu_{Yh})^2 t_{h11}(\mu_{Yh}, \mu_{Xh}) \\ & + 2(\bar{y}_h - \mu_{Yh})(\bar{x}_h - \mu_{Xh}) t_{h12}(\mu_{Yh}, \mu_{Xh}) \\ & + (\bar{x}_h - \mu_{Xh})^2 t_{h22}(\mu_{Yh}, \mu_{Xh}) \end{aligned} \right\} \\ & + \frac{1}{6} \left\{ \begin{aligned} & (\bar{y}_h - \mu_{Yh})^3 t_{h111}(\bar{y}_h^*, \bar{x}_h^*) \\ & + 3(\bar{y}_h - \mu_{Yh})(\bar{x}_h - \mu_{Xh})^2 t_{h122}(\bar{y}_h^*, \bar{x}_h^*) \\ & + 3(\bar{y}_h - \mu_{Yh})^2 (\bar{x}_h - \mu_{Xh}) t_{h112}(\bar{y}_h^*, \bar{x}_h^*) \\ & + (\bar{x}_h - \mu_{Xh})^3 t_{h222}(\bar{y}_h^*, \bar{x}_h^*) \end{aligned} \right\} \end{aligned} \right] \quad (2.32)$$

where $\bar{y}_h^* = \{\mu_{Yh} + \theta(\bar{y}_h - \mu_{Yh})\}$, $\bar{x}_h^* = \{\mu_{Xh} + \theta(\bar{x}_h - \mu_{Xh})\}$, $0 < \theta < 1$ and θ may depend on $(\bar{y}_h^*, \bar{x}_h^*)$, and $t_{hijk}(\bar{y}_h^*, \bar{x}_h^*)$ denotes the third order partial derivative of $t_h(\bar{y}_h^*, \bar{x}_h^*)$ with respect to (\bar{y}_h, \bar{x}_h) at the point $(\bar{y}_h, \bar{x}_h) = (\bar{y}_h^*, \bar{x}_h^*)$.

Taking the expectation of (2.32) the bias of the estimator t_S up to the terms of the order n^{-1} is obtained

$$B(t_S) = \frac{1}{2} \sum_{h=1}^L W_h \left\{ \begin{aligned} & \text{Var}(\bar{x}_h) t_{h22}(\mu_{Yh}, \mu_{Xh}) + \\ & 2\text{Cov}(\bar{y}_h, \bar{x}_h) t_{h12}(\mu_{Yh}, \mu_{Xh}) \end{aligned} \right\},$$

or

$$B(t_S) = \frac{1}{2} \sum_{h=1}^L \frac{W_h^2}{n_h} \left\{ \begin{aligned} & \frac{\sigma_{Xh}^2}{\theta_{Xh}} t_{h22}(\mu_{Yh}, \mu_{Xh}) \\ & + 2\sigma_{YXh} t_{h12}(\mu_{Yh}, \mu_{Xh}) \end{aligned} \right\} \\ = \frac{1}{2} \sum_{h=1}^L \frac{W_h^2 \sigma_{Yh}^2}{n_h \theta_{Xh}} \left\{ \begin{aligned} & t_{h22}(\mu_{Yh}, \mu_{Xh}) \\ & + 2\beta_{YXh} t_{h12}(\mu_{Yh}, \mu_{Xh}) \end{aligned} \right\} \quad (2.33)$$

Up to the terms of order n^{-1} , the MSE of t_S is :

$$\text{MSE}(t_S) = \frac{1}{2} \sum_{h=1}^L W_h^2 \left\{ \begin{aligned} & \text{Var}(\bar{y}_h) \\ & + \text{Var}(\bar{x}_h) t_{h2}^2(\mu_{Yh}, \mu_{Xh}) \\ & + 2\text{Cov}(\bar{y}_h, \bar{x}_h) t_{h12}(\mu_{Yh}, \mu_{Xh}) \end{aligned} \right\} \\ = \sum_{h=1}^L \frac{W_h^2}{n_h} \left\{ \begin{aligned} & \frac{\sigma_{Yh}^2}{\theta_{Yh}} + \frac{\sigma_{Xh}^2}{\theta_{Xh}} t_{h2}^2(\mu_{Yh}, \mu_{Xh}) \\ & + 2\sigma_{YXh} t_{h2}(\mu_{Yh}, \mu_{Xh}) \end{aligned} \right\} \\ = \sum_{h=1}^L \frac{W_h^2}{n_h} \left[\begin{aligned} & \frac{\sigma_{Yh}^2}{\theta_{Yh}} + \frac{\sigma_{Xh}^2}{\theta_{Xh}} t_{h2}(\mu_{Yh}, \mu_{Xh}) \\ & \left\{ t_{h2}(\mu_{Yh}, \mu_{Xh}) + 2\beta_{YXh} \theta_{Xh} \right\} \end{aligned} \right] \quad (2.34)$$

The $\text{MSE}(t_S)$ is minimized for

$$t_{h12}(\mu_{Yh}, \mu_{Xh}) = - \left(\frac{\sigma_{YXh}}{\sigma_{Xh}^2} \right) \theta_{Xh} \quad (2.35) \\ = -\beta_{YXh} \theta_{Xh}$$

Thus, the resulting minimum MSE of t_S is given by

$$\text{min. MSE}(t_S) = \sum_{h=1}^L \frac{W_h^2 \sigma_{Yh}^2}{n_h \theta_{Yh}} (1 - \rho_h^2 \theta_{Xh} \theta_{Yh}) \quad (2.36)$$

