Consistent estimation of regression parameters under replicated ultrastructural model with non-normal errors

Shalabh* , Chandra Mani Paudel and Narinder Kumar

Department of Mathematics & Statistics, Indian Institute of Technology, Kanpur – 208016, India; Department of Statistics Tribhuvan University PN Campus, Pokhra, Nepal; Department of Statistics Panjab University Chandigarh-160014, India

(Received 22 May 2006; final version received 16 October 2007)

This article discusses the construction and efficiency properties of consistent estimators of regression parameters under replicated ultrastructural model with not necessarily normally distributed measurement errors. The variances of measurement errors associated with the study and explanatory variables are estimated from the replicated sample observations and are used for the consistent estimation of regression parameters. The asymptotic efficiency properties of the estimators are derived and analysed. The finite sample performance of the estimators is empirically studied through a Monte Carlo simulation.

Keywords: measurement errors; direct regression; reverse regression; replicated data; ultrastructural model; non-normal distribution

1. Introduction

The statistical modelling of data generally assumes that the observations on variables are correctly observed. In many practical situations, this assumption is violated and the data is contaminated with measurement errors. In the context of regression analysis, the least squares estimators of regression parameters have minimum variance in the class of linear and unbiased estimators. This remains true only as long as the observations are recorded without measurement errors. The same least squares estimators become biased as well as inconsistent estimators when the observations are recorded without measurement errors in the data. The consistent estimators of regression parameters can be derived by employing the methods of moments or maximum likelihood (under normality) and using some additional information in terms of measurement error variances, reliability ratios etc. Different types of additional information yield different
estimators, see e.g. Judge et al. [1], Fuller [2] and Cheng and Van Ness [3] for more details. The properties of the consistent estimators of regression parameters under known measurement error variances are studied in the literature under an unreplicated ultrastructural model, see e.g. Cheng and Van Ness [4,5] and Srivastava and Shalabh [6–8].

The additional information to derive the consistent estimators is obtained from different external sources like past experience of the experimenter or from some similar kind of experiments conducted in the past etc. The reliability of such information is always subjected to some uncertainties or sometimes it is not even available. Moreover, Kleeper and Leamer [9, p. 163] clearly states that; ‘Frequently, the additional information needed for identification is either not available or is not widely shared by researchers in the field’. Such problems can be avoided if this information can be derived from the sample itself. For instance, the measurement error variances can be estimated if replicated observations on the true values of study and explanatory variables are available. The estimated measurement error variances can then be utilized further to construct the consistent estimators of regression parameters. The maximum likelihood estimation of the regression parameters in a replicated structural model under normally distributed measurement errors is studied by Chan and Mak [10] and Isogawa [11]. They have not given any closed form of the estimators but their estimators are the solution of some equations. Also, Yum [12] studied the properties of estimators of regression parameters under a functional model, see also Wang [13], Buonaccorsi [14] and Schafer and Purdy [15] and references cited therein for other work in replicated measurement error models. Most of the literature in measurement error models deals either with the functional or the structural form. Dolby [16] synthesized both the forms in an ultrastructural model, see also Gleser [17,18].

The efficiency properties of the estimators under measurement error models are generally studied assuming that the measurement errors are normally distributed. When there is a departure from the normal distribution, the statistical inferences become invalid. When measurement errors are not necessarily normally distributed, Ketellapper [19], Srivastava and Shalabh [6–8] and Shalabh et al. [20] studied the performance of few estimators of regression parameters under the specification of an unreplicated ultrastructural model.

A natural question arises about the performance of consistent estimators of regression parameters under an ultrastructural model with replicated observations when measurement errors variances are estimated from the sample. Use of which of the measurement error variances – either associated with study variable or explanatory variable – yield better estimates is another interesting issue. Another question crops up at the design stage that how many replications are sufficient in practice to provide a good estimate of the unknown parameters? This article is an attempt to answer such queries.

We consider an ultrastructural model and assume the availability of replicated observations on the true values of explanatory and study variables in this paper. The measurement error variances associated with the explanatory and study variables are estimated using the replicated data and used further to construct the consistent estimators of regression parameters. The large sample asymptotic efficiency properties of these estimators are derived and analysed. Only the existence and finiteness of the first four central moments of distribution of various errors is assumed without associating any form of the probability distribution with them. Which of the information can be utilized to obtain a more efficient estimator is also addressed. The minimum number of replications needed in practice for the efficient estimation of regression parameters is addressed through a numerical illustration.

The organization of this article is as follows. The model and the estimators of the parameters are described in Section 2. In Section 3, the large sample asymptotic properties of the estimators are discussed. The finite sample properties of the estimators are studied through a Monte Carlo simulation experiment in Section 4. The derivations of main results are presented in the appendix.
2. The model and the estimators

The $i$th true but unobservable values of study variable $Y_i$ and explanatory variable $X_i$ are linearly related as:

$$Y_i = \alpha + \beta X_i; \quad i = 1, 2, \ldots, n$$

(1)

where $\alpha$ and $\beta$ are the unknown intercept and slope parameters, respectively of the model.

The $r$ replicated observations $y_{ij}$ and $x_{ij}$ on $Y_i$ and $X_i$, respectively are available as

$$
\begin{align*}
    y_{ij} &= Y_i + u_{ij}; \quad j = 1, 2, \ldots, r \\
    x_{ij} &= X_i + v_{ij}; \quad j = 1, 2, \ldots, r
\end{align*}
$$

(2)

where $u_{ij}$ and $v_{ij}$ are the measurement errors associated with $y_{ij}$ and $x_{ij}$, respectively.

Further, we assume that

$$X_i = m_i + w_i,$$

(3)

where $m_i$ is the mean of $X_i$ and $w_i$ is the associated random error component. This completes the specification of an ultrastructural model with replicated observations, see Dolby [16]. When $X_i$’s are non-stochastic, then the ultrastructural model reduces to the functional form of measurement error model. When $X_i$’s are stochastic with same mean and finite variance, then the ultrastructural model reduces to the structural form of measurement error model.

The measurement errors $u_{ij}$’s are assumed to be identically and independently distributed with mean 0, variance $\sigma_u^2$, third central moment $\sigma_u^3\gamma_{1u}$ and fourth central moment $(\gamma_{2u} + 3)\sigma_u^4$. Similarly, $v_{ij}$’s are assumed to be identically and independently distributed with mean 0, variance $\sigma_v^2$, third central moment $\sigma_v^3\gamma_{1v}$ and fourth central moment $(\gamma_{2v} + 3)\sigma_v^4$. The random error component $w_i$’s are also assumed to be identically and independently distributed with mean 0, variance $\sigma_w^2$, third and fourth central moments as $\sigma_w^3\gamma_{1w}$ and $(\gamma_{2w} + 3)\sigma_w^4$, respectively. The quantities $\gamma_{1z}$ and $\gamma_{2z}$ denote the Pearson’s coefficients of skewness and kurtosis, respectively associated with the distribution of a random variable $z$. Further, $u_{ij}$, $v_{ij}$ and $w_i$ are assumed to be jointly statistically independent of each other and also independent with the true values of the variables.

