Journal of Multivariate Analysis 102 (2011) 264–280

Contents lists available at ScienceDirect

Journal of Multivariate Analysis

journal homepage: www.elsevier.com/locate/jmva

# Restricted estimation in multivariate measurement error regression model

Kanchan Jain *, Sukhbir Singh, Suresh Sharma

*Department of Statistics, Panjab University, Chandigarh-160014, India*

ARTICLE INFO

*Article history:*
Received 25 February 2010
Available online 16 September 2010

*AMS subject classifications:*
62J05
62H12

*Keywords:*
Measurement error
Multivariate regression
Reliability matrix
Linear restrictions
Consistent estimators

ABSTRACT

We study a multivariate ultrastructural measurement error (MUME) model with more than one response variable. This model is a synthesis of multivariate functional and structural models. Three consistent estimators of regression coefficients, satisfying the exact linear restrictions have been proposed. Their asymptotic distributions are derived under the assumption of a non-normal measurement error and random error components. A simulation study is carried out to investigate the small sample properties of the estimators. The effect of departure from normality of the measurement errors on the estimators is assessed.



## 1. Introduction

In any statistical study, it is very important that the data are measured accurately. Though all efforts are made to ensure this, yet in some practical situations, the basic assumption of the measurability of observations is often violated. This happens due to various reasons. These reasons can be inaccurate responses of people or incomplete observability of some variables. For example, variables like air pollutant levels, counts of some hormones in the body and rainfall etc. cannot be truly observed.

Measurement error is the difference between the observed and the true value of a variable. When such an error creeps in and is significant, the statistical procedures designed for the "no measurement error" case do not lead to correct and valid decisions. For example, the least squares estimator (LSE) is the best linear unbiased estimator of the regression coefficient if there is no measurement error. The same LSE becomes biased and inconsistent in the presence of measurement error. Hence, the study of effect of measurement error gains importance.

When the predictors cannot be measured truly, some additional information is required for consistent estimation of regression coefficients. When there is only one predictor, this additional information could be the knowledge of the measurement error variance, ratio of error variances or reliability ratio etc. When there are $p$ predictors, this information is available in the form of a variance-covariance matrix of measurement errors, a reliability matrix or instrumental variables. More details can be found in [1,5–7,17,19].

In some cases, there is some prior information available about the regression coefficients, e.g. the availability of exact linear restrictions on regression coefficients. This information is obtained from past experience of the experimenter and/or long association with the study. Utilization of this information leads to an improvement in the LSE of regression coefficients. In the without measurement error case, a consistent LSE satisfying the exact linear restrictions is provided in [15]. Shalabh [20] considered the consistent estimation of regression coefficients under exact linear restrictions for the univariate ultrastructural measurement error model with more than one predictor.

* Corresponding author.
*E-mail addresses:* jaink14@gmail.com, jaink@pu.ac.in, jaink14@yahoo.com (K. Jain).

0047-259X/$ – see front matter 
doi:10.1016/j.jmva.2010.09.004

In this paper, we study the multivariate ultrastructural measurement error (MUME) multiple regression model. We propose some consistent estimators that satisfy the exact linear restrictions on regression parameters. This model is a synthesis of the multivariate functional and structural models (Ref. [3]). The multiple regression model is a special case of this model. Such models arise in real life as often more than one correlated variable is observed. For example, in medical sciences, generally more than one body characteristic of the subjects under study is recorded. These characteristics can be cholesterol level and blood pressure etc. and the interest is to relate these to the amount of different nutrients in the daily diet. Similarly, in botanical studies, the observed characteristics of a plant are leaf diameter, mass of root ball and diameter of bloom etc. We wish to regress these on the amount of several minerals present in the soil and the amount of light and water received by each plant. It is quite possible that the variables involved in the study may possess measurement errors.

The paper is structured in five sections. Section 2 specifies the MUME multiple regression model and lists additional statistical assumptions. In Section 3, we propose three consistent estimators of regression parameters satisfying the exact linear restrictions. These estimators are derived under the knowledge of reliability matrix. Section 4 includes the asymptotic distribution of the suggested estimators. In Section 5, Monte-Carlo simulations are carried out to study the finite sample properties of the estimators. The effect on the properties of the proposed estimators is assessed in case of departure from normality of the measurement errors. Appendix provides few definitions and the detailed derivations of theoretical results.

## 2. Model specifications

The multivariate regression model (without measurement error) is specified as

$$Z_{n\times q} = D_{n\times p}B_{p\times q} + E_{n\times q}, \tag{2.1}$$

where $Z$, $D$, $B$ and $E$ are matrices of dependent variables, independent predictors, regression coefficients and equation error terms respectively. For $i, k = 1, 2, \ldots, q$; $j = 1, 2, \ldots, n$, we write

$$\begin{aligned}
Z &= \left[Z_{(1)}, Z_{(2)}, \ldots, Z_{(q)}\right] \quad \text{where } Z_{(i)} = [z_{1i}, z_{2i}, \ldots, z_{ni}]';\\
D &= \left[D_{(1)}, D_{(2)}, \ldots, D_{(n)}\right]' \quad \text{with } D_{(j)} = \left[d_{j1}, d_{j2}, \ldots, d_{jp}\right]';\\
B &= \left[B_{(1)}, B_{(2)}, \ldots, B_{(q)}\right] \quad \text{with } B_{(k)} = \left[\beta_{1k}, \beta_{2k}, \ldots, \beta_{pk}\right]';\\
E &= \left[E_{(1)}, E_{(2)}, \ldots, E_{(q)}\right] \quad \text{with } E_{(i)} = [\epsilon_{1i}, \epsilon_{2i}, \ldots, \epsilon_{ni}]'.
\end{aligned}$$

Let $\Sigma_o$ be a matrix with $(i, k)$th element $\sigma_{\epsilon ik} = \text{cov}\left(\epsilon_{ji}, \epsilon_{jk}\right)$ and diagonal elements $\sigma^2_{\epsilon i} = \text{var}\left(\epsilon_{ji}\right)$ for $j = 1, 2, \ldots, n$ (Ref. [22]). Then $\Sigma_E = [I_n \otimes \Sigma_o]$ denotes the variance-covariance matrix of $E$. This indicates that $q$ observations on the $j$th trial have variance-covariance matrix $\Sigma_o$ but observations from different trials are uncorrelated. For observable matrices $Z$ and $D$, the system of equations given by (2.1) can be estimated consistently (Ref. [13]).