Theorem 2.1

Based on the previous discussion, the following theorem is put forth. To the first degree of approximation,

$$MSE(t_S) \geq \sum_{h=1}^L \frac{W_h^2 \sigma_{Yh}^2}{n_h \theta_{Yh}} (1 - \rho_h^2 \theta_{Xh} \theta_{Yh}),$$

with equality holding if

$$t_{h12}(\mu_{Yh}, \mu_{Xh}) = -\beta_{YXh} \theta_{Xh}.$$

Note the lower bound of the MSEs of the separate class of estimators t_S at (2.30) in the variance of the optimum separate difference estimator (OSDE)

$$t_{dOS} = \sum_{h=1}^L W_h \{ \bar{y}_h + d_{0h} (\mu_{Yh} - \bar{x}_h) \}$$

with

$$d_{0h} = \beta_{YXh} \theta_{Xh},$$

which shows that the estimators belonging to the class of separate estimators t_S at (2.30) are asymptotically no more efficient than the optimum difference estimator (ODE) t_{dOS} .

Any parametric function $t_h(\bar{y}_h, \bar{x}_h)$ satisfying conditions (i) and (ii) can define an μ_{Yh} . The class of such estimators is very large. For example, the following estimators:

$$t_{S1} = \sum_{h=1}^L W_h \bar{y}_h \left(\frac{\mu_{Yh}}{\bar{x}_h} \right)^{\alpha_h},$$

$$t_{S2} = \sum_{h=1}^L W_h \bar{y}_h \left\{ 2 - \left(\frac{\bar{x}_h}{\mu_{Yh}} \right)^{\alpha_h} \right\},$$

$$t_{S3} = \sum_{h=1}^L W_h \bar{y}_h \frac{\mu_{Xh}}{\{ \mu_{Xh} + \alpha_h (\bar{x}_h - \mu_{Xh}) \}},$$

$$t_{S4} = \sum_{h=1}^L W_h \{ \bar{y}_h + \alpha_h (\bar{x}_h - \mu_{Xh}) \},$$

are particular members of a proposed class of estimator where α_h is a suitably chosen

constant. The optimum value of constant α_h in $t_{Sj}, j = 1$ to 4 , which minimizes the mean squared error of the resulting estimator are obtained from (2.36).

It follows from (2.7), (2.8), (2.16) and (2.34) that the proposed separate class of estimators t_S is more efficient than:

i. the usual unbiased estimator \bar{y}_{st} if

$$\begin{aligned} & \min. \{ 0, -\beta_{YXh} \theta_{Xh} \} \\ & < t_{h2}(\mu_{Yh}, \mu_{Xh}) < \\ & \max. \{ 0, -\beta_{YXh} \theta_{Xh} \} \end{aligned} \quad (2.37)$$

ii. the separate ratio estimator \bar{y}_{RS} if

$$\begin{aligned} & \min. \{ -R_h, (R_h - 2\beta_{YXh} \theta_{Xh}) \} \\ & < t_{h2}(\mu_{Yh}, \mu_{Xh}) < \\ & \max. \{ -R_h, (R_h - 2\beta_{YXh} \theta_{Xh}) \} \end{aligned} \quad (2.38)$$

iii. the separate product estimator \bar{y}_{PS} if

$$\begin{aligned} & \min. \{ R_h, -(R_h + 2\beta_{YXh} \theta_{Xh}) \} \\ & < t_{h2}(\mu_{Yh}, \mu_{Xh}) < \\ & \max. \{ R_h, -(R_h + 2\beta_{YXh} \theta_{Xh}) \} \end{aligned} \quad (2.39)$$

A Combined Ratio Estimator in Stratified Random Sampling in the Presence of Measurement Errors

For the estimation of population mean, μ_Y , the following combined ratio estimator is defined in the presence of measurement errors:

$$t_{RC} = \bar{y}_{st} \frac{\mu_X}{\bar{x}_{st}} \quad (2.40)$$

To the first degree of approximation, the bias and mean squared error of the combined ratio estimator t_{RC} are respectively given by

$$\begin{aligned}
 B(t_{RC}) &= \mu_Y \sum_{h=1}^L \frac{W_h^2}{n_h} \left[\frac{(\sigma_{Xh}^2 + \sigma_{Vh}^2)}{\mu_X^2} - \frac{\rho_h \sigma_{Xh} \sigma_{Yh}}{\mu_X \mu_Y} \right] \\
 &= \mu_Y \sum_{h=1}^L \frac{W_h^2}{n_h} \left[\frac{\sigma_{Xh}^2}{\mu_X^2 \theta_{Xh}} - \frac{\beta_{YXh} \sigma_{Xh}^2}{\mu_X \mu_Y} \right] \\
 &= \mu_Y \sum_{h=1}^L \frac{W_h^2 \sigma_{Xh}^2}{n_h} \left[\frac{R}{\mu_X^2 \theta_{Xh}} - \frac{\beta_{YXh}}{\mu_X} \right] \\
 &= \left(\frac{1}{\mu_X} \right) \sum_{h=1}^L \frac{W_h^2 \sigma_{Xh}^2}{n_h \theta_{Xh}} (R - \beta_{YXh} \theta_{Xh})
 \end{aligned} \tag{2.41}$$

and

$$\begin{aligned}
 MSE(t_{RC}) &= \\
 &\sum_{h=1}^L \frac{W_h^2}{n_h} \left[\frac{\sigma_{Yh}^2}{\theta_{Yh}} + R \left(\frac{\sigma_{Xh}^2}{\theta_{Xh}} \right) (R - 2\beta_{YXh} \theta_{Xh}) \right]
 \end{aligned} \tag{2.42}$$