The direct least squares estimators of $\beta$ obtained by regressing $y_{ij}$ on $x_{ij}$ and $\overline{y}_i$ on $\overline{x}_i$ are $b_1 = S_{xy}/S_{xx}$ and $b_2 = B_{xy}/B_{xx}$, respectively. Both $b_1$ and $b_2$ are inconsistent for $\beta$. After adjusting them for their inconsistencies, the resulting consistent estimators of $\beta$ corresponding to $b_1$ and $b_2$ are

$$b_3 = \frac{S_{xy}}{S_{xx} - \hat{\sigma}_v^2}; \quad S_{xx} > \hat{\sigma}_v^2$$

(4)

and

$$b_4 = \frac{B_{xy}}{B_{xx} - (\hat{\sigma}_v^2/r)}; \quad B_{xx} > \frac{\hat{\sigma}_v^2}{r},$$

(5)

respectively, where $\hat{\sigma}_v^2 = r(S_{xx} - B_{xx})/(r - 1)$, $S_{xy} = \sum_{i=1}^n \sum_{j=1}^r (x_{ij} - \overline{x})(y_{ij} - \overline{y})/nr$, $S_{xx} = \sum_{i=1}^n \sum_{j=1}^r (x_{ij} - \overline{x})^2/nr$, $B_{xy} = \sum_{i=1}^n (\overline{x}_i - \overline{x})(\overline{y}_i - \overline{y})/n$, $B_{xx} = \sum_{i=1}^n (\overline{x}_i - \overline{x})^2/n$, $\overline{x} = \sum_{i=1}^n \sum_{j=1}^r x_{ij}/nr$, $\overline{y} = \sum_{i=1}^n \sum_{j=1}^r y_{ij}/nr$, $\overline{x}_i = \sum_{j=1}^r x_{ij}/r$ and $\overline{y}_i = \sum_{j=1}^r y_{ij}/r$.

The consistent estimators of $\alpha$ corresponding to $b_3$ and $b_4$ are

$$a_3 = \overline{y} - b_3\overline{x}$$

(6)

and

$$a_4 = \overline{y} - b_4\overline{x},$$

(7)

respectively.
Alternatively, the reverse regression of $x_{ij}$ on $y_{ij}$ and $x_i$ on $y_i$ yields the least squares estimators of $\beta$ as $b_5 = S_{yy}/S_{xy}$ and $b_6 = B_{yy}/B_{xy}$, respectively, where $S_{yy} = \sum_{i=1}^{n} \sum_{j=1}^{r} (y_{ij} - \bar{y})^2 / nr$ and $B_{yy} = \sum_{i=1}^{n} (\bar{y}_i - \bar{y})^2 / n$. Both $b_5$ and $b_6$ are inconsistent for $\beta$. When they are adjusted for their inconsistencies, the resulting consistent estimators of $\beta$ corresponding to $b_5$ and $b_6$ are

$$b_7 = \frac{S_{yy} - \hat{\sigma}_u^2}{S_{xy}}; \quad S_{yy} > \hat{\sigma}_u^2$$

and

$$b_8 = \frac{B_{yy} - (\hat{\sigma}_u^2 / r)}{B_{xy}}; \quad B_{yy} - \frac{\hat{\sigma}_u^2}{r} > 0,$$

respectively, where $\hat{\sigma}_u^2 = r(S_{yy} - B_{yy})/(r - 1)$.

The consistent estimators of $\alpha$ corresponding to $b_7$ and $b_8$ are given by

$$a_7 = \bar{y} - b_7 S_{xy}$$

and

$$a_8 = \bar{y} - b_8 \bar{x},$$

respectively.

Note that the estimators $a_3, b_3, a_4, b_4, a_7, b_7, a_8$ and $b_8$ can also be obtained by the method of moments.

### 3. Asymptotic properties of the estimators

The exact efficiency properties of the estimators described in Section 2 are difficult to derive, see e.g. Cheng and Van Ness [3] for more details. We therefore employ the large sample asymptotic approximation theory to study the efficiency properties of these estimators assuming that the errors are not necessarily normally distributed.

In order to study the asymptotic properties of the estimators, we assume that $S_{nm} = \sum_{i=1}^{n} (m_i - \bar{m})^2 / n$ tends to a finite quantity as $n \to \infty$. This assumption is required to avoid the presence of any trend in the data, see e.g. Schneeweiss [21]. Further, we assume that $n$ grows large but $r$ remains fixed.

Let

$$\theta = \frac{\sigma^2_v}{S_{nm} + \sigma^2_w + \sigma^2_v}; \quad 0 \leq \theta \leq 1,$$

and

$$q = \frac{\sigma_u^2}{\beta^2 \sigma_v^2}; \quad q \leq 0.$$

When there is no measurement error in the explanatory variable, then $\sigma^2_v = 0$. This gives that $\theta = 0$ which implies that the ultrastructural model reduces to the classical regression model. Thus a non-zero value of $\theta$ serves as a measure of departure of ultrastructural model from the classical regression model.

The large sample asymptotic efficiency properties of the estimators are stated in the following theorems.
Theorem 1 The large sample asymptotic approximation of relative bias (RB) of consistent estimators of $\beta$ and the bias (B) of consistent estimators of $\alpha$ up to $O(n^{-1})$ are given by

$$\text{RB}(b_3) = E\left( \frac{b_3 - \beta}{\beta} \right) = \frac{\theta}{nr(1-\theta)} \left[ 3 + \frac{2\theta}{(r-1)(1-\theta)} \right]$$ (12)

$$\text{RB}(b_4) = E\left( \frac{b_4 - \beta}{\beta} \right) = \frac{\theta}{nr(1-\theta)} \left[ 3 + \frac{2\theta}{(r-1)(1-\theta)} \right]$$ (13)

$$\text{RB}(b_7) = E\left( \frac{b_7 - \beta}{\beta} \right) = \frac{\theta}{nr(1-\theta)} \left[ (1 - 2q) + \frac{q\theta}{(1-\theta)} \right]$$ (14)

$$\text{RB}(b_8) = E\left( \frac{b_8 - \beta}{\beta} \right) = \frac{\theta}{nr(1-\theta)} \left[ (1 - 2q) + \frac{q\theta}{r(1-\theta)} \right]$$ (15)

$$B(a_3) = E(a_3 - \alpha) = -\frac{\tilde{m}\theta\beta}{nr(1-\theta)} \left[ 3 + \frac{2\theta}{(r-1)(1-\theta)} \right]$$ (16)

$$B(a_4) = E(a_4 - \alpha) = -\frac{\tilde{m}\theta\beta}{nr(1-\theta)} \left[ 3 + \frac{2\theta}{(r-1)(1-\theta)} \right]$$ (17)

$$B(a_7) = E(a_7 - \alpha) = -\frac{\tilde{m}\theta\beta}{nr(1-\theta)} \left[ (1 - 2q) + \frac{q\theta}{(1-\theta)} \right]$$ (18)

$$B(a_8) = E(a_8 - \alpha) = -\frac{\tilde{m}\theta\beta}{nr(1-\theta)} \left[ (1 - 2q) + \frac{q\theta}{r(1-\theta)} \right] .$$ (19)

Proof of the Theorem 1 is outlined in the appendix.