We assume that $Z$ is observable but $D$ is unobservable and can be observed only through $X$ with additional measurement error $\Delta$ as

$$X_{n\times p} = D_{n\times p} + \Delta_{n\times p}, \tag{2.2}$$

where $X = \left[X_{(1)}, X_{(2)}, \ldots, X_{(n)}\right]'$ with $X_{(j)} = \left[x_{j1}, x_{j2}, \ldots, x_{jp}\right]'$;

$$\Delta = \left[\Delta_{(1)}, \Delta_{(2)}, \ldots, \Delta_{(n)}\right]' \quad \text{with } \Delta_{(j)} = \left[\delta_{j1}, \delta_{j2}, \ldots, \delta_{jp}\right]' \quad \text{for } j = 1, 2, \ldots, n.$$

When $Z$ is also measured with error, then without loss of generality, this additional measurement error can be combined with matrix $E$.

Let

$$D_{n\times p} = M_{n\times p} + \Psi_{n\times p}, \tag{2.3}$$

where $M = \left[M_{(1)}, M_{(2)}, \ldots, M_{(n)}\right]'$ with $M_{(j)} = \left[M_{j1}, M_{j2}, \ldots, M_{jp}\right]'$ is a matrix of fixed components and $\Psi = \left[\Psi_{(1)}, \Psi_{(2)}, \ldots, \Psi_{(n)}\right]'$ with $\Psi_{(j)} = \left[\psi_{j1}, \psi_{j2}, \ldots, \psi_{jp}\right]'$ is a matrix of random components.

The structural form of the measurement error model arises when all rows of $M$ are identical. In this case, rows of $X$ will be independent and identically distributed with some multivariate distribution. For a null matrix $\Psi$, the matrix $X$ is fixed but measured with error. This case specifies a functional measurement error model. When both $\Psi$ and $\Delta$ are null matrices, we get the specifications of a classical regression model.

For known matrices $R_1$, $R_2$ and $\theta$, it is supposed that some prior information is available regarding $B$ in the form of linear restrictions expressed as

$$(R_1)_{r_1\times p}\, B_{p\times q}\, (R_2)_{q\times r_2} = \theta_{r_1\times r_2}. \tag{2.4}$$

$R_1$ and $R_2$ respectively impose exact linear restrictions on the parameters of individual equations and across equations. This information can be available from past experience of the experimenter with similar studies or studies conducted in the past and/or long association with the study. An example of exact linear restrictions is the Cobb–Douglas production function in economics with the condition of constant returns to scale. This condition gives the sum of regression coefficients as unity.

The knowledge of ratios of dependent variables or expressing some dependent variables as a linear combination of others can be incorporated using $R_2$. This prior information is available from the application of reduced rank regression methods in previous studies featuring the same dependent variables (Ref. [16]).

Eqs. (2.1)–(2.4) specify the multivariate ultrastructural measurement error (MUME) multiple regression model with exact linear restrictions on the coefficient matrix $B$. No distributional assumption is imposed on the measurement error and the random error components except the finiteness of the first four moments of their distributions.

Let $\gamma_{1c}$ and $\gamma_{2c}$ be the Pearson's coefficient of skewness and kurtosis of a random variable $C$. For $i = 1, 2, \dots, q$ and $j = 1, 2, \dots, n$, the following assumptions are made:

1. Elements of vector $E_{(i)} = [\epsilon_{1i}, \epsilon_{2i}, \dots, \epsilon_{ni}]'$ are independent with mean 0, variance $\sigma^2_{\epsilon i}$, third moment $\gamma_{1\epsilon i}\sigma^3_{\epsilon i}$ and fourth moment $(\gamma_{2\epsilon i} + 3)\sigma^4_{\epsilon i}$;
2. $\delta_{ji}$ are independent and identically distributed random variables with mean 0, variance $\sigma^2_\delta$, third moment $\gamma_{1\delta}\sigma^3_\delta$ and fourth moment $(\gamma_{2\delta} + 3)\sigma^4_\delta$;
3. $\psi_{ji}$ are independent and identically distributed random variables with mean 0, variance $\sigma^2_\psi$, third moment $\gamma_{1\psi}\sigma^3_\psi$ and fourth moment $(\gamma_{2\psi} + 3)\sigma^4_\psi$;
4. $\Delta$, $\Psi$ and $E$ are mutually independent;
5. $M_{(n)} \to \sigma_M$ as $n \to \infty$.

Assumption 5 implies that

$$\lim_{n\to\infty} n^{-1}M'M = \lim_{n\to\infty} n^{-1}\sum_{j=1}^{n} M_{(j)}M'_{(j)} = \sigma_M\sigma'_M = \text{finite};$$

$$\lim_{n\to\infty} n^{-1}M'a_n = \lim_{n\to\infty} n^{-1}\sum_{j=1}^{n} M_{(j)} = \sigma_M = \text{finite},$$

where $a_n$ is an $(n \times 1)$ unit vector. In the long run, the possibility of any trend in the observations is avoided by using assumption 5 (Refs. [17,18]).

## 3. Estimation of parameters

It is a well known fact that when predictors are measured with error, we need some prior information for a consistent estimation of parameters (Ref. [5]). Gleser [6] proposed a consistent estimator of regression coefficients using the reliability matrix associated with predictors.

For $\Sigma_X = [n^{-1}M'M + \sigma^2_\psi I_p + \sigma^2_\delta I_p]$ and $\Sigma_D = [n^{-1}M'M + \sigma^2_\psi I_p]$, it can be shown that

$$\lim \Sigma_X = [\sigma_M\sigma'_M + \sigma^2_\psi I_p + \sigma^2_\delta I_p] = \Sigma \quad \text{(say)} \tag{3.1}$$

and

$$\lim \Sigma_D = \left[\sigma_M\sigma'_M + \sigma^2_\psi I_p\right] = \Sigma - \sigma^2_\delta I_p, \tag{3.2}$$

using assumption 5 and Lemma A.1. The reliability matrix $K = \Sigma^{-1}\left(\Sigma - \sigma^2_\delta I_p\right)$ is a generalization of the reliability ratio used in psychometric studies (Refs. [1,5]). A consistent estimator of $K$ can be utilized to construct consistent estimators of regression coefficients. Gleser [6] suggested a consistent estimator of $B$ as

$$\widehat{B}_1 = K_X^{-1}\widehat{B}, \tag{3.3}$$

where $K_X = \Sigma_X^{-1}\Sigma_D$ is a consistent estimator of $K$ and $\widehat{B} = (X'X)^{-1}(X'Z)$ is the LSE of $B$.

