Class(es) of Factor-Type Estimator(s) in Presence of Measurement Error

Diwakar Shukla
Dr. Hari Singh Gour University, Sagar, M. P., India

Sharad Pathak
Dr. Hari Singh Gour University, Sagar, M. P., India

Narendra Singh Thakur
Banasthali University, Rajasthan, India

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
DOI: 10.22237/jmasm/1351742700
Available at: http://digitalcommons.wayne.edu/jmasm/vol11/iss2/6

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.
When data is collected via sample survey it is assumed whatever is reported by a respondent is correct. However, given the issues of prestige bias, personal respect and honor, respondents’ self-reported data often produces over- or under-estimated values as opposed to true values regarding the variables under question. This causes measurement error to be present in sample values. This article considers the factor-type estimator as an estimation tool and examines its performance under a measurement error model. Expressions of optimization are derived and theoretical results are supported by numerical examples.

Key words: Measurement error, factor-type estimator, bias, mean squared error.

Introduction
Sample surveys result in an efficiency of estimators on the basis of collected or simulated data. Data for analyses may originate from various sampling sources, such as, simple random sampling, stratified sampling, systematic sampling or cluster sampling. Estimation methods are typically analyzed under the assumption that observations collected are true and without error; however, real life data, gathered through sample surveys contains errors due to memory failure, prestige bias, over reporting patterns, unwillingness to respond, desire for secrecy and other reasons. The deviation between true and observed values is error and is technically termed measurement error. Measurement error may be characterized as the difference between the value of a variable provided by the respondent and the true value of the same variable. The total survey error of a statistic with measurement error has both fixed error (bias) and variable error (variance) over repeated trials of the survey (Cochran, 2005; Sukhatme, et al., 1984). Figure 1 illustrates the concept of measurement error.

There are two possibilities for incompleteness in a survey: incorrect response or non-response. Measurement bias provides a systematic pattern in the difference between the respondent’s answers to a question and the correct answer. For example, a respondent may forget to report a few specific income sources resulting in total reported income being lower than actual. Measurement variance reflects random variation in answers provided to an interviewer while asking the same question, that is, often the same respondent provides different answers to the same question when asked repeatedly. Several methods are available in the survey sampling literature to handle non-response, including the revisit method, imputation methods, auxiliary sources utilization method and the neighboring units manipulation methods, however, when a respondent provides incorrect information regarding a variable, additional techniques are required. This study considers this aspect and deals with mean estimation under measurement error.

Manisha and Singh (2001) examined population mean estimation in the presence of measurement errors; they provided an effect of measurement errors on a new estimator obtained as a combination of ratio and mean per unit estimator. Shalabh (1997) studied a ratio method of estimation in the presence of measurement errors. Singh and Shukla (1987) presented a


Study Notations and Assumptions
Assume a set of information obtained via a simple random sampling procedure on three characteristics $Y$, $X_1$ and $X_2$. Suppose $(y_i, x_{i1}, x_{i2})$ are observational values and $(Y_i, X_{i1}, X_{i2})$ are corresponding true values for the characteristics respectively. Notations for this study are:

- $\bar{Y}, \bar{X}_1$ and $\bar{X}_2$: Population parameters;
- $\bar{y}, \bar{x}_1$ and $\bar{x}_2$: Mean per unit estimates for a simple random sample of size $n$;
- $n$: Sample size;
- $f$: Sampling friction ($f = n/N$);
- $N$: Population size;
- $U_i$: Measurement error for $Y$;
- $V_i$: Measurement error for $X_1$;
- $T_i$: Measurement error for $X_2$;
- $\sigma_U^2$, $\sigma_Y^2$ and $\sigma_T^2$: Variances for measurement error;
- $\sigma_{Y_i}^2$, $\sigma_{X_{i1}}^2$ and $\sigma_{X_{i2}}^2$: Variances of variable $Y$, $X_1$ and $X_2$ respectively;
- $\rho_{01}$: Correlation between variable $Y$ and $X_1$;
- $\rho_{02}$: Correlation between variable $Y$ and $X_2$;
- $\rho_{12}$: Correlation between variable $X_1$ and $X_2$;
CLASSES OF FACTOR-TYPE ESTIMATORS IN PRESENCE OF MEASUREMENT ERROR

$C_Y = \frac{\sigma_Y}{\bar{Y}}$: Coefficient of variation for variable $Y(C_0)$;

$C_{X1} = \frac{\sigma_{X1}}{\bar{X}_1}$: Coefficient of variation for variable $X_1(C_1)$; and

$C_{X2} = \frac{\sigma_{X2}}{\bar{X}_2}$: Coefficient of variation for variable $X_2(C_2)$.