From (2.8) and (2.42):

$$\begin{aligned}
 \text{Var}(\bar{y}_{st}) - MSE(t_{RC}) &= \\
 &-\sum_{h=1}^L \frac{W_h^2}{n_h} \left(\frac{\sigma_{Xh}^2}{\theta_{Xh}} \right) R (R - 2\beta_{YXh} \theta_{Xh}),
 \end{aligned}$$

which is positive if

$$\frac{\sum_{h=1}^L \frac{W_h^2}{n_h} \beta_{YXh} \left(\frac{\sigma_{Xh}^2}{\theta_{Xh}} \right)}{R \sum_{h=1}^L \frac{W_h^2}{n_h} \sigma_{Xh}^2} > \frac{1}{2} \tag{2.43}$$

When the data are recorded without error then the expression (2.42) reduces to:

$$MSE(t_{RC})_t = \sum_{h=1}^L \frac{W_h^2}{n_h} [\sigma_{Yh}^2 + R \sigma_{Xh}^2 (R - 2\beta_{YXh})] \tag{2.44}$$

From (2.42) and (2.44):

$$\begin{aligned}
 MSE(\bar{y}_{RC}) - MSE(t_{RC})_t &= \\
 &\sum_{h=1}^L \frac{W_h^2}{n_h} \left[\left(\frac{1 - \theta_{Yh}}{\theta_{Yh}} \right) \sigma_{Yh}^2 + R^2 \left(\frac{1 - \theta_{Xh}}{\theta_{Xh}} \right) \sigma_{Xh}^2 \right]
 \end{aligned} \tag{2.45}$$

which is always positive. It follows that the presence of measurement errors in both the variables X and Y inflates the $MSE(t_{RC})$.

A Combined Product Estimator in Stratified Random Sampling in the Presence of Measurement Errors

The following combined product estimator is defined for the population mean μ_Y in the presence of measurement errors as:

$$t_{PC} = \bar{y}_{st} \frac{\bar{X}_{st}}{\mu_X} \tag{2.46}$$

The exact bias and mean squared error to the first degree of approximation of the combined ratio estimator t_{PC} are respectively given by

$$B(t_{PC}) = \mu_Y \sum_{h=1}^L \left(\frac{W_h^2}{n_h} \right) \frac{\beta_{YXh} \sigma_{Xh}^2}{\mu_X} \tag{2.47}$$

and

$$MSE(t_{PC}) = \sum_{h=1}^L \frac{W_h^2}{n_h} \left[\frac{\sigma_{Yh}^2}{\theta_{Yh}} + R \left(\frac{\sigma_{Xh}^2}{\theta_{Xh}} \right) (R + 2\beta_{YXh} \theta_{Xh}) \right] \tag{2.48}$$

From (2.8) and (2.48):

$$\begin{aligned}
 \text{Var}(\bar{y}_{st}) - MSE(t_{PC}) &= \\
 &-\sum_{h=1}^L \frac{W_h^2}{n_h} \left(\frac{\sigma_{Xh}^2}{\theta_{Xh}} \right) R (R + 2\beta_{YXh} \theta_{Xh})
 \end{aligned} \tag{2.49}$$

which is positive if

$$\frac{\sum_{h=1}^L \frac{W_h^2}{n_h} \beta_{YXh} \left(\frac{\sigma_{Xh}^2}{\theta_{Xh}} \right)}{R \sum_{h=1}^L \frac{W_h^2}{n_h} \sigma_{Xh}^2} < -\frac{1}{2} \quad (2.50)$$

Assuming that the observations for X and Y be recorded without error, then expression (2.48) reduces to:

$$MSE(t_{PC})_t = \sum_{h=1}^L \frac{W_h^2}{n_h} \left[\sigma_{Yh}^2 + R \sigma_{Xh}^2 (R + 2\beta_{YXh}) \right] \quad (2.51)$$

which can be obtained from (2.48) by setting $\theta_{Yh} = \theta_{Xh} = 1$. From (2.48) and (2.51):

$$MSE(t_{PC}) - MSE(t_{PC})_t = \sum_{h=1}^L \frac{W_h^2}{n_h} \left[\left(\frac{1 - \theta_{Yh}}{\theta_{Yh}} \right) \sigma_{Yh}^2 + R^2 \left(\frac{1 - \theta_{Xh}}{\theta_{Xh}} \right) \sigma_{Xh}^2 \right] \quad (2.52)$$

which is always positive. Expression (2.52) is the same as that obtained in (2.45). Thus, the presence of measurement errors in both variables X and Y are responsible for increasing the MSE of the combined product estimator t_{PC} .