Since

$$\operatorname{p\,lim} \widehat{B}_1 = B, \tag{3.4}$$

hence the estimator $\widehat{B}_1$ is consistent. But it does not satisfy the exact linear restrictions given by (2.4).

In the following discussion, we propose three consistent restricted estimators of $B$ using (3.3).

### 3.1. First proposed consistent estimator of B

In the without measurement error case, $D$ is observable. Hence a restricted estimator of $B$ is obtained by minimizing

$$G = \operatorname{tr}\left[(Z - DB)'(Z - DB)\right], \tag{3.5}$$

with respect to $B$ under (2.4) (Ref. [10]). Our aim is to derive a restricted estimator in the measurement error case. Since $D$ is unobservable, we observe $X = D + \Delta$ with additional measurement error $\Delta$. Minimizing

$$G_1 = \operatorname{tr}\left[(Z - XB)'(Z - XB)\right], \tag{3.6}$$

under (2.4) leads to the following estimator

$$\widehat{B}_r = \widehat{B} + (X'X)^{-1}R_1'\left[R_1(X'X)^{-1}R_1'\right]^{-1}\left(\theta - R_1\widehat{B}R_2\right)(R_2'R_2)^{-1}R_2'. \tag{3.7}$$

Using assumptions 1–5 and Lemma A.1, we observe that

$$\text{p}\lim \widehat{B}_r \neq B. \tag{3.8}$$

Hence we have to search for another restricted estimator which is consistent as well. For this purpose, we define

$$G_2 = G_1 - \text{tr}\left[B'(n\Sigma_X)\left(I_p - K_X\right)B\right]. \tag{3.9}$$

The motivation for defining $G_2$ emanates from the following conditional expectation

$$\begin{aligned} E\{G_1|Z, D\} &= E\left\{\text{tr}\left[((Z - DB) - \Delta B)'((Z - DB) - \Delta B)\right]|Z, D\right\} \\ &= G + \text{tr}\left[B'\left(n\sigma_\delta^2 I_p\right)B\right] \quad \text{(using assumption 2)} \\ &= G + \text{tr}\left[B'(n\Sigma_X)\left(I_p - K_X\right)B\right]. \end{aligned} \tag{3.10}$$

The last equality follows, since $\Sigma_X\left(I_p - K_X\right) = \Sigma_X\left(I_p - \Sigma_X^{-1}\Sigma_D\right) = \Sigma_X - \Sigma_D = \sigma_\delta^2 I_p$.

Replacing the unknown $\Sigma_X$ in $G_2$, and using (3.1), (3.9) and Lemma A.1(vi),

$$\widehat{G}_2 = G_1 - \text{tr}\left[B'\left(X'X\right)\left(I_p - K_X\right)B\right]. \tag{3.11}$$

In order to obtain the restricted estimator, we use the method of Lagrangian multipliers and minimize

$$G_3 = \widehat{G}_2 - 2\lambda_1'(\theta - R_1BR_2)\lambda_2, \tag{3.12}$$

where $\lambda_1$ and $\lambda_2$ are vectors of Lagrangian multipliers of order $(r_1 \times 1)$ and $(r_2 \times 1)$. Solving the equations

$$\frac{\partial G_3}{\partial B} = X'XK_XB - X'Z - R_1'\lambda_1\lambda_2'R_2' = 0; \tag{3.13}$$

$$\frac{\partial G_3}{\partial \lambda_1} = (\theta - R_1BR_2)\lambda_2 = 0; \tag{3.14}$$

$$\frac{\partial G_3}{\partial \lambda_2} = (\theta - R_1BR_2)'\lambda_1 = 0, \tag{3.15}$$

we obtain the estimator

$$\widehat{B}_2 = \widehat{B}_1 + K_X^{-1}(X'X)^{-1}R_1'\left[R_1K_X^{-1}(X'X)^{-1}R_1'\right]^{-1}\left(\theta - R_1\widehat{B}_1R_2\right)(R_2'R_2)^{-1}R_2'. \tag{3.16}$$

Using (2.4), (3.3) and (3.4), it is observed that $\text{p}\lim \widehat{B}_2 = B$ and $R_1\widehat{B}_2R_2 = \theta$. Hence $\widehat{B}_2$ is a consistent estimator satisfying the linear restrictions (2.4).

**Remark 3.1.** Let $W = \left(X'X\right)K_X$ be the weight matrix. Then minimization of the weighted function

$$G_W = \text{tr}\left[\left(\widehat{B}_1 - B\right)'W\left(\widehat{B}_1 - B\right)\right] \tag{3.17}$$

with respect to $B$ and under the restrictions (2.4), gives the resulting estimator as $\widehat{B}_2$.

This observation leads to the proposal of two more restricted estimators of $B$.

### 3.2. Second proposed consistent estimator of B

Let the weight matrix be $W = (X'X)$. Another restricted estimator of $B$ is obtained by minimizing $\text{tr}\left[\left(\widehat{B}_1 - B\right)'\left(X'X\right)\left(\widehat{B}_1 - B\right)\right]$ with respect to $B$ when $R_1BR_2 = \theta$. Using the Lagrangian multiplier method, we equate to zero, the first order derivatives of

$$\text{tr}\left[\left(\widehat{B}_1 - B\right)'\left(X'X\right)\left(\widehat{B}_1 - B\right)\right] - 2\lambda_1'(\theta - R_1BR_2)\lambda_2, \tag{3.18}$$

with respect to $B$, $\lambda_1$ and $\lambda_2$. This results in the estimator

$$\widehat{B}_3 = \widehat{B}_1 + (X'X)^{-1}R_1'\left[R_1(X'X)^{-1}R_1'\right]^{-1}\left(\theta - R_1\widehat{B}_1R_2\right)(R_2'R_2)^{-1}R_2'. \tag{3.19}$$

Consistency of the estimator $\widehat{B}_3$ is implied by (3.4) and assumptions 1–5. This consistent estimator also satisfies (2.4), since $R_1\widehat{B}_3R_2 = \theta$.