New notations are:

\[
W_Y = \frac{\sum(Y_i - \bar{Y})}{\sqrt{n}},
\]

\[
W_{X1} = \frac{\sum(X_{1i} - \bar{X}_1)}{\sqrt{n}},
\]

\[
W_{X2} = \frac{\sum(X_{2i} - \bar{X}_2)}{\sqrt{n}},
\]

\[
W_U = \frac{\sum(U_i)}{\sqrt{n}},
\]

\[
W_V = \frac{\sum(V_i)}{\sqrt{n}},
\]

\[
W_T = \frac{\sum(T_i)}{\sqrt{n}}.
\]

existing estimators: mean per unit estimator

The mean per unit (or mean) estimator is a well-known estimator, and in the setup of measurement error, $\bar{y}_i = n^{-1} \sum_i (U_i + Y_i)$, is shown in (3.2). The bias for $\bar{y}$ is zero, that is,

\[
E(\bar{y}) = \frac{1}{n} \sum_i (U_i + Y_i) = \bar{Y} \tag{4.1a}
\]

and the variance is

\[
Variance(\bar{y}) = \frac{\sigma_y^2}{n} \left[ 1 + \frac{\sigma_U^2}{\sigma_Y^2} \right] \tag{4.1b}
\]

To estimate $\bar{Y}$, the sample statistic $\bar{y}$, which provides an unbiased estimator, can be used. In mean per unit estimator $\bar{y}$, no additional information is required. Several methods exist for using the auxiliary $X$ characteristic.
Existing Estimators: Shalabh (1997) Estimator
Shalabh (1997) proposed an estimator that is a ratio estimator studied under measurement error.

\[ t_R = \frac{\bar{Y}}{\bar{X}} \mu_X \]  

(4.2)

Where the bias of \( t_R \) is

\[ B(t_R) = \frac{\mu_Y}{n} \left[ C_X (C_X - \rho C_Y) + \frac{\sigma^2}{\mu_X} \right] \]  

(4.2a)

and the mean squared error is

\[ \text{MSE} (t_R) = \frac{\sigma^2_Y}{n} \left[ 1 - \frac{C_X}{C_Y} \left( 2\rho - \frac{C_X}{C_Y} \right) \right] + \frac{1}{n} \left[ \sigma^2_Y + \left( \frac{\mu_Y}{\mu_X} \right)^2 \sigma^2_Y \right] \]  

(4.2b)

where \( \mu_X \) denotes the population mean of \( X \).

Existing Estimators: Manisha and Singh (2001) Estimator
Manisha and Singh (2001) proposed the estimator

\[ \bar{y}_\theta = \theta t_R + (1 - \theta) \bar{y} \]  

(4.3)

where the bias of \( \bar{y}_\theta \) is

\[ B(\bar{y}_\theta) = \theta \left[ \frac{\mu_Y}{n \mu_X} \left( \sigma^2_Y + \sigma^2_Y \right) - \frac{1}{n \mu_X} \rho \sigma_X \sigma_Y \right] \]  

(4.3a)

and the mean squared error is

\[ B(\bar{y}_\theta) = \frac{\sigma^2_Y}{n} \left[ 1 - \theta \frac{C_X}{C_Y} \left( 2\rho - \theta \frac{C_X}{C_Y} \right) \right] + \frac{1}{n} \left[ \theta^2 \frac{\mu^2_Y}{\mu_X^2} \sigma^2_Y + \sigma^2_Y \right] \]  

(4.3b)

Proposed Estimator(s)
The two parameter F-T estimators proposed are:

\[ \bar{y}_{FT1}^* = \bar{y} T_1 T_2 \]
\[ \bar{y}_{FT2}^* = \bar{y} T_1 T_2^{-1} \]
\[ \bar{y}_{FT3}^* = \bar{y} T_1^{-1} T_2 \]

(5.1)

where \( f = n/N \) and

\[ T_1 = \left( \frac{A_1 + C_1}{(A_1 + fB)X_i + fB_iX_i} \right) \]
\[ A = (K - 1)(K - 2) \]
\[ B = (K - 1)(K - 4) \]
\[ C = (K - 2)(K - 3)(K - 4) \]

(5.1b)

Thus,

\[ \bar{y}_{FT1}^* = \bar{y} \left( \frac{(A_1 + C_1)X_i + fB_iX_i}{(A_1 + fB_i)X_i + C_iX_i} \right) \]
\[ \bar{y}_{FT2}^* = \bar{y} \left( \frac{(A_1 + C_1)X_i + fB_iX_i}{(A_1 + fB_i)X_i + C_iX_i} \right) \]
\[ \bar{y}_{FT3}^* = \bar{y} \left( \frac{(A_1 + C_1)X_i + fB_iX_i}{(A_1 + fB_i)X_i + C_iX_i} \right) \]

(5.2a

(5.2b

(5.2c

where \( \theta \) is a characterizing scalar and \( U \) and \( V \) are measurement errors corresponding to \( Y \) and \( X \) respectively.
Note that there is a combination of \( iK \) where \((1, 2)\) \( i \) is constant, it is important to choose suitably so that the resultant mean squared error of the proposed estimators may be minimized to the greatest extent. Using the proposed estimator many different estimators may be obtained because an estimator exists for each combination of \((K_1, K_2)\).