Firstly we analyse the bias of the estimators of $\beta$ and $\alpha$. The relative bias of $b_3$ and $b_4$ are positive and equal whereas the magnitude of relative bias of $b_8$ is always smaller than that of $b_7$ up to the order of approximation. The estimators $b_8$ and $b_7$ are positively biased if $q \leq 1/2$. The magnitudes of bias of $a_3$ and $a_4$ are equal and their signs depend on the signs of $\beta$ and $\tilde{m}$. If $\tilde{m}$ and $\beta$ have same signs, then $a_3$ and $a_4$ are negatively biased whereas if $\tilde{m}$ and $\beta$ have opposite signs, then they are positively biased. When $\tilde{m}$ and $\beta$ have same signs, then $a_7$ and $a_8$ are negatively biased if $q \leq 1/2$ and positively biased if $q > 1/2$. If $\tilde{m}$ and $\beta$ have opposite signs, the nature of bias of $a_7$ and $a_8$ depends upon the values of $q$ and $\theta$. The estimator $a_8$ has smaller magnitude of bias than $a_7$ when $\tilde{m}$ and $\beta$ have same signs. If $\tilde{m}$ and $\beta$ have opposite signs, then $a_7$ has smaller magnitude of bias than $a_8$.

The differences in the relative bias of estimators arising from the utilization of $\hat{\sigma}_u^2$ and $\hat{\sigma}_v^2$ are

$$\text{RB}(b_7) - \text{RB}(b_3) = \frac{\theta}{nr(1-\theta)} \left[ \frac{q\theta}{(1-\theta)} - 2 \left( 1 + q + \frac{\theta}{(r-1)(1-\theta)} \right) \right]$$

and

$$\text{RB}(b_8) - \text{RB}(b_4) = \frac{\theta}{nr(1-\theta)} \left[ \frac{q\theta}{r(1-\theta)} - 2 \left( 1 + q + \frac{\theta}{(r-1)(1-\theta)} \right) \right] ,$$

respectively and whence

$$\text{RB}(b_3) < \text{RB}(b_7), \text{ when } \theta > \frac{2}{3}$$

and

$$\text{RB}(b_4) < \text{RB}(b_8), \text{ when } \theta > \frac{2r}{2r + 1} .$$
These conditions provide a guideline to use either \( \hat{\sigma}^2_u \) or \( \hat{\sigma}^2_v \) for a small sample bias efficient estimator. Similar dominance conditions can also be derived for the estimators of \( \alpha \).

**Theorem 2** The large sample asymptotic approximation of the relative mean-squared error (MSE) (RM) of consistent estimators of \( \beta \) and the MSE (M) of consistent estimators of \( \alpha \) up to the order of approximation \( O(n^{-1}) \) are given by

\[
RM(b_3) = E \left( \frac{b_3 - \beta}{\beta} \right)^2 = \frac{\theta^2}{nr(1-\theta)^2} \left[ q + \frac{2}{(r-1)} + \frac{(1-\theta)(1+q)}{\theta} \right] \tag{20}
\]

\[
RM(b_4) = E \left( \frac{b_4 - \beta}{\beta} \right)^2 = \frac{\theta^2}{nr(1-\theta)^2} \left[ q + \frac{2}{(r-1)} + \frac{(1-\theta)(1+q)}{\theta} \right] \tag{21}
\]

\[
RM(b_7) = E \left( \frac{b_7 - \beta}{\beta} \right)^2 = \frac{\theta^2}{nr(1-\theta)^2} \left[ q + \frac{(1-\theta)(1+q)}{\theta} + \frac{2q^2}{(r-1)} \right] \tag{22}
\]

\[
RM(b_8) = E \left( \frac{b_8 - \beta}{\beta} \right)^2 = \frac{\theta^2}{nr(1-\theta)^2} \left[ q + \frac{(1-\theta)(1+q)}{\theta} + \frac{2q^2}{(r-1)} \right] \tag{23}
\]

\[
M(a_3) = E(a_3 - \alpha)^2 = \frac{\beta^2 \sigma_v^2}{nr} \left[ 1 + q + \frac{\nu^2}{\sigma_v^2(1-\theta)^2} \left( q + \frac{(1-\theta)(1+q)}{\theta} \right) + \frac{2}{(r-1)} \right] \tag{24}
\]

\[
M(a_4) = E(a_4 - \alpha)^2 = \frac{\beta^2 \sigma_v^2}{nr} \left[ 1 + q + \frac{\nu^2}{\sigma_v^2(1-\theta)^2} \left( q + \frac{(1-\theta)(1+q)}{\theta} \right) + \frac{2}{(r-1)} \right] \tag{25}
\]

\[
M(a_7) = E(a_7 - \alpha)^2 = \frac{\beta^2 \sigma_v^2}{nr} \left[ 1 + q + \frac{\nu^2}{\sigma_v^2(1-\theta)^2} \left( q + \frac{(1-\theta)(1+q)}{\theta} \right) + \frac{2q^2}{(r-1)} \right] \tag{26}
\]

\[
M(a_8) = E(a_8 - \alpha)^2 = \frac{\beta^2 \sigma_v^2}{nr} \left[ 1 + q + \frac{\nu^2}{\sigma_v^2(1-\theta)^2} \left( q + \frac{(1-\theta)(1+q)}{\theta} \right) + \frac{2q^2}{(r-1)} \right] \tag{27}
\]

The proof of Theorem 2 is outlined in appendix.

Now we analyse the MSE of the estimators of \( \beta \) and \( \alpha \). The relative mean squared errors of \( b_4 \) and \( b_8 \) are always smaller than that of \( b_3 \) and \( b_7 \), respectively. Similarly, the mean squared errors of \( a_4 \) and \( a_8 \) are always smaller than \( a_3 \) and \( a_7 \), respectively.

The difference in the relative MSE is

\[
RM(b_3) - RM(b_7) = RM(b_4) - RM(b_8) = \frac{2\theta^2(1-q^2)}{nr(r-1)(1-\theta)^2}.
\]

Thus we conclude that under the MSE criterion, \( b_3 \) and \( b_4 \) are superior to \( b_7 \) and \( b_8 \), respectively if \( 0 \leq q < 1 \). When \( q = 1 \), they are equally efficient. For \( q > 1 \), \( b_7 \) and \( b_8 \) are superior to \( b_3 \) and \( b_4 \) respectively.

For illustration, consider a set of values with \( S_{mm} = 0.23, \sigma_w^2 = 0.1, \beta = 0.3, n = 43 \) and \( r = 4 \). The asymptotic relative efficiencies of \( b_3 \) over \( b_7 \) and \( b_4 \) over \( b_8 \) are computed and presented in Table 1 when \( 0 \leq q < 1 \). Similarly, the asymptotic relative gain in the efficiencies of \( b_7 \) over \( b_3 \) and \( b_8 \) over \( b_4 \) are presented in Table 2 when \( q > 1 \).
Table 1. Relative gain in efficiencies of $b_3$ over $b_7$ and of $b_4$ over $b_8$ when $0 \leq q < 1$.

<table>
<thead>
<tr>
<th>Relative efficiency of $b_3$ over $b_7$</th>
<th>Relative efficiency of $b_4$ over $b_8$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{\sigma}_u^2$</td>
<td>$\hat{\sigma}_u^2$</td>
</tr>
<tr>
<td>----------------------------------------</td>
<td>----------------------------------------</td>
</tr>
<tr>
<td>$0.01$ 29.27 63.98 80.18</td>
<td>$0.01$ 30.18 65.03 80.91</td>
</tr>
<tr>
<td>$0.09$ 5.77 49.81 71.60</td>
<td>$0.09$ 6.82 56.50 77.24</td>
</tr>
<tr>
<td>$0.25$ – 25.78 56.40</td>
<td>$0.25$ – 33.87 67.89</td>
</tr>
<tr>
<td>$0.49$ – – 36.58</td>
<td>$0.49$ – – 49.80</td>
</tr>
</tbody>
</table>

Note: The cells with hyphen (–) in the table show that the condition $0 \leq q < 1$ is not fulfilled.