A Combined Difference Estimator in Stratified Random Sampling in the presence of Measurement Errors

A combined difference estimator in stratified random sampling is defined in the presence of measurement errors for a population mean, μ_Y , as

$$t_{dC} = \bar{y}_{st} + d(\mu_X - \bar{x}_{st}) \quad (2.53)$$

where d is a suitably chosen constant. It can be seen that the combined difference estimator t_{dC} is unbiased. The variance of t_{dC} is given by

$$\begin{aligned} \text{Var}(t_{dC}) &= \\ &= \text{Var}(\bar{y}_{st}) + d^2 (\bar{x}_{st}) - 2d \text{Cov}(\bar{y}_{st}, \bar{x}_{st}) \\ &= \sum_{h=1}^L \frac{W_h^2}{n_h} \left[\frac{\sigma_{Yh}^2}{\theta_{Yh}} + d^2 \left(\frac{\sigma_{Xh}^2}{\theta_{Xh}} \right) - 2d \sigma_{YXh} \right] \quad (2.54) \\ &= \sum_{h=1}^L \frac{W_h^2}{n_h} \left(\frac{\sigma_{Yh}^2}{\theta_{Yh}} \right) + d^2 \sum_{h=1}^L \frac{W_h^2}{n_h} \left(\frac{\sigma_{Xh}^2}{\theta_{Xh}} \right) \\ &\quad - 2d \sum_{h=1}^L \frac{W_h^2}{n_h} \beta_{YXh} \sigma_{Xh}^2 \end{aligned}$$

which is minimized for

$$d = \frac{\sum_{h=1}^L \frac{W_h^2}{n_h} \beta_{YXh} \sigma_{Xh}^2}{\sum_{h=1}^L \frac{W_h^2}{n_h} \left(\frac{\sigma_{Xh}^2}{\theta_{Xh}} \right)} \quad (2.55)$$

Thus, the resulting minimum variance of t_{dC} is given by

$$\min \text{Var}(t_{dC}) = \sum_{h=1}^L \frac{W_h^2}{n_h} \left(\frac{\sigma_{Yh}^2}{\theta_{Yh}} \right) - \frac{\left\{ \sum_{h=1}^L \frac{W_h^2}{n_h} \beta_{YXh} \sigma_{Xh}^2 \right\}^2}{\sum_{h=1}^L \frac{W_h^2}{n_h} \left(\frac{\sigma_{Xh}^2}{\theta_{Xh}} \right)} \quad (2.56)$$

Assume the data associated with variables X and Y are recorded without error; in such a case, the expressions (2.54) reduce to:

$$\min \text{Var}(t_{dC}) = \sum_{h=1}^L \frac{W_h^2}{n_h} \left(\frac{\sigma_{Yh}^2}{\theta_{Yh}} \right) - \frac{\left\{ \sum_{h=1}^L \frac{W_h^2}{n_h} \beta_{YXh} \sigma_{Xh}^2 \right\}^2}{\sum_{h=1}^L \frac{W_h^2}{n_h} \left(\frac{\sigma_{Xh}^2}{\theta_{Xh}} \right)} \quad (2.57)$$

From (2.54) and (2.57):

$$\begin{aligned} \text{Var}(t_{dC}) - \text{Var}(t_{dC})_t = & \\ \sum_{h=1}^L \frac{W_h^2}{n_h} \sigma_{Yh}^2 \left(\frac{1 - \theta_{Yh}}{\theta_{Yh}} \right) + d^2 \sum_{h=1}^L \frac{W_h^2}{n_h} \sigma_{Xh}^2 \left(\frac{1 - \theta_{Xh}}{\theta_{Xh}} \right) & \end{aligned} \quad (2.58)$$

which is always positive. It follows from (2.56) that the presence of measurement errors in both variables X and Y enhances the variance of t_{dC} regardless of the value of d .

The $\text{Var}(t_{dC})$ at (2.57) is minimized for

$$d = \frac{\sum_{h=1}^L \frac{W_h^2}{n_h} \beta_{YXh} \sigma_{Xh}^2}{\sum_{h=1}^L \frac{W_h^2}{n_h} \sigma_{Xh}^2} \quad (2.59)$$

Combining (2.59) with (2.57) results in the minimum value of $\text{Var}(\bar{y}_{st})$ as

$$\begin{aligned} \min \text{Var}(t_{dC})_t = \sum_{h=1}^L \frac{W_h^2}{n_h} \sigma_{Yh}^2 - \frac{\left(\sum_{h=1}^L \frac{W_h^2}{n_h} \beta_{YXh} \sigma_{Xh}^2 \right)^2}{\sum_{h=1}^L \frac{W_h^2}{n_h} \sigma_{Xh}^2} & \end{aligned} \quad (2.60)$$

From (2.56) and (2.58):

$$\begin{aligned} \min \text{Var}(t_{dC}) - \min \text{Var}(t_{dC})_t = & \\ \sum_{h=1}^L \frac{W_h^2}{n_h} \sigma_{Yh}^2 \left(\frac{1 - \theta_{Yh}}{\theta_{Yh}} \right) & \\ + \left\{ \sum_{h=1}^L \frac{W_h^2}{n_h} \beta_{YXh} \sigma_{Xh}^2 \right\}^2 \frac{\sum_{h=1}^L \frac{W_h^2}{n_h} \sigma_{Xh}^2 \left(\frac{1 - \theta_{Xh}}{\theta_{Xh}} \right)}{\left\{ \sum_{h=1}^L \frac{W_h^2}{n_h} \sigma_{Xh}^2 \right\} \left\{ \sum_{h=1}^L \frac{W_h^2}{n_h} \left(\frac{\sigma_{Xh}^2}{\theta_{Xh}} \right) \right\}} & \end{aligned} \quad (2.61)$$

It can be observed from (2.61) that the difference $[\min \text{Var}(t_{dC}) - \min \text{Var}(t_{dC})_t]$ is always positive. Thus the presence of measurement error in both variables X and Y

inflates the variance of t_{dC} at their optimum conditions, which disturbs the optimal properties of the difference estimator t_{dC} .