Properties of the Proposed Estimator(s)

For the approximation assume that:

\[
\delta_0 = \frac{1}{\sqrt{n}} (W_Y + W_U); \\
\delta_1 = \frac{1}{\sqrt{n}} (W_{X_1} + W_Y); \\
\delta_2 = \frac{1}{\sqrt{n}} (W_{X_2} + W_T); \\
\theta_i = \alpha_i - \beta_i; \\
f = \frac{n}{N}; \\
\alpha_i = \frac{fB_i}{A_i + fB_i + C_i}; \\
\beta_i = \frac{C_i}{A_i + fB_i + C_i}.
\]

Theorem 6.1

The estimator \( \hat{Y}_{FT1} \) up to first order of approximation can be expressed as:

\[
\hat{Y}_{FT1} = \bar{y} + \delta_0 + \frac{\bar{y} \theta_1}{\bar{X}_1} \delta_1 + \frac{\bar{y} \theta_2}{\bar{X}_2} \delta_2 - \bar{y} \frac{\beta_1 \theta_1}{\bar{X}_1^2} \delta_1^2 - \frac{\bar{y} \beta_2 \theta_2}{\bar{X}_2^2} \delta_2^2 + \theta_1 \bar{X}_1 \delta_0 \delta_1 + \theta_2 \bar{X}_2 \delta_0 \delta_2 + \bar{y} \theta_1 \theta_2 \delta_1 \delta_2
\]

and the bias of \( \hat{Y}_{FT1} \) is:

<table>
<thead>
<tr>
<th>( K_1 )</th>
<th>( K_2 )</th>
<th>( t_1 = \bar{y} \frac{\bar{X}_1}{\bar{x}_1} )</th>
<th>( t_2 = \bar{y} \frac{\bar{X}_1}{\bar{x}_1} \frac{\bar{x}_2}{\bar{x}_2} )</th>
<th>( t_3 = \bar{y} \frac{\bar{X}_1}{\bar{x}_1} \frac{N \bar{X}_2 - n \bar{x}_2}{(N-n) \bar{x}_2} )</th>
<th>( t_4 = \bar{y} \frac{\bar{X}_1}{\bar{x}_1} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>( t_5 = \bar{y} \frac{\bar{X}_1}{\bar{x}_1} \frac{\bar{x}_2}{\bar{x}_2} )</td>
<td>( t_6 = \bar{y} \frac{\bar{X}_1}{\bar{x}_1} \frac{\bar{x}_2}{\bar{x}_2} )</td>
<td>( t_7 = \bar{y} \frac{\bar{X}_1}{\bar{x}_1} \frac{N \bar{X}_2 - n \bar{x}_2}{(N-n) \bar{x}_2} )</td>
<td>( t_8 = \bar{y} \frac{\bar{X}_1}{\bar{x}_1} )</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>( t_9 = \bar{y} \frac{N \bar{X}_1 - n \bar{x}_1}{(N-n) \bar{X}_1} \frac{\bar{x}_2}{\bar{x}_2} )</td>
<td>( t_{10} = \bar{y} \frac{N \bar{X}_1 - n \bar{x}_1}{(N-n) \bar{X}_1} \frac{\bar{x}_2}{\bar{x}_2} )</td>
<td>( t_{11} = \bar{y} \frac{N \bar{X}_1 - n \bar{x}_1}{(N-n) \bar{X}_1} \frac{N \bar{X}_2 - n \bar{x}_2}{(N-n) \bar{x}_2} )</td>
<td>( t_{12} = \bar{y} \frac{N \bar{X}_1 - n \bar{x}_1}{(N-n) \bar{X}_1} )</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>( t_{13} = \bar{y} \frac{\bar{x}_2}{\bar{x}_2} )</td>
<td>( t_{14} = \bar{y} \frac{\bar{x}_2}{\bar{x}_2} )</td>
<td>( t_{15} = \bar{y} \frac{N \bar{X}_2 - n \bar{x}_2}{(N-n) \bar{x}_2} )</td>
<td>( \bar{y} )</td>
</tr>
</tbody>
</table>
Bias(\(\bar{y}_{FT1}\)) = \(\frac{n}{\bar{y}} \{\theta_1 \rho_{01} C_0 C_1 + \theta_1 \rho_{02} C_0 C_2 + \theta_1 \rho_{12} C_1 C_2 - \beta_1 \theta_1 C_1^2 \left(1 + \frac{\sigma_Y^2}{\sigma_{x_1}}\right) - \beta_2 \theta_2 C_2^2 \left(1 + \frac{\sigma_Y^2}{\sigma_{x_2}}\right)\}\)