Table 2. Relative gain in efficiency of $b_7$ over $b_3$ and of $b_8$ over $b_4$ when $q > 1$.

<table>
<thead>
<tr>
<th>Relative efficiency of $b_7$ over $b_3$</th>
<th>Relative efficiency of $b_8$ over $b_4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{\sigma}_u^2$</td>
<td>$\hat{\sigma}_u^2$</td>
</tr>
<tr>
<td>----------------------------------------</td>
<td>----------------------------------------</td>
</tr>
<tr>
<td>$0.01$ – – –</td>
<td>$0.01$ – – –</td>
</tr>
<tr>
<td>$0.09$ – – –</td>
<td>$0.09$ – – –</td>
</tr>
<tr>
<td>$0.25$ 26.80 – –</td>
<td>$0.25$ 26.80 – –</td>
</tr>
<tr>
<td>$0.49$ 49.53 5.45 –</td>
<td>$0.49$ 49.53 5.45 –</td>
</tr>
</tbody>
</table>

Note: The cells with hyphen (–) in the table show that the condition $q > 1$ is not fulfilled.

The asymptotic relative gain in efficiency of $b_3$ over $b_7$ and $b_4$ over $b_8$ increases as $\hat{\sigma}_u^2$ increases for a given $\hat{\sigma}_v^2$ (see Table 1). On the other hand, the asymptotic relative gain in efficiency of $b_3$ over $b_7$ and $b_4$ over $b_8$ decreases as $\hat{\sigma}_v^2$ increases for fixed $\hat{\sigma}_u^2$. It is clear from Table 2 that the asymptotic relative gain in efficiency of $b_7$ over $b_3$ and of $b_8$ over $b_4$ increases as $\hat{\sigma}_v^2$ increases for a given $\hat{\sigma}_u^2$. The reverse hold true when $\hat{\sigma}_u^2$ is fixed, i.e., the asymptotic relative gain in efficiency of $b_7$ over $b_3$ and $b_8$ over $b_4$ decreases as $\hat{\sigma}_u^2$ increases.

We do not observe any influence of non-normality of distribution of error terms on the bias and MSE of the estimators up to order $O(n^{-1})$. This effect may precipitate if the higher order approximations for bias (or relative bias) and MSE (or relative MSE) are considered.

It follows from the application of central limit theorem that the consistent estimators of $\alpha$ and $\beta$ are asymptotically normally distributed. More precisely, the asymptotic distribution of $\sqrt{n}(a_l - \alpha)$, $(l = 3, 4, 7, 8)$ is normal with mean $0$ and asymptotic variance $M(a_l)$ as stated in Theorem 2. Similarly, the asymptotic distribution of $\sqrt{n}(b_l - \beta)$, $(l = 3, 4, 7, 8)$ is normal with mean $0$ and asymptotic variance $\beta^2RM(b_l)$ as stated in Theorem 2.

4. Monte Carlo simulation

An interesting question arises at the design stage of an experiment is that how many replications are sufficient to construct a good consistent estimator? Such information will help the experimenter in striking a balance between the cost of experiment and the efficiency of estimators. We have tried to address this issue based on the simulation results from a Monte Carlo experiment. Without loss of generality and for illustration, the measurement errors and random errors are assumed to follow a normal distribution. The results based on 5000 iterations are summarized in Table 3. A curve between the number of replication ($r$) and MSE of different consistent estimators is presented in Figure 1.
Table 3. MSE of $b_3, b_4, b_7$ and $b_8$ with respect to $r$ under normal distribution with $\sigma_u = 0.3, \sigma_v = 0.7, \sigma_w = 0.1$ and $n = 43$.

<table>
<thead>
<tr>
<th>Number of replication ($r$)</th>
<th>MSE of estimators</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$b_3$</td>
</tr>
<tr>
<td>2</td>
<td>12.0866</td>
</tr>
<tr>
<td>3</td>
<td>1.3832</td>
</tr>
<tr>
<td>4</td>
<td>0.2934</td>
</tr>
<tr>
<td>6</td>
<td>0.0161</td>
</tr>
<tr>
<td>9</td>
<td>0.0096</td>
</tr>
<tr>
<td>13</td>
<td>0.0065</td>
</tr>
<tr>
<td>18</td>
<td>0.0041</td>
</tr>
</tbody>
</table>

It is clear from Table 3 and Figure 1 that as $r$ increases, the MSE of all consistent estimators decreases. When $r > 5$, then all the estimators have small and gradually diminishing variability. Thus it can be suggested that the minimum number of replications needed in this case is 6.

The large sample asymptotic approximation theory is employed to derive the bias and MSE of the estimators. The large sample asymptotic theory provides an idea about the behaviour of the distribution of estimators in large samples. So the derived results do not give any idea about the distribution of the estimators over the whole sample space and in the finite samples.

A Monte Carlo simulation experiment is conducted to study the behaviour of estimators in finite samples. Following distributions of measurement errors are considered to study the non-normality effect on the efficiency properties of estimators.

1. Normal distribution (having zero skewness and zero kurtosis).
2. Central $t$-distribution (having zero skewness and non-zero kurtosis).
3. Standard Weibull distribution (having non-zero skewness and non-zero kurtosis) with density

$$f(x) = \left(\frac{cx^{c-1}}{b^c}\right) \exp\left[-\left(\frac{x}{b}\right)^c\right], \quad 0 \leq x < \infty, \quad b > 0, \quad c > 0$$

where scale parameter $b = 1$ and $c$ is the shape parameter.
4. Gamma distribution (having non-zero skewness and non-zero kurtosis) with density

\[ f(x) = \frac{e^{-x}x^{\lambda-1}}{\Gamma(\lambda)}; \quad \lambda > 0, \quad 0 < x < \infty. \]

5. Exponential distribution (having non-zero skewness and non-zero kurtosis) with density

\[ f(x) = \theta e^{-\theta x}; \quad x \geq 0, \quad \theta > 0. \]

The idea behind using different distributions of measurement errors is as follows. While generating the random variables, we scaled the random variables from different distributions to have the same mean and same variance. The distributions under consideration now differ only in the sense of skewness and kurtosis. For instance, when we compare the simulation results between normal and \( t \)-distribution, the difference in the values may be contributed as arising mainly due to the peakedness of the distribution. Similarly, comparison between the simulated values from the normal and Weibull, gamma or exponential distributions will give an idea about the joint effect of skewness and kurtosis. If there is no effect of departure from normality, then the results under \( t \), Weibull, exponential or gamma distributions are expected to be almost same as the results under normal distribution.

A set of values for \( m_i \)'s is generated for which \( S_{nn} = 0.23 \). Since the main interest lies in studying the non-normal effect of measurement errors, so the distribution of \( w_i \) is fixed to be normal with mean 0 and standard deviation \( \sigma_w = 0.01 \) in all experimental settings. Further, following choices are made for the different parameters: \( \sigma_u = 0.1, 0.2, 0.3; \sigma_v = 0.1, 0.3, 0.5, 0.7; \alpha = 1; \beta = 0.3; r = 4 \) and \( n = 15, 43 \). The sample sizes \( n = 15 \) and \( n = 43 \) are treated as small and large samples, respectively. The simulation experiment is carried out using the S-Plus statistical software for different combinations of parameters.