A Combined Class of Estimators in Stratified Random Sampling in the presence of Measurement Errors

Following the same procedure as adopted by Srivastava (1971, 1980) a class of combined estimators of the population mean μ_Y in the presence of measurement errors is suggested, such as

$$t_C = t(\bar{y}_{st}, \bar{x}_{st}), \quad (2.62)$$

where $t(\bar{y}_{st}, \bar{x}_{st})$ is a function of $(\bar{y}_{st}, \bar{x}_{st})$ such that

$$t(\mu_Y, \mu_X) = \mu_Y \quad (2.63)$$

$$\Rightarrow t_1(\mu_Y, \mu_X) = \left. \frac{\partial t_1(\mu_Y, \mu_X)}{\partial \bar{y}_{st}} \right|_{(\mu_Y, \mu_X)} = 1,$$

and satisfies the same conditions as given in 2.4 for t_S .

Expanding the function $t(\bar{y}_{st}, \bar{x}_{st})$ about the point $(\bar{y}_{st}, \bar{x}_{st}) = (\mu_Y, \mu_X)$ in a third-order Taylor's series results in

$$\begin{aligned} t_C = & \\ \mu_Y + & \left[\begin{aligned} & (\bar{y}_{st} - \mu_Y) t_1(\mu_Y, \mu_X) \\ & + (\bar{x}_{st} - \mu_X) t_2(\mu_Y, \mu_X) \\ & + \frac{1}{2} \left\{ \begin{aligned} & (\bar{y}_{st} - \mu_Y)^2 t_{11}(\mu_Y, \mu_X) \\ & + 2(\bar{y}_{st} - \mu_Y)(\bar{x}_{st} - \mu_X) t_{12}(\mu_Y, \mu_X) \\ & + (\bar{x}_{st} - \mu_X)^2 t_{22}(\mu_Y, \mu_X) \end{aligned} \right\} \\ & + \frac{1}{6} \left\{ \begin{aligned} & (\bar{y}_{st} - \mu_Y)^2 t_{111}(\bar{y}_{st}^*, \bar{x}_{st}^*) \\ & + 3(\bar{y}_{st} - \mu_Y)(\bar{x}_{st} - \mu_X)^2 t_{122}(\bar{y}_{st}^*, \bar{x}_{st}^*) \\ & + 3(\bar{y}_{st} - \mu_Y)^2 (\bar{x}_{st} - \mu_X) t_{112}(\bar{y}_{st}^*, \bar{x}_{st}^*) \\ & + (\bar{x}_{st} - \mu_X)^3 t_{222}(\bar{y}_{st}^*, \bar{x}_{st}^*) \end{aligned} \right\} \end{aligned} \right] \end{aligned}$$

or

$$\begin{aligned}
 (t_C - \mu_Y) = & \left[\begin{aligned} & (\bar{y}_{st} - \mu_Y) t_1(\mu_Y, \mu_X) \\ & + (\bar{x}_{st} - \mu_X) t_2(\mu_Y, \mu_X) \\ & + \frac{1}{2} \left\{ \begin{aligned} & (\bar{y}_{st} - \mu_Y)^2 t_{11}(\mu_Y, \mu_X) \\ & + 2(\bar{y}_{st} - \mu_Y)(\bar{x}_{st} - \mu_X) t_{12}(\mu_Y, \mu_X) \\ & + (\bar{x}_{st} - \mu_X)^2 t_{22}(\mu_Y, \mu_X) \end{aligned} \right\} \\ & + \frac{1}{6} \left\{ \begin{aligned} & (\bar{y}_{st} - \mu_Y)^3 t_{111}(\bar{y}_{st}^*, \bar{x}_{st}^*) \\ & + 3(\bar{y}_{st} - \mu_Y)(\bar{x}_{st} - \mu_X)^2 t_{122}(\bar{y}_{st}^*, \bar{x}_{st}^*) \\ & + 3(\bar{y}_{st} - \mu_Y)^2 (\bar{x}_{st} - \mu_X) t_{112}(\bar{y}_{st}^*, \bar{x}_{st}^*) \\ & + (\bar{x}_{st} - \mu_X)^3 t_{222}(\bar{y}_{st}^*, \bar{x}_{st}^*) \end{aligned} \right\} \end{aligned} \right] \tag{2.64}
 \end{aligned}$$

where $\bar{y}_{st}^* = \{\mu_Y + \xi(\bar{y}_{st} - \mu_Y)\}$, $\bar{x}_{st}^* = \{\mu_X + \xi(\bar{x}_{st} - \mu_X)\}$, and $0 < \xi < 1$. Also, ξ may depend on $(\bar{y}_{st}^*, \bar{x}_{st}^*)$ and $t_{ijk}(\bar{y}_{st}^*, \bar{x}_{st}^*)$ denotes the third order partial derivative of $t(\bar{y}_{st}, \bar{x}_{st})$ with respect to $(\bar{y}_{st}, \bar{x}_{st})$ at the point $(\bar{y}_{st}, \bar{x}_{st}) = (\bar{y}_{st}^*, \bar{x}_{st}^*)$.