\(\bar{y}_{FT1} - \bar{y} = \delta_0 + \frac{\bar{y} \theta_1}{X_1} \delta_1 + \frac{\bar{y} \theta_2}{X_2} \delta_2 \)

\(-\frac{\bar{y} \beta_1 \theta_1}{X_1^2} \delta_1^2 - \frac{\bar{y} \beta_2 \theta_2}{X_2^2} \delta_2^2 \)

\(+ \frac{\theta_1}{X_1} \delta_0 \delta_1 + \frac{\theta_2}{X_2} \delta_0 \delta_2 + \frac{\bar{y} \theta_1 \theta_2}{X_1 X_2} \delta_1 \delta_2 \)

Thus, for the solution:

\[\left[\bar{y}_{FT1} - \bar{y}\right]^2 = \left\{\delta_0 + \frac{\bar{y} \theta_1}{X_1} \delta_1 + \frac{\bar{y} \theta_2}{X_2} \delta_2 \right.\]

\(-\frac{\bar{y} \beta_1 \theta_1}{X_1^2} \delta_1^2 - \frac{\bar{y} \beta_2 \theta_2}{X_2^2} \delta_2^2 + \theta_1 \delta_0 \delta_1 \]

\(+ \frac{\theta_2}{X_2} \delta_0 \delta_2 + \frac{\bar{y} \theta_1 \theta_2}{X_1 X_2} \delta_1 \delta_2 \right\}^2 \]

and

\[E[\bar{y}_{FT1} - \bar{y}]^2 = \frac{1}{n} \left(\sigma_Y^2 + \sigma_{U}^2\right)\]

\(+\frac{\bar{y}^2}{n} \left\{\frac{\bar{y}^2 C_1^2}{\sigma_{x_1}} + \frac{\bar{y}^2 C_2^2}{\sigma_{x_2}} \right\} \]

\(+2 \theta_1 \rho_{01} C_0 C_1 + 2 \theta_2 \rho_{02} C_0 C_2 + 2 \theta_1 \rho_{12} C_1 C_2 \}

Theorem 6.2

The estimator \(\bar{y}_{FT2}\) up to first order of approximation can be expressed as:

\[\bar{y}_{FT2} = \bar{y} + \delta_0 + \frac{\bar{y} \theta_1}{X_1} \delta_1 - \frac{\bar{y} \theta_2}{X_2} \delta_2 \]

\(-\frac{\bar{y} \beta_1 \theta_1}{X_1^2} \delta_1^2 + \frac{\bar{y} \beta_2 \theta_2}{X_2^2} \delta_2^2 \]

\(+ \frac{\theta_1}{X_1} \delta_0 \delta_1 + \frac{\theta_2}{X_2} \delta_0 \delta_2 + \frac{\bar{y} \theta_1 \theta_2}{X_1 X_2} \delta_1 \delta_2 \]

The bias of \(\bar{y}_{FT2}\) is:
Bias($\hat{y}_{FT2}^*$) = 
\[ \frac{\bar{Y}}{n} \left\{ \theta_1 \rho_{01} C_0 C_1 - \theta_2 \rho_{02} C_0 C_2 - \theta_1 \theta_2 \rho_{12} C_1 C_2 \right\} 
- \beta_1 \theta_1 C_1^2 \left( 1 + \frac{\sigma^2}{\sigma_X^2} \right) + \alpha_2 \theta_2 C_2^2 \left( 1 + \frac{\sigma^2}{\sigma_X^2} \right) \]
(6.5)

and the mean squared error of $\hat{y}_{FT2}^*$ is:

MSE($\hat{y}_{FT2}^*$) = \[ \frac{1}{n}(\sigma^2_y + \sigma^2_u) \]
\[ + \frac{\bar{Y}}{n} \left[ \theta_1^2 C_1^2 \left( 1 + \frac{\sigma^2}{\sigma_X^2} \right) + \theta_2^2 C_2^2 \left( 1 + \frac{\sigma^2}{\sigma_X^2} \right) \right] \]
\[ + 2 \theta_1 \rho_{01} C_0 C_1 - 2 \theta_2 \rho_{02} C_0 C_2 - 2 \theta_1 \theta_2 \rho_{12} C_1 C_2 \].
(6.6)