The experiment is repeated 10,000 times for each experimental setting. The bias and mean squared error of the estimators are computed empirically subject to the constraints, viz., \( S_{xx} > \sigma_u^2, B_{xx} - (\sigma_u^2/r) > 0, S_{yy} > \sigma_v^2, B_{yy} - (\sigma_v^2/r) > 0 \) and \( S_{xy} \neq 0 \) in respective estimators. The densities of distribution of all the estimators are plotted for various experimental settings. Keeping in view the length of paper, only few results in Tables 1–11 and Figures 1–9 are presented here.

First we analyze the empirical bias (EB) of different estimators. The estimators \( b_3 \) and \( b_4 \) are observed to be positively biased whereas \( a_3 \) and \( a_4 \) are negatively biased. This nature is same as that found from the asymptotic results. The magnitudes of EBs of all the estimators are smaller in large sample than in small sample.

<table>
<thead>
<tr>
<th></th>
<th>( n = 15 )</th>
<th>( n = 43 )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( a_3 )</td>
<td>( a_4 )</td>
</tr>
<tr>
<td>Normal</td>
<td>EB</td>
<td>-0.0017</td>
</tr>
<tr>
<td></td>
<td>EMSE</td>
<td>0.00270</td>
</tr>
<tr>
<td>( t )</td>
<td>EB</td>
<td>-0.0019</td>
</tr>
<tr>
<td></td>
<td>EMSE</td>
<td>0.0027</td>
</tr>
<tr>
<td>Weibull</td>
<td>EB</td>
<td>-0.0029</td>
</tr>
<tr>
<td></td>
<td>EMSE</td>
<td>0.0028</td>
</tr>
<tr>
<td>Gamma</td>
<td>EB</td>
<td>-0.0025</td>
</tr>
<tr>
<td></td>
<td>EMSE</td>
<td>0.0031</td>
</tr>
<tr>
<td>Exponential</td>
<td>EB</td>
<td>-0.0006</td>
</tr>
<tr>
<td></td>
<td>EMSE</td>
<td>0.0027</td>
</tr>
</tbody>
</table>
When the dispersion of measurement errors is small, then the EB of all the estimators of $\beta$ and $\alpha$ are small (see e.g. Tables 4 and 8). As $\sigma_v$ is increased from 0.1 to 0.7, while keeping $\sigma_u$ and $\sigma_w$ fixed at 0.1, then the EB of $a_3$, $a_4$, $b_3$ and $b_4$ increases under all distributions of measurement errors. The EB of $b_7$ and $b_8$ changes from negative to positive direction whereas the EB of $a_7$ and $a_8$ changes from positive to negative direction on a real line (see e.g. Tables 8 and 9).
When \( \sigma_v \) and \( \sigma_w \) are fixed at 0.1 and \( \sigma_u \) is increased from 0.1 to 0.3, then a significant change in the magnitude of EB of \( a_7, a_8, b_7 \) and \( b_8 \) is observed but there is no such significant change in the magnitude of EB of \( a_3, a_4, b_3 \) and \( b_4 \).

Again, when \( \sigma_u \) and \( \sigma_v \) are increased from 0.1 to 0.3 and 0.1 to 0.7, respectively, then the EB of \( a_3, a_4, b_3 \) and \( b_4 \) increases in the same direction (see e.g. Tables 6 and 7). In this case, the
magnitude of EB of $a_7$, $a_8$, $b_7$ and $b_8$ also increases (see e.g. Tables 10 and 11). Thus the influence of $\sigma_u$ on the EB of $a_3$, $a_4$, $b_3$ and $b_4$ is higher in comparison to $\sigma_u$. Also, both $\sigma_u$ and $\sigma_v$ affect the EB of $a_7$, $a_8$, $b_7$ and $b_8$ significantly.

For small standard deviations of measurement errors (see Tables 4 and 8), the empirical mean squared error (EMSE) of all the estimators are quite small. As $\sigma_u$ increases from 0.1 to 0.7, while keeping $\sigma_u$ and $\sigma_w$ fixed at 0.1, the EMSE of all the estimators increases. The rate of increment in the EMSE of $a_3$, $a_4$, $b_3$ and $b_4$ is higher than in $a_7$, $a_8$, $b_7$ and $b_8$. There is a significant difference in the EMSE of estimators when measurement errors arise from different distributions. This clearly indicates that the departure from normality does affects the efficiency properties of the estimators.

When $\sigma_u$ increases from 0.1 to 0.3 while $\sigma_u$ and $\sigma_w$ are fixed at 0.1, then no significant change is observed in the EMSE of $a_3$, $a_4$, $b_3$ and $b_4$, but a significant change is observed in the EMSE of $a_7$, $a_8$, $b_7$ and $b_8$. This effect is much higher in small sample than in large sample.

Again, when $\sigma_u$ and $\sigma_v$ increase from 0.1 to 0.3 and 0.1 to 0.7, respectively while keeping $\sigma_w$ fixed at 0.1, then the EMSE of all the estimators increases (see, e.g. Tables 6, 7, 10 and 11). The estimators have higher EMSE in small sample than in large sample. Also, a significant difference is observed in the EMSEs under different distributions of measurement errors. The non-normality effect of measurement errors towards the efficiency properties of the estimators is higher in small sample than in large sample.

When EMSE is compared with asymptotic MSE of the estimator, it is observed that they are quite close in large sample cases with small measurement errors. When $\sigma_u$ increases, then the EMSE of estimators increases in small sample. Hence the empirical results and asymptotic results differ significantly for higher value of $\sigma_u$ and $\sigma_v$ in small sample. Keeping in view the length of paper, they are not reported here.

The density of the estimators are plotted and compared with corresponding least square estimators for various experimental setting. Keeping in view the length of paper, all results are not reported but only few are presented in Figures 2–9. This will give an idea about the nature of inconsistency and also an idea to the experimenter about the deviations in the results if least squares estimators is used, ignoring the measurement errors.

First we consider the case of small standard deviations of measurement errors ($\sigma_u = \sigma_v = \sigma_w = 0.1$). In this case, the densities of $b_1$, $b_2$, $b_3$ and $b_4$ are similar. Some dis-similarities in the density graphs of these estimators appear in large sample under gamma distributed measurement errors (see, Figure 2). Among the density plots of $b_5$, $b_6$, $b_7$ and $b_8$, the density plots of $b_7$ and $b_8$ are almost same. The peakedness is higher in large sample than in small sample. Comparing the peakedness in the density curves, we observe that $b_5$ and $b_6$ have lowest and highest peakedness, respectively, when the measurement errors follow normal, $t$, Weibull and exponential distributions.

Table 11. EB and EMSE of $a_7$, $a_8$, $b_7$ and $b_8$ with $\sigma_u = 0.3$, $\sigma_v = 0.5$, $\sigma_w = 0.1$.