### 3.3. Third proposed consistent estimator of B

We take the weight matrix $W = I_p$ in (3.17). Minimizing $\operatorname{tr}\left[\left(\widehat{B}_1 - B\right)'\left(\widehat{B}_1 - B\right)\right]$ with respect to $B$ under (2.4), the proposed estimator is found to be

$$\widehat{B}_4 = \widehat{B}_1 + R_1'\left[R_1R_1'\right]^{-1}\left(\theta - R_1\widehat{B}_1R_2\right)(R_2'R_2)^{-1}R_2'. \tag{3.20}$$

Using (2.4) and (3.4), this estimator is found to be consistent and $R_1\widehat{B}_4R_2 = \theta$.

## 4. Large sample properties of estimators

The derivation of the exact distributions of estimators $\widehat{B}_j, j = 1, 2, 3, 4$ is difficult. Even if obtained, the complexity of the expressions makes it very difficult to draw inferences. Hence in the following theorem, we derive the large sample distributions for these estimators.

**Theorem 4.1.** $\left[n^{\frac{1}{2}}\left(\widehat{B}_j - B\right)\right], j = 1, 2, 3, 4$ *have an asymptotic Matrix Normal distribution that is*

$$\left[n^{\frac{1}{2}}\left(\widehat{B}_j - B\right)\right] \xrightarrow{d} MN_{p,q}\left(O_{p\times q}, A_j\left(\Lambda_{qp\times qp}\right)A_j'\right). \tag{4.1}$$

$O_{p\times q}$ *is the mean matrix with all elements zero. The matrix $\Lambda$ is such that when it is partitioned into square matrices of order $(q \times q)$, the $(s, t)$th partitioned matrix for $s, t = 1, \ldots, p$ is given by*

$$\begin{aligned}\Lambda_{st} = {} & \left(\sigma_\delta^2 + \sigma_\psi^2\right)\left\{\operatorname{tr}\left(e_se_t'\right)\left[\Sigma_o + \sigma_\delta^2B'KB\right] + B'\overline{K}e_te_s'\sigma_M\sigma_M'\overline{K}B\right\} + \sigma_\delta^2B'e_te_s'\left(\sigma_\psi^2\overline{K} - \sigma_\delta^2K\right)B \\ & + e_t'\sigma_M\sigma_M'e_s\left(\Sigma_o + \sigma_\delta^2B'K^2B + \sigma_\psi^2B'\overline{K}^2B\right) + N(\gamma).\end{aligned} \tag{4.2}$$

*In the above matrix,*

$e_s$ : *a $(p \times 1)$ vector with the sth element as one and all other elements are zero;*
$N(\gamma)$ : *function of coefficients of skewness and kurtosis as given in Appendix;*

$$\begin{aligned}A_1 &= \left((\Sigma K)^{-1} \otimes I_q\right);\\ A_2 &= A_1 - \left(\left((\Sigma K)^{-1}R_1'\left[R_1(\Sigma K)^{-1}R_1'\right]^{-1}R_1(\Sigma K)^{-1}\right) \otimes \left(R_2\left[R_2'R_2\right]^{-1}R_2'\right)\right);\\ A_3 &= A_1 - \left(\left(\Sigma^{-1}R_1'\left[R_1\Sigma^{-1}R_1'\right]^{-1}R_1(\Sigma K)^{-1}\right) \otimes \left(R_2\left[R_2'R_2\right]^{-1}R_2'\right)\right);\\ A_4 &= A_1 - \left(\left(R_1'\left[R_1R_1'\right]^{-1}R_1(\Sigma K)^{-1}\right) \otimes \left(R_2\left[R_2'R_2\right]^{-1}R_2'\right)\right).\end{aligned}$$

*The proof is given in the Appendix.*

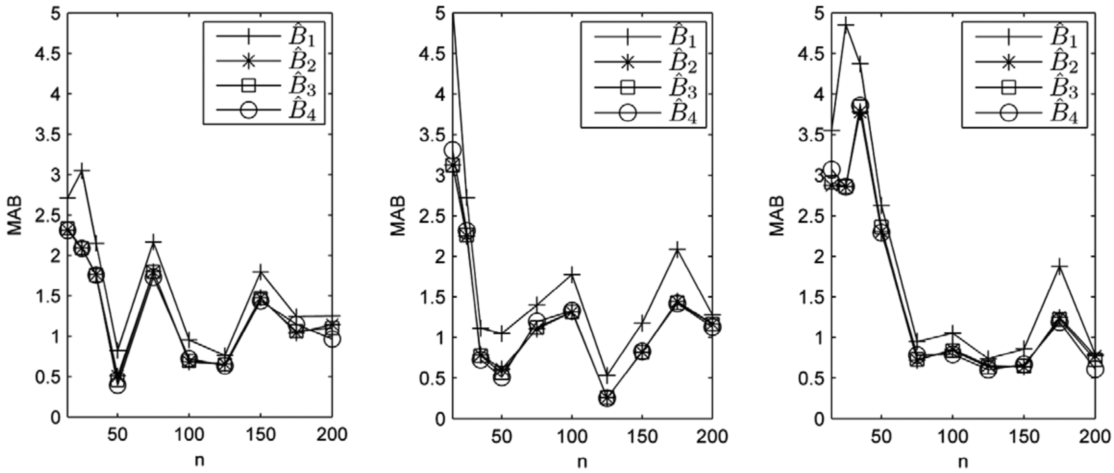
From (4.1), it is clear that all the estimators are asymptotically unbiased. This holds as asymptotic mean of $\left[n^{\frac{1}{2}}\left(\widehat{B}_j - B\right)\right], j = 1, 2, 3, 4$ is a matrix with all elements zero. From (4.2), we observe that the deviation from normality of the random error term $\epsilon_{ji}$ does not affect the asymptotic behavior of the estimators. This further indicates that the non-normality of elements of the measurement error matrix $\Delta$ and the random component matrix $\Psi$ of true predictors affect the asymptotic variance-covariance matrix of estimators. The function $N(\gamma)$ suggests that the non-normality inflates the variance of estimators. When elements of $\Delta$ and $\Psi$ are normally distributed, the function $N(\gamma)$ vanishes from the expression of the asymptotic variance-covariance matrix.

The next section discusses the small sample properties of the estimators through a simulation study.

## 5. Simulation study

Large sample theory tells about the behavior of the estimators' distributions only in the central part. So the overall distributions are studied for finite samples using Monte-Carlo simulations. We use MATLAB for simulation purposes. To assess the effect of departure from normality of measurement error and random error component, the distributions considered are

I. Normal Distribution (with no skewness and kurtosis);
II. Central $t$-Distribution (with zero skewness and non-zero kurtosis);
III. Gamma Distribution (with non-zero skewness and kurtosis).