Proof 6.2

From (5.2b) the proposed estimator is

\[ \hat{y}_{FT2}^* = \frac{\bar{Y}}{(A + fB)(A + fB_1)} \]
\[ \left[ (A + fB_1)X_1 + C_1 \right] \left[ (A + fB_2)X_2 + C_2 \right] \]
\[ \left[ (A + fB_1)X_1 + C_1 \right] \left[ (A + fB_2)X_2 + fB_2X_2 \right] \]
\[ \frac{\bar{Y} (A + C_1)X_1 + fB_1X_1 (A + fB_2)X_2 + C_2X_2}{(A + fB_1)X_1 + C_1X_1 (A + fB_2)X_2 + fB_2X_2} \]
\[ \frac{\bar{Y} + \delta_0 + \frac{\bar{Y} \theta_1}{X_1} \delta_1 - \frac{\bar{Y} \theta_2}{X_2} \delta_2 \left( 1 + \frac{\sigma^2}{\sigma_X^2} \right) + \alpha_2 \theta_2 C_2^2 \left( 1 + \frac{\sigma^2}{\sigma_X^2} \right) \right\} \]
\[ + \left\{ 1 + \frac{\alpha_2 \delta_2}{X_2} \right\} \left\{ 1 + \frac{\beta_1 \delta_1}{X_1} \right\} \]
(6.7)

Solving the equations results in:

\[ \hat{y}_{FT2}^* = \bar{Y} + \delta_0 + \frac{\bar{Y} \theta_1}{X_1} \delta_1 - \frac{\bar{Y} \theta_2}{X_2} \delta_2 \]
\[ - \frac{\bar{Y} \beta_1 \theta_1}{X_1^2} \delta_1^2 + \frac{\bar{Y} \alpha_2 \theta_2}{X_2^2} \delta_2^2 \]
\[ + \frac{\theta_1}{X_1} \delta_1 \delta_2 + \frac{\theta_2}{X_2} \delta_0 \delta_2 \]
\[ - \frac{\bar{Y} \theta_1 \theta_2}{X_1 X_2} \delta_1 \delta_2 \]
(6.8)

Based on the solution:

\[ E\left[ \hat{y}_{FT2}^* - \bar{Y} \right]^2 = \frac{1}{n}(\sigma^2_y + \sigma^2_u) \]
\[ + \frac{\bar{Y}^2}{n} \left[ \theta_1^2 C_1^2 \left( 1 + \frac{\sigma^2}{\sigma_X^2} \right) + \theta_2^2 C_2^2 \left( 1 + \frac{\sigma^2}{\sigma_X^2} \right) \right] \]
\[ + 2 \theta_1 \rho_{01} C_0 C_1 - 2 \theta_2 \rho_{02} C_0 C_2 - 2 \theta_1 \theta_2 \rho_{12} C_1 C_2 \].
(6.9)
Theorem 6.3

The estimator $\hat{y}_{FT3}^*$, up to first order of approximation, can be expressed as:

$$\hat{y}_{FT3}^* = \bar{Y} + \delta_0 - \frac{\bar{Y}\theta_1}{X_1} \delta_1 + \frac{\bar{Y}\theta_2}{X_2} \delta_2$$

$$+ \frac{\bar{Y}\alpha_1}{X_1^2} \delta^2 - \frac{\bar{Y}\beta_1}{X_2^2} \delta^2$$

$$- \frac{\theta_1}{X_1} \delta_0 \delta_1 + \frac{\theta_2}{X_2} \delta_0 \delta_2 - \frac{\bar{Y}\theta_1 \theta_2}{X_1 X_2} \delta \delta_2$$

(6.7)

and the bias of $\hat{y}_{FT3}^*$ is:

$$\text{Bias}(\hat{y}_{FT3}^*) = \frac{\bar{Y}}{n} \left\{ \theta_2 \rho_{02} C_0 C_2 - \theta_2 \theta_0 C_0 C_1 \right\}$$

$$- \theta_1 \theta_2 \rho_{12} C_1 C_2 + \alpha_1 \theta_1 \theta_1 \left( 1 + \sigma_Y^2 \right)$$

$$- \beta_2 \theta_2 \theta_2 \left( 1 + \frac{\sigma_Y^2}{\sigma_{X_1}^2} \right) \right\}.$$  

(6.8)

The mean squared error of $\hat{y}_{FT3}^*$ is:

$$\text{MSE}(\hat{y}_{FT3}^*) = \frac{1}{n} \left( \sigma_Y^2 + \sigma_0^2 \right)$$

$$+ \frac{\bar{Y}^2}{n} \left\{ \theta_1^2 C_1 \left( 1 + \frac{\sigma_Y^2}{\sigma_{X_1}^2} \right) + \theta_2^2 C_2 \left( 1 + \frac{\sigma_Y^2}{\sigma_{X_2}^2} \right) \right\}$$

$$- 2 \theta_1 \theta_2 \rho_{01} C_0 C_1 + 2 \theta_2 \rho_{02} C_0 C_2 - 2 \theta_1 \theta_2 \rho_{12} C_1 C_2 \}. $$