Taking the expectation of (2.64) provides the bias of the estimator t_C up to the terms of the order n^{-1} ,

$$\begin{aligned}
 B(t_C) = & \frac{1}{2} \left\{ \begin{aligned} & \text{Var}(\bar{x}_{st}) t_{22}(\mu_Y, \mu_X) \\ & + 2\text{Cov}(\bar{y}_{st}, \bar{x}_{st}) t_{12}(\mu_Y, \mu_X) \end{aligned} \right\}, \\
 B(t_C) = & \frac{1}{2} \sum_{h=1}^L \frac{W_h^2}{n_h} \left\{ \begin{aligned} & \frac{\sigma_{Xh}^2}{\theta_{Xh}} t_{22}(\mu_Y, \mu_X) \\ & + 2\sigma_{YXh} t_{12}(\mu_Y, \mu_X) \end{aligned} \right\} \\
 = & \frac{1}{2} \left[\begin{aligned} & \sum_{h=1}^L \frac{W_h^2 \sigma_{Yh}^2}{n_h \theta_{Xh}} t_{22}(\mu_Y, \mu_X) \\ & + 2 \sum_{h=1}^L \frac{W_h^2}{n_h} \beta_{YXh} \sigma_{Xh}^2 t_{12}(\mu_Y, \mu_X) \end{aligned} \right] \tag{2.65}
 \end{aligned}$$

Squaring both sides of (2.64) and neglecting terms $(\bar{y}_{st} - \mu_Y)$ and $(\bar{x}_{st} - \mu_X)$ having power greater than two results in

$$(t_C - \mu_Y)^2 = \left[\begin{aligned} & (\bar{y}_{st} - \mu_Y)^2 t_1^2(\mu_Y, \mu_X) \\ & + (\bar{x}_{st} - \mu_X) t_2^2(\mu_Y, \mu_X) \\ & + 2\text{Cov}(\bar{y}_{st}, \bar{x}_{st}) t_{12}(\mu_Y, \mu_X) \end{aligned} \right]$$

Noting that $t_1(\mu_Y, \mu_X) = 1$ and taking the expectation of both sides of the above expression, provides the mean squared error of the class of combined estimators t_C as

$$\text{MSE}(t_C) = \left[\begin{aligned} & \text{Var}(\bar{y}_{st}) + \text{Var}(\bar{x}_{st}) t_{22}(\mu_Y, \mu_X) \\ & + 2\text{Cov}(\bar{y}_{st}, \bar{x}_{st}) t_{12}(\mu_Y, \mu_X) \end{aligned} \right],$$

or

$$\begin{aligned}
 \text{MSE}(t_C) = & \sum_{h=1}^L \frac{W_h^2}{n_h} \left\{ \begin{aligned} & \frac{\sigma_{Yh}^2}{\theta_{Yh}} + \frac{\sigma_{Xh}^2}{\theta_{Xh}} t_2^2(\mu_Y, \mu_X) \\ & + 2\sigma_{YXh} t_2(\mu_Y, \mu_X) \end{aligned} \right\} \\
 = & \left[\begin{aligned} & \sum_{h=1}^L \frac{W_h^2}{n_h} \frac{\sigma_{Yh}^2}{\theta_{Yh}} + \sum_{h=1}^L \frac{W_h^2}{n_h} \frac{\sigma_{Xh}^2}{\theta_{Xh}} t_2^2(\mu_Y, \mu_X) \\ & + 2 \sum_{h=1}^L \frac{W_h^2}{n_h} \beta_{YXh} \sigma_{Xh}^2 t_2(\mu_Y, \mu_X) \end{aligned} \right] \tag{2.66}
 \end{aligned}$$

which is minimized for

$$t_2(\mu_Y, \mu_X) = - \frac{\sum_{h=1}^L \frac{W_h^2}{n_h} \beta_{YXh} \sigma_{Xh}^2}{\sum_{h=1}^L \frac{W_h^2}{n_h} \left(\frac{\sigma_{Xh}^2}{\theta_{Xh}} \right)} \tag{2.67}$$

Thus the resulting minimum MSE of t_C is given by

$$\min .MSE(t_c) = \sum_{h=1}^L \frac{W_h^2}{n_h} \left(\frac{\sigma_{Yh}^2}{\theta_{Yh}} \right) - \frac{\left\{ \sum_{h=1}^L \frac{W_h^2}{n_h} \beta_{YXh} \sigma_{Xh}^2 \right\}^2}{\sum_{h=1}^L \frac{W_h^2}{n_h} \left(\frac{\sigma_{Xh}^2}{\theta_{Xh}} \right)} \quad (2.68)$$

Theorem 2.2

Based on the above, the following theorem is put forth. To the first degree of approximation,

$$MSE(t_c) \geq \left[\sum_{h=1}^L \frac{W_h^2}{n_h} \left(\frac{\sigma_{Yh}^2}{\theta_{Yh}} \right) - \frac{\left\{ \sum_{h=1}^L \frac{W_h^2}{n_h} \beta_{YXh} \sigma_{Xh}^2 \right\}^2}{\sum_{h=1}^L \frac{W_h^2}{n_h} \left(\frac{\sigma_{Xh}^2}{\theta_{Xh}} \right)} \right]$$

with equality holding if

$$t_2(\mu_Y, \mu_X) = - \frac{\sum_{h=1}^L \frac{W_h^2}{n_h} \beta_{YXh} \sigma_{Xh}^2}{\sum_{h=1}^L \frac{W_h^2}{n_h} \left(\frac{\sigma_{Xh}^2}{\theta_{Xh}} \right)}$$

Note that the lower bound of the MSE of the combined class of estimators t_c at (2.62) is the variance of the optimum combined difference estimator (OCDE)

$$t_{d0C} = \bar{y}_{st} + d_0(\mu_X - \bar{x}_{st}) \quad (2.69)$$

with

$$d_0 = \frac{\sum_{h=1}^L \frac{W_h^2}{n_h} \beta_{YXh} \sigma_{Xh}^2}{\sum_{h=1}^L \frac{W_h^2}{n_h} \left(\frac{\sigma_{Xh}^2}{\theta_{Xh}} \right)},$$

which demonstrates that the estimators belonging to the class of combined estimators t_c at (2.62) are asymptotically no more efficient than the optimum difference estimator (ODE) t_{d0C} at (2.69).