(a) Normal-Dist. (b) $t$-Dist. (c) Gamma-Dist.

**Fig. 1.** MAB of all estimators.

The elements of matrices $\Delta$ and $\Psi$ are randomly generated from the univariate versions of the above distributions. For simulations, we consider a system where the variance-covariance matrix of the random disturbance matrix $E$ is such that $\sigma_{\epsilon ik} = \rho\sigma_\epsilon^2$ and $\sigma_{\epsilon i}^2 = \sigma_\epsilon^2$, $i$ and $k$ varies from 1 to $q$. $\rho$ is the common correlation coefficient among the columns of $E$. The elements of matrix $E$ are generated using Multivariate Normal, Multivariate $t$ and Multivariate Gamma distributions to incorporate the appropriate correlation among components (Ref. [9]).

Using definition (A.3), we define the mean bias matrix (MBM) and the mean square error matrix (MSEM) respectively as

$$\mathrm{MBM}\left(\widehat{B}\right) = E\left(\widehat{B} - B\right) \quad \text{and}$$

$$\mathrm{MSEM}\left(\widehat{B}\right) = E\left\{\mathrm{vec}\left(\widehat{B} - B\right)' \times \left(\mathrm{vec}\left(\widehat{B} - B\right)'\right)'\right\}.$$

Simulations are performed for various sample sizes to compute the estimators $\widehat{B_j}, j = 1, 2, 3, 4$ for different combinations of $(\rho, \sigma_\epsilon^2, \sigma_\psi^2, \sigma_\delta^2)$. These are specified as (0.8, 0.5, 0.5, 0.5), (0.8, 0.5, 0.5, 1.0), (0.8, 0.5, 1.0, 0.5), (0.8, 0.5, 1.0, 1.0), (0.8, 1.0, 0.5, 0.5), (0.8, 1.0, 0.5, 1.0), (0.8, 1.0, 1.0, 0.5), (0.8, 1.0, 1.0, 1.0).

The matrices $R_1, R_2, \theta$ and $B$ are fixed a priori as

$$B = \begin{bmatrix} 2.4 & 12.4 & -1.9 \\ 1.3 & -5.3 & 6.5 \\ 1.9 & 7.5 & -1.3 \end{bmatrix}; \qquad R_1 = \begin{bmatrix} 0.3 & 0.5 & 0.8 \\ 0.5 & 0.7 & 0.8 \end{bmatrix}; \qquad R_2 = \begin{bmatrix} 1.0 & 0.3 \\ 1.0 & 0.4 \\ 0.5 & 0.4 \end{bmatrix}; \qquad \theta = \begin{bmatrix} 10.78 & 4.35 \\ 13.40 & 5.51 \end{bmatrix}.$$

The random variables following the considered distributions have been scaled suitably to have mean zero and variances specified in different combinations. These distributions differ only with respect to skewness and kurtosis. The experiment is run 10,000 times for each combination and MBM and MSEM are computed empirically for $\widehat{B_j}, j = 1, 2, 3$ and 4. To compare the MBM of estimators, we compute the Frobenius norm of MBM defined as $\|\mathrm{MBM}\|_F = \sqrt{\mathrm{tr}\left\{\mathrm{MBM}' \times \mathrm{MBM}\right\}}$. This is called the mean absolute bias (MAB). Similarly the MSEMs of estimators are compared using trace of the MSEM denoted by tr(MSEM).

For Normal, $t$ and Gamma distribution, Fig. 1 shows the plots of MAB of all estimators for various sample sizes and $\left(\rho, \sigma_\epsilon^2, \sigma_\psi^2, \sigma_\delta^2\right) = (0.8, 0.5, 0.5, 1.0)$.

Looking at Fig. 1, it is observed that MAB fluctuates a lot for different sample sizes and may not lead to clear conclusions. Hence instead of mean bias, we prefer to use the concept of median bias. The median bias matrix (MedBM) and median square error matrix (MedSEM) are defined as

$$\mathrm{MedBM}\left(\widehat{B}\right) = \left(\mathrm{median}\left(\widehat{B}\right) - B\right) \quad \text{and}$$

$$\mathrm{MedSEM}\left(\widehat{B}\right) = \mathrm{median}\left\{\mathrm{vec}\left(\widehat{B} - B\right)' \times \left(\mathrm{vec}\left(\widehat{B} - B\right)'\right)'\right\}.$$

The median absolute bias (MedAB) is the Frobenius norm of MedBM. The simulations results for MedAB and tr(MedSEM), of all estimators for different sample sizes, distributions and parametric combinations are displayed in Tables 1–8.

Tables 1 and 2 lead to the conclusion that the bias and variance of estimators increase as the measurement error variance $\sigma_\delta^2$ increases. Similarly, a comparison of Tables 1 and 3 reveals that the bias and variances of estimators decrease as the variance of true explanatory variables increases.

**Table 1**
MedAB and tr(MedSEM) of estimators for $\left(\rho, \sigma_\epsilon^2, \sigma_\psi^2, \sigma_\delta^2\right) = (0.8, 0.5, 0.5, 0.5)$.

| | | $n = 15$ | | | | $n = 100$ | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | $\widehat{B}_1$ | $\widehat{B}_2$ | $\widehat{B}_3$ | $\widehat{B}_4$ | $\widehat{B}_1$ | $\widehat{B}_2$ | $\widehat{B}_3$ | $\widehat{B}_4$ |
| Normal | MedAB | 0.2258 | 0.1752 | 0.1757 | 0.2199 | 0.0341 | 0.0282 | 0.0283 | 0.0371 |
| | tr(MedSEM) | 4.9215 | 2.7934 | 2.7964 | 3.0420 | 0.5469 | 0.3467 | 0.3475 | 0.3589 |
| $t$ | MedAB | 0.3001 | 0.2346 | 0.2330 | 0.2953 | 0.0667 | 0.0535 | 0.0545 | 0.0653 |
| | tr(MedSEM) | 5.0825 | 2.9088 | 2.9125 | 3.1289 | 0.6017 | 0.3780 | 0.3781 | 0.3918 |
| Gamma | MedAB | 0.2950 | 0.2477 | 0.2440 | 0.2962 | 0.0522 | 0.0413 | 0.0412 | 0.0489 |
| | tr(MedSEM) | 5.1615 | 2.9686 | 2.9716 | 3.1850 | 0.5773 | 0.3742 | 0.3744 | 0.3780 |