(6.9)

Proof 6.3

From (5.2c) the proposed estimator is

$$\hat{y}_{FT3}^* = \left\{ \bar{Y} + \delta_0 \right\} \left\{ 1 + \frac{\alpha_2 \delta}{X_2} \right\} \left\{ 1 + \frac{\beta \delta}{X_1} \right\}$$

$$\left\{ 1 + \frac{\sigma_1^2}{\sigma_{X_1}^2} \right\} \left\{ 1 + \frac{\sigma_2^2}{\sigma_{X_2}^2} \right\},$$

which results in

$$\hat{y}_{FT3}^* = \bar{Y} + \delta_0 - \frac{\bar{Y}\theta_1}{X_1} \delta_1 + \frac{\bar{Y}\theta_2}{X_2} \delta_2$$

$$+ \frac{\bar{Y}\alpha_1}{X_1^2} \delta^2 - \frac{\bar{Y}\beta_2}{X_2^2} \delta^2$$

$$- \frac{\theta_1}{X_1} \delta_0 \delta_1 + \frac{\theta_2}{X_2} \delta_0 \delta_2 - \frac{\bar{Y}\theta_1 \theta_2}{X_1 X_2} \delta \delta_2,$$

and

$$E\left[ \hat{y}_{FT3}^* - \bar{Y} \right] =$$

$$\begin{bmatrix}
\delta_0 - \bar{Y}\theta_1 / X_1 + \bar{Y}\theta_2 / X_2 \\
\bar{Y}\alpha_1 / X_1 \delta^2 - \bar{Y}\beta_2 / X_2 \delta^2 \\
- \bar{Y}\theta_1 / X_1 \delta_0 \delta_1 + \bar{Y}\theta_2 / X_2 \delta_0 \delta_2 - \bar{Y}\theta_1 \theta_2 / X_1 X_2 \delta \delta_2
\end{bmatrix}.$$
CLASSES OF FACTOR-TYPE ESTIMATORS IN PRESENCE OF MEASUREMENT ERROR

\[
E[\bar{Y}_{FT3} - \bar{Y}] = \frac{\bar{Y}}{n} \{ \theta_2 \rho_{02} C_0 C_2 - \theta_1 \rho_{01} C_0 C_1 - \theta_1 \theta_2 \rho_{12} C_1 C_2 + \alpha_1 \theta_1 C^2_i \left(1 + \frac{\sigma^2_v}{\sigma^2_{x_i}} \right) \\
- \beta_2 \theta_2 C^2_i \left(1 + \frac{\sigma^2_t}{\sigma^2_{x_i}} \right) \}.
\]

From which follows

\[
[\bar{Y}_{FT3} - \bar{Y}] \gamma^2 = \left[ \delta_0 - \frac{\bar{Y} \theta_1}{X^2_1} \delta_1 + \frac{\bar{Y} \theta_2 \rho_{12}}{X^2_2} \delta_2 + \frac{\bar{Y} \alpha \theta_1}{X^2_1} \delta^2_1 \\
- \frac{\bar{Y} \beta \theta_2 \rho_{12}}{X^2_2} \delta^2_2 - \frac{\theta_1}{X^2_1} \delta_1 \delta_1 + \frac{\theta_2}{X^2_2} \delta_2 \delta_2 - \frac{\bar{Y} \theta_1 \theta_2 \rho_{12}}{X^2_1 X^2_2} \delta_1 \delta_2 \right]^2
\]

and

\[
E[\bar{Y}_{FT3} - \bar{Y}] \gamma^2 = E \left[ \delta_0 - \frac{\bar{Y} \theta_1}{X^2_1} \delta_1 + \frac{\bar{Y} \theta_2 \rho_{12}}{X^2_2} \delta_2 \\
+ \frac{\bar{Y} \alpha \theta_1}{X^2_1} \delta^2_1 - \frac{\bar{Y} \beta \theta_2 \rho_{12}}{X^2_2} \delta^2_2 - \frac{\theta_1}{X^2_1} \delta_1 \delta_1 \\
+ \frac{\theta_2}{X^2_2} \delta_2 \delta_2 - \frac{\bar{Y} \theta_1 \theta_2 \rho_{12}}{X^2_1 X^2_2} \delta_1 \delta_2 \right]^2.
\]