<table>
<thead>
<tr>
<th></th>
<th>$n = 15$</th>
<th>$n = 43$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$a_7$</td>
<td>$a_8$</td>
</tr>
<tr>
<td>Normal EB</td>
<td>0.0858</td>
<td>−0.0211</td>
</tr>
<tr>
<td>EMSE</td>
<td>146.228</td>
<td>55.9765</td>
</tr>
<tr>
<td>$t$ EB</td>
<td>0.1864</td>
<td>0.0143</td>
</tr>
<tr>
<td>EMSE</td>
<td>161.8690</td>
<td>28.6893</td>
</tr>
<tr>
<td>Weibull EB</td>
<td>−0.0675</td>
<td>−0.0169</td>
</tr>
<tr>
<td>EMSE</td>
<td>260.876</td>
<td>13.5878</td>
</tr>
<tr>
<td>Gamma EB</td>
<td>0.0013</td>
<td>0.0283</td>
</tr>
<tr>
<td>EMSE</td>
<td>11.3684</td>
<td>6.4170</td>
</tr>
<tr>
<td>Exponential EB</td>
<td>0.02560</td>
<td>−0.0072</td>
</tr>
<tr>
<td>EMSE</td>
<td>25.6922</td>
<td>15.5339</td>
</tr>
<tr>
<td>Distribution</td>
<td>n = 15</td>
<td>n = 43</td>
</tr>
<tr>
<td>--------------</td>
<td>--------</td>
<td>--------</td>
</tr>
<tr>
<td>Normal</td>
<td><img src="image1" alt="Graph" /></td>
<td><img src="image2" alt="Graph" /></td>
</tr>
<tr>
<td>$t$</td>
<td><img src="image3" alt="Graph" /></td>
<td><img src="image4" alt="Graph" /></td>
</tr>
<tr>
<td>Weibull</td>
<td><img src="image5" alt="Graph" /></td>
<td><img src="image6" alt="Graph" /></td>
</tr>
<tr>
<td>Gamma</td>
<td><img src="image7" alt="Graph" /></td>
<td><img src="image8" alt="Graph" /></td>
</tr>
<tr>
<td>Exponential</td>
<td><img src="image9" alt="Graph" /></td>
<td><img src="image10" alt="Graph" /></td>
</tr>
</tbody>
</table>

Figure 2. Density plots of $b_1$, $b_2$, $b_3$ and $b_4$ with $\sigma_u = 0.1, \sigma_v = 0.1, \sigma_w = 0.1$. 
<table>
<thead>
<tr>
<th>Distribution</th>
<th>$n = 15$</th>
<th>$n = 43$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Normal</strong></td>
<td><img src="image" alt="Normal distribution" /></td>
<td><img src="image" alt="Normal distribution" /></td>
</tr>
<tr>
<td><strong>t</strong></td>
<td><img src="image" alt="t distribution" /></td>
<td><img src="image" alt="t distribution" /></td>
</tr>
<tr>
<td><strong>Weibull</strong></td>
<td><img src="image" alt="Weibull distribution" /></td>
<td><img src="image" alt="Weibull distribution" /></td>
</tr>
<tr>
<td><strong>Gamma</strong></td>
<td><img src="image" alt="Gamma distribution" /></td>
<td><img src="image" alt="Gamma distribution" /></td>
</tr>
<tr>
<td><strong>Exponential</strong></td>
<td><img src="image" alt="Exponential distribution" /></td>
<td><img src="image" alt="Exponential distribution" /></td>
</tr>
</tbody>
</table>

Figure 3. Density plots of $b_1$, $b_2$, $b_3$ and $b_4$ with $\sigma_u = 0.1$, $\sigma_v = 0.5$, $\sigma_w = 0.1$. 
Figure 4. Density plots of $b_1$, $b_2$, $b_3$ and $b_4$ with $\sigma_u = 0.2$, $\sigma_v = 0.5$, $\sigma_w = 0.1$. 

<table>
<thead>
<tr>
<th>Distribution</th>
<th>$n = 15$</th>
<th>$n = 43$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td>![Density Plot]</td>
<td>![Density Plot]</td>
</tr>
<tr>
<td>$t$</td>
<td>![Density Plot]</td>
<td>![Density Plot]</td>
</tr>
<tr>
<td>Weibull</td>
<td>![Density Plot]</td>
<td>![Density Plot]</td>
</tr>
<tr>
<td>Gamma</td>
<td>![Density Plot]</td>
<td>![Density Plot]</td>
</tr>
<tr>
<td>Exponential</td>
<td>![Density Plot]</td>
<td>![Density Plot]</td>
</tr>
<tr>
<td>Distribution</td>
<td>$n = 15$</td>
<td>$n = 43$</td>
</tr>
<tr>
<td>--------------</td>
<td>----------</td>
<td>----------</td>
</tr>
<tr>
<td>Normal</td>
<td><img src="image1" alt="Normal Density Plot" /></td>
<td><img src="image2" alt="Normal Density Plot" /></td>
</tr>
<tr>
<td>t</td>
<td><img src="image3" alt="t Density Plot" /></td>
<td><img src="image4" alt="t Density Plot" /></td>
</tr>
<tr>
<td>Weibull</td>
<td><img src="image5" alt="Weibull Density Plot" /></td>
<td><img src="image6" alt="Weibull Density Plot" /></td>
</tr>
<tr>
<td>Gamma</td>
<td><img src="image7" alt="Gamma Density Plot" /></td>
<td><img src="image8" alt="Gamma Density Plot" /></td>
</tr>
<tr>
<td>Exponential</td>
<td><img src="image9" alt="Exponential Density Plot" /></td>
<td><img src="image10" alt="Exponential Density Plot" /></td>
</tr>
</tbody>
</table>

Figure 5. Density plots of $b_1$, $b_2$, $b_3$ and $b_4$ with $\sigma_u = 0.3$, $\sigma_v = 0.5$, $\sigma_w = 0.1$. 
<table>
<thead>
<tr>
<th>Distribution</th>
<th>$n = 15$</th>
<th>$n = 43$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td><img src="image1" alt="Density plots" /></td>
<td><img src="image2" alt="Density plots" /></td>
</tr>
<tr>
<td>$t$</td>
<td><img src="image3" alt="Density plots" /></td>
<td><img src="image4" alt="Density plots" /></td>
</tr>
<tr>
<td>Weibull</td>
<td><img src="image5" alt="Density plots" /></td>
<td><img src="image6" alt="Density plots" /></td>
</tr>
<tr>
<td>Gamma</td>
<td><img src="image7" alt="Density plots" /></td>
<td><img src="image8" alt="Density plots" /></td>
</tr>
<tr>
<td>Exponential</td>
<td><img src="image9" alt="Density plots" /></td>
<td><img src="image10" alt="Density plots" /></td>
</tr>
</tbody>
</table>

Figure 6. Density plots of $b_5$, $b_6$, $b_7$ and $b_8$ with $\sigma_w = 0.1$, $\sigma_v = 0.1$, $\sigma_u = 0.1$. 
<table>
<thead>
<tr>
<th>Distribution</th>
<th>$n = 15$</th>
<th>$n = 43$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td><img src="image1" alt="Normal Distribution" /></td>
<td><img src="image2" alt="Normal Distribution" /></td>
</tr>
<tr>
<td>$t$</td>
<td><img src="image3" alt="t Distribution" /></td>
<td><img src="image4" alt="t Distribution" /></td>
</tr>
<tr>
<td>Weibull</td>
<td><img src="image5" alt="Weibull Distribution" /></td>
<td><img src="image6" alt="Weibull Distribution" /></td>
</tr>
<tr>
<td>Gamma</td>
<td><img src="image7" alt="Gamma Distribution" /></td>
<td><img src="image8" alt="Gamma Distribution" /></td>
</tr>
<tr>
<td>Exponential</td>
<td><img src="image9" alt="Exponential Distribution" /></td>
<td><img src="image10" alt="Exponential Distribution" /></td>
</tr>
</tbody>
</table>

Figure 7. Density plots of $b_5$, $b_6$, $b_7$ and $b_8$ with $\sigma_w = 0.1$, $\sigma_v = 0.5$, $\sigma_u = 0.1$. 
Figure 8. Density plots of \( b_5, b_6, b_7 \) and \( b_8 \) with \( \sigma_w = 0.1, \sigma_v = 0.5, \sigma_u = 0.2 \).
Figure 9. Density plots of $b_5$, $b_6$, $b_7$ and $b_8$ with $\sigma_w = 0.1$, $\sigma_v = 0.5$, $\sigma_u = 0.3$. 
Under gamma distributed measurement errors, $b_5$ and $b_6$ have much flatter density curves than $b_7$ and $b_8$. The departure from the central value of density curves from $\beta (=0.3)$ is much higher in $b_5$ than in other estimators (see Figure 6).