**Table 2**
MedAB and tr(MedSEM) of estimators for $\left(\rho, \sigma_\epsilon^2, \sigma_\psi^2, \sigma_\delta^2\right) = (0.8, 0.5, 0.5, 1.0)$.

| | | $n = 15$ | | | | $n = 100$ | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | $\widehat{B}_1$ | $\widehat{B}_2$ | $\widehat{B}_3$ | $\widehat{B}_4$ | $\widehat{B}_1$ | $\widehat{B}_2$ | $\widehat{B}_3$ | $\widehat{B}_4$ |
| Normal | MedAB | 0.4315 | 0.3494 | 0.3476 | 0.4255 | 0.0751 | 0.0660 | 0.0688 | 0.0810 |
| | tr(MedSEM) | 10.5905 | 5.7986 | 5.8392 | 6.4087 | 1.1040 | 0.6588 | 0.6646 | 0.6883 |
| $t$ | MedAB | 0.5012 | 0.4112 | 0.4126 | 0.5012 | 0.1131 | 0.0995 | 0.0998 | 0.1122 |
| | tr(MedSEM) | 11.6394 | 6.3656 | 6.4243 | 7.0478 | 1.2760 | 0.7928 | 0.7957 | 0.8201 |
| Gamma | MedAB | 0.5362 | 0.4559 | 0.4570 | 0.5300 | 0.0926 | 0.0725 | 0.0719 | 0.0816 |
| | tr(MedSEM) | 11.5267 | 6.4006 | 6.3943 | 6.9241 | 1.2716 | 0.8015 | 0.8024 | 0.8099 |

**Table 3**
MedAB and tr(MedSEM) of estimators for $\left(\rho, \sigma_\epsilon^2, \sigma_\psi^2, \sigma_\delta^2\right) = (0.8, 0.5, 1.0, 0.5)$.

| | | $n = 15$ | | | | $n = 100$ | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | $\widehat{B}_1$ | $\widehat{B}_2$ | $\widehat{B}_3$ | $\widehat{B}_4$ | $\widehat{B}_1$ | $\widehat{B}_2$ | $\widehat{B}_3$ | $\widehat{B}_4$ |
| Normal | MedAB | 0.1695 | 0.1463 | 0.1456 | 0.1717 | 0.0322 | 0.0173 | 0.0179 | 0.0258 |
| | tr(MedSEM) | 4.0290 | 2.3501 | 2.3504 | 2.5303 | 0.4745 | 0.3067 | 0.3060 | 0.3137 |
| $t$ | MedAB | 0.2462 | 0.1960 | 0.1967 | 0.2307 | 0.0473 | 0.0378 | 0.0382 | 0.0447 |
| | tr(MedSEM) | 4.0866 | 2.4043 | 2.4090 | 2.6061 | 0.5103 | 0.3278 | 0.3284 | 0.3357 |
| Gamma | MedAB | 0.2004 | 0.1427 | 0.1436 | 0.1834 | 0.0266 | 0.0147 | 0.0152 | 0.0225 |
| | tr(MedSEM) | 4.1269 | 2.4700 | 2.4701 | 2.6496 | 0.5011 | 0.3240 | 0.3235 | 0.3286 |

**Table 4**
MedAB and tr(MedSEM) of estimators for $\left(\rho, \sigma_\epsilon^2, \sigma_\psi^2, \sigma_\delta^2\right) = (0.8, 0.5, 1.0, 1.0)$.

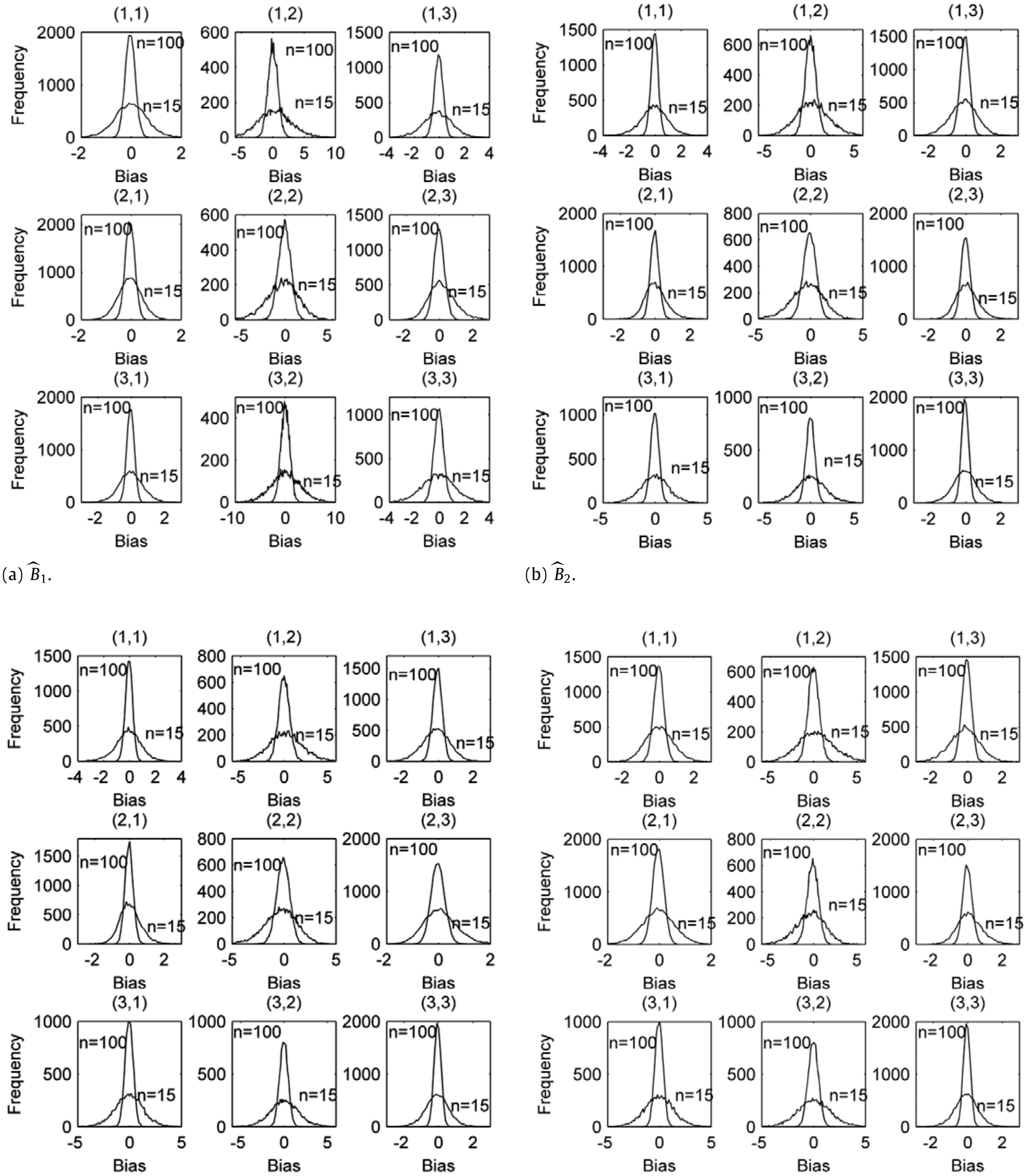
| | | $n = 15$ | | | | $n = 100$ | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | $\widehat{B}_1$ | $\widehat{B}_2$ | $\widehat{B}_3$ | $\widehat{B}_4$ | $\widehat{B}_1$ | $\widehat{B}_2$ | $\widehat{B}_3$ | $\widehat{B}_4$ |
| Normal | MedAB | 0.3013 | 0.2253 | 0.2241 | 0.2783 | 0.0607 | 0.0382 | 0.0395 | 0.0522 |
| | tr(MedSEM) | 8.5831 | 4.8597 | 4.9021 | 5.3184 | 0.9417 | 0.5954 | 0.5952 | 0.6095 |
| $t$ | MedAB | 0.4307 | 0.3379 | 0.3400 | 0.4163 | 0.0794 | 0.0701 | 0.0711 | 0.0785 |
| | tr(MedSEM) | 9.2685 | 5.2174 | 5.2455 | 5.7161 | 1.0759 | 0.6801 | 0.6828 | 0.7022 |
| Gamma | MedAB | 0.4266 | 0.3479 | 0.3404 | 0.4058 | 0.0755 | 0.0529 | 0.0536 | 0.0630 |
| | tr(MedSEM) | 9.0768 | 5.2846 | 5.2878 | 5.6606 | 1.0473 | 0.6711 | 0.6703 | 0.6785 |