The solution of which results in:

\[
E[\bar{Y}_{FT3} - \bar{Y}] \gamma^2 = \frac{1}{n} \left( \sigma^2_v + \sigma^2_u \right) \\
+ \frac{\bar{Y}^2}{n} \left[ \theta^2_1 C^2_i \left(1 + \frac{\sigma^2_v}{\sigma^2_{x_i}} \right) + \theta^2_2 C^2_i \left(1 + \frac{\sigma^2_t}{\sigma^2_{x_i}} \right) \\
- 2\theta_1 \rho_{01} C_0 C_1 + 2\theta_2 \rho_{02} C_0 C_2 - 2\theta_1 \theta_2 \rho_{12} C_1 C_2 \right].
\]

Minimum Mean Squared Error & Optimal Choices for the Proposed Estimator(s)

The mean squared error of the proposed estimators \(\bar{Y}_{FT1}\), \(\bar{Y}_{FT2}\) and \(\bar{Y}_{FT3}\) shown in (6.3), (6.6) and (6.9) respectively, are functions with unknown parameter \(\theta_i\); \(i = 1, 2\), whereas \(\theta_i\) is a function of \(K\) solely. Thus, it is practical to calculate an optimum value of \(K\) in such a way that the mean squared error of the resultant proposed estimator becomes least.

Consider \(\bar{Y}_{FT1}\), notice the minimum mean squared error. On differentiation of \(MSE(\bar{Y}_{FT1})\) with respect to \(\theta_1\) and \(\theta_2\) and equating to zero (assuming \(\theta_i \neq 0\)), two simultaneous equations result:

\[
C^2_1 \left(1 + \frac{\sigma^2_v}{\sigma^2_{x_1}} \right) \theta_1 + \rho_{12} C_1 C_2 \theta_2 + \rho_{01} C_0 C_1 = 0
\]

for \(\frac{\partial}{\partial \theta_1} \left[ E[\bar{Y}_{FT1} - \bar{Y}]^2 \right] = 0 \)

(7.1)

and

\[
C^2_2 \left(1 + \frac{\sigma^2_v}{\sigma^2_{x_2}} \right) \theta_2 + \rho_{12} C_1 C_2 \theta_1 + \rho_{02} C_0 C_2 = 0
\]

For \(\frac{\partial}{\partial \theta_2} \left[ E[\bar{Y}_{FT1} - \bar{Y}]^2 \right] = 0 \).

(7.2)

From (7.1) and (7.2) the values of \(\theta_1\) and \(\theta_2\) are:

\[
\hat{\theta}_{(1)} = \frac{C_0}{C_1} \left[ \frac{\rho_{02} \rho_{12} - \rho_{01} \left(1 + \frac{\sigma^2_v}{\sigma^2_{x_1}} \right) \left(1 + \frac{\sigma^2_v}{\sigma^2_{x_2}} \right) - \rho_{12}^2}{\left(1 + \frac{\sigma^2_v}{\sigma^2_{x_1}} \right) \left(1 + \frac{\sigma^2_v}{\sigma^2_{x_2}} \right) - \rho_{12}^2} \right]
\]

(7.3)

The optimum values \(\hat{\theta}_{(1)} = \Delta_1\) and \(\hat{\theta}_{(2)} = \Delta_2\), for example, provide a minimum
mean squared error to \( \hat{y}_{FT1}^* \), where the second derivative is positive. Similarly, \( \hat{\theta}_3 = \hat{\theta}_1 \); \( \hat{\theta}_4 = (-1) \hat{\theta}_2 \) and \( \hat{\theta}_5 = (-1) \hat{\theta}_1 \); \( \hat{\theta}_6 = \hat{\theta}_2 \) are optimal choices corresponding to \( \hat{y}_{FT2}^* \) and \( \hat{y}_{FT3}^* \) respectively. These \( \theta_\bullet \) provide polynomials in terms of \( K \) to produce values for which the mean squared error will be optimum.

Empirical Study
This illustration demonstrates how to evaluate the gain in efficiencies (in terms of mean squared error) obtained by the proposed estimators. To evaluate the performance of the various estimators discussed, a population is considered (see Appendix A); required information is shown in Table 8.1.

Table 8.1: Population Parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \bar{Y} )</td>
<td>63.396</td>
<td>( n )</td>
<td>50</td>
</tr>
<tr>
<td>( \bar{X}_1 )</td>
<td>48.136</td>
<td>( N )</td>
<td>250</td>
</tr>
<tr>
<td>( \bar{X}_2 )</td>
<td>56.364</td>
<td>( f )</td>
<td>0.2</td>
</tr>
<tr>
<td>( C_0 )</td>
<td>0.2899</td>
<td>( \rho_{01} )</td>
<td>0.8544</td>
</tr>
<tr>
<td>( C_1 )</td>
<td>0.4637</td>
<td>( \rho_{02} )</td>
<td>0.8249</td>
</tr>
<tr>
<td>( C_2 )</td>
<td>0.4085</td>
<td>( \rho_{12} )</td>
<td>0.8289</td>
</tr>
</tbody>
</table>