When $\sigma_v$ is increased from 0.1 to 0.7, while keeping $\sigma_u$ and $\sigma_w$ fixed at 0.1, then the density curves of $b_3$ and $b_4$ are more flat than that of $b_1$ and $b_2$ (see e.g. Figure 3). In case of $b_5$, $b_6$, $b_7$ and $b_8$, the peakedness of the density curves decreases (see e.g. Figure 7). The density curves of $b_5$ and $b_6$ are more flat than the density curves of $b_7$ and $b_8$. There is a significant difference in the peakedness of the curves depending on the distribution of measurement errors and sample sizes.

No significant changes in the density plots of $b_1$, $b_2$, $b_3$ and $b_4$ is observed when $\sigma_u$ increases while keeping $\sigma_v$ and $\sigma_w$ fixed at 0.1. Some changes in the peakedness of density curves of $b_5$, $b_6$, $b_7$ and $b_8$ are observed in small sample depending on the distribution of measurement errors. But in large sample case, such significant change is observed only in the case of gamma distributed measurement errors.

When $\sigma_u$ increases from 0.1 to 0.3 and $\sigma_v$ from 0.1 to 0.7 while keeping $\sigma_w$ fixed at 0.1, then the peakedness of density curves of $b_3$ and $b_4$ decreases significantly in both the small and large samples (see e.g. Figures 4 and 5). The density curves of $b_5$, $b_6$, $b_7$ and $b_8$ become more flat in these cases. Such change are more significant in small sample. The estimators $b_7$ and $b_8$ have more peaked density curves than the density curves of $b_5$ and $b_6$ (see, e.g. Figures 8 and 9). Thus $b_3$ and $b_4$ are less biased than $b_1$ and $b_2$. When standard deviations of measurement errors become high, then $b_3$ and $b_4$ acquire higher variabiliy than $b_1$ and $b_2$. On the other hand, the estimators $b_7$ and $b_8$ are superior to $b_5$ and $b_6$ in every aspect under consideration. Also the efficiency properties of the estimators under direct and reverse regression approaches are affected by the skewness and kurtosis of the measurement errors distributions.

References

A. Appendix

Let

\[ X = \text{col.}(X_1, X_2, \ldots, X_n), \]
\[ Y = \text{col.}(Y_1, Y_2, \ldots, Y_n), \]
\[ u = \text{col.}(u_1, \ldots, u_r, u_{2r}, \ldots, u_{nr}), \]
\[ v = \text{col.}(v_1, \ldots, v_{1r}, v_{2r}, \ldots, v_{1nr}), \]
\[ w = \text{col.}(w_1, w_2, \ldots, w_n), \]
\[ m = \text{col.}(m_1, m_2, \ldots, m_n), \]
\[ e_r = \text{col.}(1, 1, \ldots, 1), \quad (r \times 1) \text{ column vector with each element being unity}, \]
\[ e_n = \text{col.}(1, 1, \ldots, 1), \quad (n \times 1) \text{ column vector with each element being unity}, \]
\[ e_{nr} = \text{col.}(1, 1, \ldots, 1), \quad (nr \times 1) \text{ column vector with each element being unity}, \]
\[ x = \text{col.}(x_1, x_{1r}, x_{2r}, \ldots, x_{nr}), \]
\[ y = \text{col.}(y_1, y_{1r}, y_{2r}, \ldots, y_{nr}), \]
\[ A = I_n - \frac{1}{nr} e_{nr} e_{nr}^T, \quad B = \frac{1}{r} \left( I_r \otimes e_r^T - \frac{1}{n} e_{nr} e_{nr}^T \right), \]
\[ C = I_n - \frac{1}{n} e_{nr} e_{nr}^T, \quad \text{and} \quad D = \frac{1}{r} \left( I_r \otimes e_r e_r^T - \frac{1}{nr} e_{nr} e_{nr}^T \right), \]

where \( X = \text{col.}(Z_1, Z_2, \ldots, Z_n) \), denotes a \( n \times 1 \) column vector and \( \otimes \) denotes the Kronecker product operator of matrices.

Note that

\[ BB^T = \frac{1}{r} C, \quad B' B = \frac{1}{r} D, \quad AD = D \quad \text{and} \quad CB = B. \]

Here, the matrices \( A, C \) and \( D \) are idempotent with \( \text{tr} A = (nr - 1) \), \( \text{tr} C = (n - 1) \) and \( \text{tr} D = (n - 1) \).

Further, we define the following quantities and each is of order of \( O(n) \):

\[ Q_u = \frac{e_u^T u}{\sqrt{nr}} = \sqrt{n} \tilde{u}, \quad Q_v = \frac{e_v^T v}{\sqrt{nr}} = \sqrt{n} \tilde{v} \quad \text{and} \quad Q_w = \frac{e_w^T w}{\sqrt{n}} = \sqrt{n} \tilde{w}. \]

Now we express,

\[ S_{xy} = \frac{1}{nr} \sum_{i=1}^{n} \sum_{j=1}^{r} (x_{ij} - \bar{x})(y_{ij} - \bar{y}) \]
\[ = \beta S_{mm} + \frac{\beta}{n} w'Cw + \frac{2\beta}{n} m'Cw + \frac{1}{n} (m + w)'Bu + \frac{\beta}{r} (m + w)'Bv + \frac{1}{nr} v'Au \]
\[ = \beta \sigma_v^2 \left[ \frac{(1 - \theta)}{\theta} + \frac{1}{\sqrt{r}} (g_{xy} + t_{xy}) \right], \]

where

\[ g_{xy} = \frac{1}{\sqrt{n} \sigma_v} \left[ \frac{1}{\sqrt{nr} \beta} (m + w)'Bu + (m + w)'Bv + 2m'Cw + (w'Cw - n \sigma_v^2) \right] \]

and

\[ t_{xy} = \frac{1}{\sqrt{nr} \beta \sigma_v} v'Au. \]
Similarly,

\[ S_{xx} = \sigma_v^2 \left[ \frac{1}{\theta} + \frac{1}{\sqrt{n}} (g_{xx} + t_{xx}) \right], \]

\[ S_{yy} = \beta^2 \sigma_v^2 \left[ \frac{(1 - \theta)}{\theta} + q + \frac{1}{\sqrt{n}} (g_{yy} + t_{yy}) \right], \]

\[ B_{xy} = \beta \sigma_v^2 \left[ \frac{(1 - \theta)}{\theta} + \frac{1}{\sqrt{n}} (g_{xy} + t_{xy}) \right], \]

\[ B_{xx} = \sigma_v^2 \left[ \frac{(1 - \theta)}{\theta} + \frac{q}{r} + \frac{1}{\sqrt{n}} (g_{xx} + t_{xx}) \right], \]

and \[ B_{yy} = \beta^2 \sigma_v^2 \left[ \frac{(1 - \theta)}{\theta} + \frac{q}{r} + \frac{1}{\sqrt{n}} (g_{yy} + t_{yy}) \right] \]

where \( g_{xx} = \frac{1}{n \sigma_v^2} \left[ 2m'Cw + 2(m + w)'Bv + (w'Cw - n \sigma_v^2) \right], \)