We also perform a graphical analysis of simulation results to study the properties of estimators in depth. Only a few graphs are presented here. Fig. 2 shows the frequency curves of bias in the elements of different estimators. The results are for $\left(\rho, \sigma_\epsilon^2, \sigma_\psi^2, \sigma_\delta^2\right) = (0.8, 0.5, 0.5, 1.0)$ when elements of $E$, $\Psi$ and $\Delta$ are normally distributed.

In Fig. 2, the frequency curves are more peaked and concentrated around zero for $n = 100$ as compared to $n = 15$ for all elements of different estimators. Hence it can be concluded that estimators become unbiased and consistent as the sample size increases.

The tr(MedSEM) and MedAB of all estimators against the sample size are plotted in Figs. 3 and 4. In these figures, $\left(\rho, \sigma_\epsilon^2, \sigma_\psi^2, \sigma_\delta^2\right) = (0.8, 0.5, 0.5, 1.0)$ and elements of $E$, $\Psi$ and $\Delta$ are (a) normally distributed, (b) $t$ distributed and (c) gamma distributed.

Figs. 3 and 4 show that for all estimators, tr(MedSEM) and MedAB gradually decrease to zero as the sample size increases. This validates the theoretical findings that the estimators are asymptotically consistent and unbiased.




(a) $\widehat{B}_1$.

(b) $\widehat{B}_2$.

(c) $\widehat{B}_3$.

(d) $\widehat{B}_4$.

**Fig. 2.** Frequency curves of bias for different estimators.

**Table 5**
MedAB and tr(MedSEM) of estimators for $\left(\rho, \sigma_\epsilon^2, \sigma_\psi^2, \sigma_\delta^2\right) = (0.8, 1.0, 0.5, 0.5)$.

| | | $n = 15$ | | | | $n = 100$ | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | $\widehat{B}_1$ | $\widehat{B}_2$ | $\widehat{B}_3$ | $\widehat{B}_4$ | $\widehat{B}_1$ | $\widehat{B}_2$ | $\widehat{B}_3$ | $\widehat{B}_4$ |
| Normal | MedAB | 0.2500 | 0.1953 | 0.1970 | 0.2456 | 0.0401 | 0.0334 | 0.0321 | 0.0401 |
| | tr(MedSEM) | 4.8627 | 2.7922 | 2.8024 | 3.0469 | 0.5430 | 0.3429 | 0.3439 | 0.3564 |
| $t$ | MedAB | 0.3285 | 0.2523 | 0.2562 | 0.3143 | 0.0525 | 0.0487 | 0.0493 | 0.0556 |
| | tr(MedSEM) | 5.1624 | 2.9335 | 2.9487 | 3.2262 | 0.6128 | 0.3856 | 0.3871 | 0.3992 |
| Gamma | MedAB | 0.3279 | 0.2582 | 0.2604 | 0.3176 | 0.0606 | 0.0506 | 0.0519 | 0.0635 |
| | tr(MedSEM) | 5.2223 | 3.0127 | 3.0147 | 3.2321 | 0.6077 | 0.3832 | 0.3824 | 0.3861 |

**Table 6**
MedAB and tr(MedSEM) of estimators for $(\rho, \sigma_\epsilon^2, \sigma_\psi^2, \sigma_\delta^2) = (0.8, 1.0, 0.5, 1.0)$.