Table 8.2: Percent Relative Efficiency of Various Estimators with respect to Mean per Unit Estimator

<table>
<thead>
<tr>
<th>Estimator(s)</th>
<th>( PRE (\bullet) ) with respect to MPU</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( MSE(\hat{y}_{FT1}) )</td>
</tr>
<tr>
<td>( \hat{y} )</td>
<td>100</td>
</tr>
<tr>
<td>( t_1 )</td>
<td>40.75</td>
</tr>
<tr>
<td>( t_2 )</td>
<td>27.41</td>
</tr>
<tr>
<td>( t_3 )</td>
<td>93.26</td>
</tr>
<tr>
<td>( t_4 )</td>
<td>36.30</td>
</tr>
<tr>
<td>( t_5 )</td>
<td>21.35</td>
</tr>
<tr>
<td>( t_6 )</td>
<td>11.53</td>
</tr>
<tr>
<td>( t_7 )</td>
<td>23.87</td>
</tr>
<tr>
<td>( t_8 )</td>
<td>21.95</td>
</tr>
<tr>
<td>( t_9 )</td>
<td>38.35</td>
</tr>
<tr>
<td>( t_{10} )</td>
<td>28.38</td>
</tr>
<tr>
<td>( t_{11} )</td>
<td>104.25</td>
</tr>
<tr>
<td>( t_{12} )</td>
<td>106.28</td>
</tr>
<tr>
<td>( t_{13} )</td>
<td>42.40</td>
</tr>
<tr>
<td>( t_{14} )</td>
<td>25.71</td>
</tr>
<tr>
<td>( t_{15} )</td>
<td>106.23</td>
</tr>
<tr>
<td>( \text{Opt} (\hat{y}_{FT1}^*) )</td>
<td>113.05</td>
</tr>
<tr>
<td>( \text{Opt} (\hat{y}_{FT2}^*) )</td>
<td>92.28</td>
</tr>
<tr>
<td>( \text{Opt} (\hat{y}_{FT3}^*) )</td>
<td>95.02</td>
</tr>
</tbody>
</table>
Results

Three different approaches were examined as tools for estimating in the presence of measurement error. Results indicate that the proposed approaches are effective and efficient over many existing strategies. The multiple choices for $K$ are accessible via:

\[
(\Delta_1 + 1)K_1^3 + (f\Delta_1 - f - 8\Delta_1 - 9)K_1^2 \\
+ (23\Delta_1 - 5f\Delta_1 + 5f + 26)K_1 \\
+ (4f\Delta_1 - 22\Delta_1 - 4f - 24) = 0
\]

and

\[
(\Delta_2 + 1)K_2^3 + (f\Delta_2 - f - 8\Delta_2 - 9)K_2^2 \\
+ (23\Delta_2 - 5f\Delta_2 + 5f + 26)K_2 \\
+ (4f\Delta_2 - 22\Delta_2 - 4f - 24) = 0.
\]

Polynomials (9.1) and (9.2), which are obtained from (7.3), provide three roots for lesser mean squared error. As discussed for $K_1 = K_2 = 4$ the proposed classes provide mean per unit estimators, thus those values are unbiased estimators.

For the estimator $\bar{y}_{FT1}^*$, the optimum values of the characterizing scalar are $(K_1)_1 = 4.4951$, $(K_1)_2 = 3.1167$, $(K_1)_3 = 1.8111$, $(K_2)_1 = 4.5063$, $(K_2)_2 = 3.1133$ and $(K_2)_3 = 1.8096$. For $\bar{y}_{FT2}$, the values are $(K_1)_4 = (K_1)_1$, $(K_1)_5 = (K_1)_2$, $(K_1)_6 = (K_1)_3$ and $(K_2)_4 = 1.8857$, and for the $\bar{y}_{FT3}$ estimator values are, $(K_1)_7 = 11.5039$, $(K_2)_7 = (K_2)_1$, $(K_2)_8 = (K_2)_2$ and $(K_2)_9 = (K_2)_3$ with the remaining imaginary roots.

Tables 8.2 and 8.3 show that the proposed estimator is efficient over many currently used estimators, including the Manisha and Singh (2001) and the Shalabh (1997).

Conclusion

Based on study results, the proposed estimator(s) have several benefits over estimators currently used in research, including:

1. For different values of the characterizing scalar, there now exists a new estimation tool; and

2. The proposed class(es) provides a wide range for selecting the constant scalar by solving the associated polynomials and for root values estimators attains minimum mean squared error.

The proposed methodology is more effective, practicable and efficient, and may be recommended for use in practice.

References