Any parametric function $t(\bar{y}_{st}, \bar{x}_{st})$ satisfying the regularity conditions as described for t_c , can define a μ_Y . The class of such estimators is very large. For example, the following estimators:

$$t_{C1} = \bar{y}_{st} \left(\frac{\mu_Y}{\bar{x}_{st}} \right)^\psi,$$

$$t_{C2} = \bar{y}_{st} \left\{ 2 - \left(\frac{\bar{x}_{st}}{\mu_Y} \right)^\psi \right\},$$

$$t_{C3} = \bar{y}_{st} \frac{\mu_X}{\left\{ \mu_X + \psi(\bar{x}_{st} - \mu_X) \right\}},$$

$$t_{C4} = \left\{ \bar{y}_{st} + \psi(\bar{x}_{st} - \mu_X) \right\},$$

are particular members of the proposed class of estimator, where ψ is a suitably chosen constant. The optimum value of constant ψ in $t_{Cj}, j = 1$ to 4 which minimizes the mean squared error of the resulting estimator are obtained from (2.68).

It follows from (2.42), (2.8), (2.48) and (2.66) that the proposed separate class of estimators t_c is more efficient than:

- i. the usual unbiased estimator \bar{y}_{st} if

$$\min. \left\{ 0, \frac{-2 \sum_{h=1}^L p_h \beta_{YXh} \theta_{Xh}}{\sum_{h=1}^L p_h} \right\} < t_2(\mu_Y, \mu_X) < \quad (2.70)$$

$$\max. \left\{ 0, \frac{-2 \sum_{h=1}^L p_h \beta_{YXh} \theta_{Xh}}{\sum_{h=1}^L p_h} \right\}$$

- ii. the combined ratio estimator \bar{y}_{CS} if

$$\min. \left\{ -R, R - 2 \left(\frac{\sum_{h=1}^L p_h \beta_{YXh} \theta_{Xh}}{\sum_{h=1}^L p_h} \right) \right\} < t_2(\mu_{Yh}, \mu_{Xh}) < \quad (2.71)$$

$$\max. \left\{ -R, R - 2 \left(\frac{\sum_{h=1}^L p_h \beta_{YXh} \theta_{Xh}}{\sum_{h=1}^L p_h} \right) \right\}$$

iii. the combined product estimator \bar{y}_{PC} if

$$\min. \left\{ -R, -R - 2 \left(\frac{\sum_{h=1}^L p_h \beta_{YXh} \theta_{Xh}}{\sum_{h=1}^L p_h} \right) \right\} < t_2(\mu_{Yh}, \mu_{Xh}) < \quad (2.72)$$

$$\max. \left\{ -R, -R - 2 \left(\frac{\sum_{h=1}^L p_h \beta_{YXh} \theta_{Xh}}{\sum_{h=1}^L p_h} \right) \right\}$$

where

$$p_h = \sum_{h=1}^L \frac{W_h^2}{n_h} \left(\frac{\sigma_{Xh}^2}{\theta_{Xh}} \right).$$

Theoretical Comparisons

From (2.7) (or (2.36)) and (2.42) (or (2.68)):

$$\text{MSE}(t_{RS} \text{ or } t_S) - \text{MSE}(t_{RC}) = \sum_{h=1}^L \frac{W_h^2}{n_h} \left(\frac{\sigma_{Xh}^2}{\theta_{Xh}} \right) \left[(R - \beta_{YXh} \theta_{Xh})^2 - (R_h - \beta_{YXh} \theta_{Xh})^2 \right]$$

which is positive if

$$(R - \beta_{YXh} \theta_{Xh})^2 > (R_h - \beta_{YXh} \theta_{Xh})^2 \quad (3.1)$$

It follows that t_{RS} will be more efficient than t_{RC} if and only if $(\beta_{YXh} \theta_{Xh})$ is nearer to R_h than to R .

From (2.16) and (2.48):

$$\text{MSE}(t_{PS}) - \text{MSE}(t_{PC}) = \sum_{h=1}^L \frac{W_h^2}{n_h} \left(\frac{\sigma_{Xh}^2}{\theta_{Xh}} \right) \left[(R + \beta_{YXh} \theta_{Xh})^2 - (R_h + \beta_{YXh} \theta_{Xh})^2 \right]$$

which is positive if

$$(R + \beta_{YXh} \theta_{Xh})^2 > (R_h + \beta_{YXh} \theta_{Xh})^2 \quad (3.2)$$

Thus, the separate product estimator t_{PC} is more efficient than the combined product estimator t_{PC} if the inequality (3.2) holds.