\( g_{yy} = \frac{1}{\sqrt{n \sigma_v^2}} \left[ 2 \beta (m + w)'Bu + 2m'Cw + (w'Cw - n \sigma_v^2) \right], \)

\( t_{xx} = \frac{1}{\sqrt{n \sigma_v^2}} (u'A \nu - n \sigma_v^2), \)

\( t_{yy} = \frac{1}{\sqrt{n \sigma_v^2}} (u'Av - n \sigma_v^2), \)

\( t_{xy}^* = \frac{1}{n \beta \sigma_v^2} u' Dv, \)

\( t_{yx}^* = \frac{1}{n \sigma_v^2} (v'Dv - n \sigma_v^2), \)

and \( t_{yy}^* = \frac{1}{n \beta \sigma_v^2} (u'Du - n \sigma_v^2). \)

Now, the estimation error of estimator of \( b_1 \) is given by

\[ b_1 - \beta = \frac{(r - 1) S_{xy}}{r B_{xx} - S_{xx}} - \beta \]

\[ = \frac{\theta \beta}{\sqrt{n(1 - \theta)}} \left[ g_{xy} + t_{xy} - g_{xx} - \frac{r}{r - 1} t_{xx}^* + \frac{1}{r - 1} t_{xx} \right] \]

\[ \times \left[ 1 + \frac{\theta}{\sqrt{n(1 - \theta)}} \left( g_{xx} + \frac{r}{r - 1} t_{xx}^* - \frac{1}{r - 1} t_{xx} \right) \right]^{-1} \]

\[ = \frac{1}{\sqrt{n}} B_{-1/2} + \frac{1}{n} B_{-1} + O_p(n^{-3/2}) \]

where \( B_{-1/2} = \frac{\theta \beta^2}{(1 - \theta)} \left[ g_{xy} + t_{xy} - g_{xx} - \frac{r}{r - 1} t_{xx}^* + \frac{1}{r - 1} t_{xx} \right] \)

and \( B_{-1} = -\frac{\beta \theta^2}{(1 - \theta)^2} \left[ g_{xx} + \frac{r}{r - 1} t_{xx}^* - \frac{1}{r - 1} t_{xx} \right] \]

\[ \times \left[ g_{xy} + t_{xy} - g_{xx} - \frac{r}{r - 1} t_{xx}^* + \frac{1}{r - 1} t_{xx} \right]. \]

Similarly, the estimation errors of \( b_2, b_3 \) and \( b_4 \) can be obtained as

\[ b_2 - \beta = \frac{1}{\sqrt{n}} H_{-1/2} + \frac{1}{n} H_{-1} + O_p(n^{-3/2}), \]

\[ b_3 - \beta = \frac{1}{\sqrt{n}} A_{-1/2} + \frac{1}{n} A_{-1} + O_p(n^{-3/2}), \]

\[ b_4 - \beta = \frac{1}{\sqrt{n}} D_{-1/2} + \frac{1}{n} D_{-1} + O_p(n^{-3/2}), \]
After evaluation of the expectations and solving, we get

\[ H_{-1/2} = \frac{\theta\beta}{(1-\theta)} \left[ g_{xy} + t_{xy}^* - g_{xx} \frac{r}{r-1} t_{xx}^* + \frac{1}{r-1} t_{xx} \right] \]

\[ H_{-1} = \frac{\theta^2\beta}{(1-\theta)^2} \left[ \left( g_{xx} + \frac{r}{r-1} t_{xx}^* - \frac{1}{r-1} t_{xx} \right) \times \left( g_{xy} + t_{xy}^* - g_{xx} - \frac{r}{r-1} t_{xx}^* + \frac{1}{r-1} t_{xx} \right) \right], \]

\[ A_{-1/2} = \frac{\theta\beta}{(1-\theta)} \left[ g_{yy} + \frac{r}{r-1} t_{yy}^* - \frac{1}{r-1} t_{yy} - g_{yy} - t_{yy} \right] \]

\[ A_{-1} = \frac{\theta^2\beta}{(1-\theta)^2} (g_{xy} + t_{xy}) \left[ g_{yy} + \frac{r}{r-1} t_{yy}^* - \frac{1}{r-1} t_{yy} - g_{yy} - t_{yy} \right], \]

\[ D_{-1/2} = \frac{\theta\beta}{(1-\theta)} \left[ g_{yy} + \frac{r}{r-1} t_{yy}^* - \frac{1}{r-1} t_{yy} - g_{yy} - t_{yy} \right] \]

\[ D_{-1} = \frac{\theta^2\beta}{(1-\theta)^2} (g_{xy} + t_{xy}) \left[ g_{yy} + \frac{r}{r-1} t_{yy}^* - \frac{1}{r-1} t_{yy} - g_{yy} - t_{yy} \right]. \]

Now consider

\[ \hat{\alpha} = \bar{y} - \hat{\beta}\bar{x} \]

\[ = \alpha - (\bar{u} - \beta\bar{v}) - (\hat{\beta} - \beta)(\bar{m} + \bar{v} + \bar{v}). \]

Therefore the estimation error of an estimator of intercept term is given as

\[ a_l - \alpha = -\bar{m}(b_l - \beta) + \frac{1}{\sqrt{n}} \left[ (Q_n - \beta Q_v) - (b_l - \beta)(Q_w + Q_v) \right], \quad l = 3, 4, 7, 8. \]

The different expectations of product of vectors and matrices, which are used to prove the results in theorems can be derived by using the distributional properties of \( u, v \) and \( w \) as well as following Srivastava and Tiwari [22].

Now, up to order \( O(n^{-1}) \),

\[ \text{Bias}(b_3) = \frac{1}{\sqrt{n}} E(B_{-1/2}) - \frac{1}{n} E(B_{-1}). \]

After evaluation of the expectations and solving, we get

\[ \text{RB}(b_3) = \frac{\theta}{nr(1-\theta)} + \frac{\theta^2}{nr(1-\theta)^2} \left[ \frac{2(1-\theta)}{\theta} + \frac{2}{r-1} \right] + O(n^{-3/2}) \]

\[ = \frac{\theta}{nr(1-\theta)} \left[ 3 + \frac{2\theta}{(r-1)(1-\theta)} \right] + O(n^{-3/2}). \]

Similarly,

\[ \text{MSE}(b_3) = \frac{1}{n} E(B_{-1/2}^2), \quad \text{up to } O(n^{-1}). \]

Evaluating the expectations and solving we get

\[ \text{RM}(b_3) = \frac{\theta^2}{n(1-\theta)} \left[ (1-\theta)(1+q) + \frac{q}{r} + \frac{1}{r(1-\theta)} \left( \frac{1}{r} \gamma_{2v} + 2 \right) + \frac{1}{r(r-1)^2} (\gamma_2 + 2) \right] \]

\[ - \frac{2}{r(r-1)^2} (\gamma_{2v} + 2) + O(n^{-2}) \]

\[ = \frac{\theta^2}{nr(1-\theta)^2} \left[ q + \frac{2}{r(1-\theta)} + \frac{(1+q)(1-\theta)}{\theta} \right] + O(n^{-2}). \]

Similarly, by using the estimation errors of respective estimators, other results given in the Theorems 1 and 2 can be derived on the similar lines.