From (2.24) and (2.56):

$$\begin{aligned} \min \text{MSE}(t_{dC}) - \min \text{MSE}(t_{dS}) &= \frac{\left\{ \sum_{h=1}^L \frac{W_h^2}{n_h} \beta_{YXh} \sigma_{Xh}^2 \right\}^2}{\sum_{h=1}^L \frac{W_h^2}{n_h} \left(\frac{\sigma_{Xh}^2}{\theta_{Xh}} \right)} - \left\{ \sum_{h=1}^L \frac{W_h^2}{n_h} \beta_{YXh}^2 \theta_{Xh} \sigma_{Xh}^2 \right\} \\ &= \frac{1}{\sum_{h=1}^L \frac{W_h^2}{n_h} \left(\frac{\sigma_{Xh}^2}{\theta_{Xh}} \right)} \left[\left\{ \sum_{h=1}^L \frac{W_h^2}{n_h} \beta_{YXh} \sigma_{Xh}^2 \right\}^2 - \left\{ \sum_{h=1}^L \frac{W_h^2}{n_h} \beta_{YXh}^2 \theta_{Xh} \sigma_{Xh}^2 \right\} \sum_{h=1}^L \frac{W_h^2}{n_h} \left(\frac{\sigma_{Xh}^2}{\theta_{Xh}} \right) \right] \\ &= \frac{\left\{ \sum_{h=1}^L p_h \beta_{YXh} \theta_{Xh} \right\}^2}{\sum_{h=1}^L p_h} - \sum_{h=1}^L p_h \beta_{YXh}^2 \theta_{Xh}^2 \\ &= \left\{ \frac{\sum_{h=1}^L p_h \beta_{YXh} \theta_{Xh}}{\sum_{h=1}^L p_h} \right\}^2 \sum_{h=1}^L p_h - \sum_{h=1}^L p_h \beta_{YXh}^2 \theta_{Xh}^2 \end{aligned}$$

EFFECT OF MEASUREMENT ERRORS ON PRODUCT ESTIMATORS

$$\begin{aligned}
 &= \sum_{h=1}^L p_h \left[\left\{ \frac{\sum_{h=1}^L p_h \beta_{YXh} \theta_{Xh}}{\sum_{h=1}^L p_h} \right\}^2 - \beta_{YXh}^2 \theta_{Xh}^2 \right] \\
 &= \sum_{h=1}^L p_h (\beta_{YXh} \theta_{Xh} - a)^2
 \end{aligned} \tag{3.3}$$

where

$$a = \frac{\sum_{h=1}^L p_h \beta_{YXh} \theta_{Xh}}{\sum_{h=1}^L p_h}$$

and

$$p_h = \frac{W_h^2 \sigma_{Xh}^2}{n_h \theta_{Xh}}.$$

Observe the expression (3.3) is always positive. Thus, unless the term $(\beta_{YXh} \theta_{Xh})$ is the same from stratum to stratum, the separate difference estimator t_{ds} (or the separate class of the estimators t_s) at its optimum condition, that is, OSDE $t_{dos} = \sum_{h=1}^L W_h \{ \bar{y}_h + d_{0h} (\mu_{Yh} - \bar{x}_h) \}$ with $d_{0h} = \beta_{YXh} \theta_{Xh}$ is more efficient than the combined difference estimator t_{dC} (or the combined class of the estimators t_C) at optimum (i.e., the OCDE $t_{d0C} = \bar{y}_{st} + d_0 (\mu_X - \bar{x}_{st})$ with

$$d_0 = \frac{\sum_{h=1}^L \frac{W_h^2}{n_h} \beta_{YXh} \sigma_{Xh}^2}{\sum_{h=1}^L \frac{W_h^2}{n_h} \left(\frac{\sigma_{Xh}^2}{\theta_{Xh}} \right)}.$$

References

Chandhok, P. K., & Han, C. P. (1990). On the efficiency of the ratio estimator under Mizuno scheme with measurement errors. *Journal of the Indian Statistical Association*, 28, 31-39.

Cochran, W. G. (1968). Errors of measurement in Statistics. *Technometrics*, 10, 637-666.

Manisha, & Singh, R. K. (2001). An estimation of population mean in the presence of measurement errors. *Journal of the Indian Society of Agricultural Statistics*, 54(1), 13-18.

Maneesha, & Singh, R. K. (2002). Role of regression estimator involving measurement errors. *Brazilian Journal of Probability and Statistics*, 16, 39-46.

Shalabh (1997). Ratio method of estimation in the presence of measurement errors, *Journal of the Indian Society of Agricultural Statistics*, 50(2), 150-155.

Singh, H. P., & Karpe, N. (2008a). Ratio-Product estimator for population mean in presence of measurement errors. *Journal of Applied Statistical Sciences*, 16(4), 49-64.

Singh, H. P., & Karpe, N. (2008b). Estimation of population variance using auxiliary information in the presence of Measurement Errors. *Statistics in Transition-New Series*, 9(3), 443-470.

Singh, H. P., & Karpe, N. (2009a). A class of estimators using auxiliary information for estimating finite population variance in presence of Measurements Errors. *Communications in Statistical Theory and Methods*, 38(5), 734-741.

Singh, H. P., & Karpe, N. (2009b). A General Procedure for Estimating the General parameter using Auxiliary Information in Presence of Measurement Errors. *Communications in Korean Journal of Applied Statistics*, (Accepted for publication).

Singh, H. P., & Karpe, N. (2009c). On the estimation of ratio and product of two population means using supplementary information in presence of measurement errors. *Statistica*, (Accepted for publication).

Srivastava, A. K., & Shalabh. (2001). Effect of measurement errors on the regression method of estimation in survey sampling. *Journal of Statistical Research*, 35(2), 35-44.

Srivastava, S. K. (1971). A generalized estimator for the mean of a finite population using multi-auxiliary information. *Journal of the American Statistical Association*, 66, 404-407.

Srivastava, S. K. (1980). A class of estimators using auxiliary information in sample surveys. *Canadian Journal of Statistics*, 8, 253-254.

Sukhatme, P. V., Sukhatme, B. V., Sukhatme, S., & Asok, C. (1984). *Sampling Theory of Surveys With Applications*. Des Moines: Iowa State University Press.