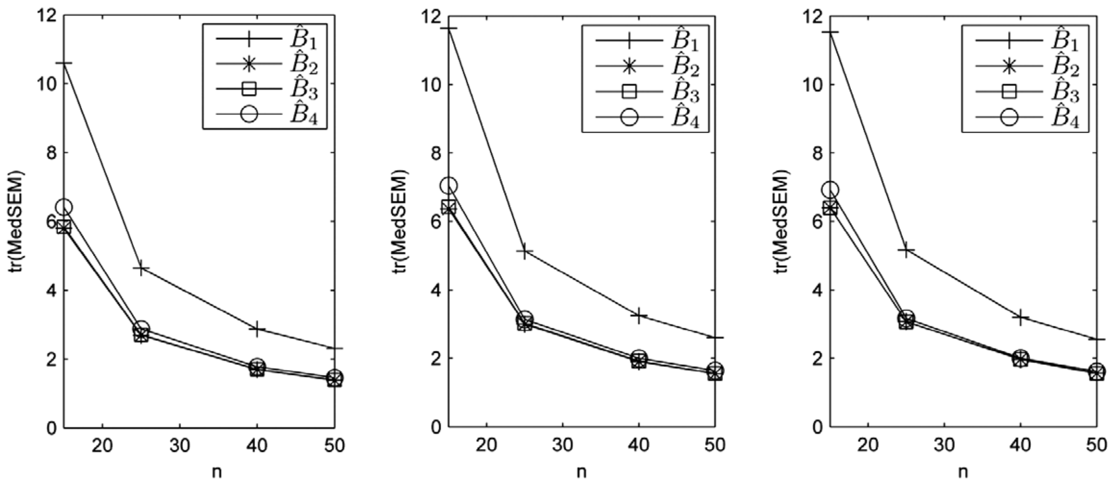
| | | $n = 15$ | | | | $n = 100$ | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | $\widehat{B}_1$ | $\widehat{B}_2$ | $\widehat{B}_3$ | $\widehat{B}_4$ | $\widehat{B}_1$ | $\widehat{B}_2$ | $\widehat{B}_3$ | $\widehat{B}_4$ |
| Normal | MedAB | 0.4420 | 0.3429 | 0.3429 | 0.4220 | 0.0725 | 0.0577 | 0.0603 | 0.0749 |
| | tr(MedSEM) | 10.7463 | 5.9002 | 5.9588 | 6.5587 | 1.0766 | 0.6724 | 0.6745 | 0.6990 |
| $t$ | MedAB | 0.5879 | 0.4724 | 0.4689 | 0.5811 | 0.1104 | 0.0887 | 0.0893 | 0.1067 |
| | tr(MedSEM) | 12.0278 | 6.5083 | 6.5860 | 7.2060 | 1.2845 | 0.7838 | 0.7887 | 0.8131 |
| Gamma | MedAB | 0.5337 | 0.4527 | 0.4534 | 0.5222 | 0.1147 | 0.0986 | 0.0974 | 0.1111 |
| | tr(MedSEM) | 11.4446 | 6.3605 | 6.3712 | 6.8856 | 1.2748 | 0.8114 | 0.8099 | 0.8225 |

**Table 7**
MedAB and tr(MedSEM) of estimators for $(\rho, \sigma_\epsilon^2, \sigma_\psi^2, \sigma_\delta^2) = (0.8, 1.0, 1.0, 0.5)$.

| | | $n = 15$ | | | | $n = 100$ | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | $\widehat{B}_1$ | $\widehat{B}_2$ | $\widehat{B}_3$ | $\widehat{B}_4$ | $\widehat{B}_1$ | $\widehat{B}_2$ | $\widehat{B}_3$ | $\widehat{B}_4$ |
| Normal | MedAB | 0.1838 | 0.1317 | 0.1272 | 0.1729 | 0.0194 | 0.0136 | 0.0137 | 0.0179 |
| | tr(MedSEM) | 4.0306 | 2.3411 | 2.3492 | 2.5595 | 0.4712 | 0.3011 | 0.3016 | 0.3101 |
| $t$ | MedAB | 0.2454 | 0.1974 | 0.1967 | 0.2428 | 0.0666 | 0.0528 | 0.0534 | 0.0616 |
| | tr(MedSEM) | 4.1819 | 2.4050 | 2.4135 | 2.5953 | 0.5197 | 0.3319 | 0.3328 | 0.3399 |
| Gamma | MedAB | 0.2372 | 0.1924 | 0.1931 | 0.2343 | 0.0507 | 0.0429 | 0.0437 | 0.0471 |
| | tr(MedSEM) | 4.1186 | 2.4061 | 2.4023 | 2.5687 | 0.5155 | 0.3351 | 0.3347 | 0.3388 |

**Table 8**
MedAB and tr(MedSEM) of estimators for $(\rho, \sigma_\epsilon^2, \sigma_\psi^2, \sigma_\delta^2) = (0.8, 1.0, 1.0, 1.0)$.

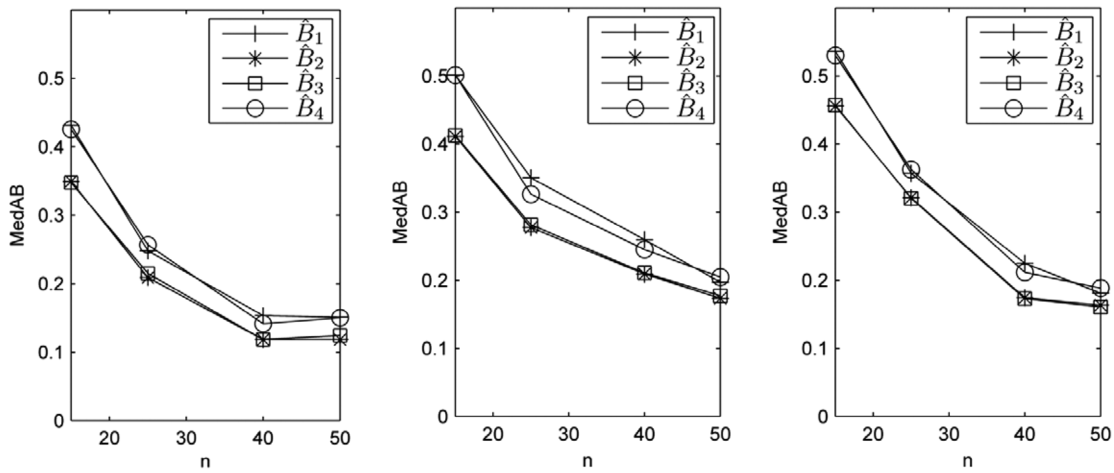
| | | $n = 15$ | | | | $n = 100$ | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | $\widehat{B}_1$ | $\widehat{B}_2$ | $\widehat{B}_3$ | $\widehat{B}_4$ | $\widehat{B}_1$ | $\widehat{B}_2$ | $\widehat{B}_3$ | $\widehat{B}_4$ |
| Normal | MedAB | 0.3142 | 0.2507 | 0.2472 | 0.3016 | 0.0432 | 0.0373 | 0.0386 | 0.0475 |
| | tr(MedSEM) | 8.2648 | 4.7693 | 4.7894 | 5.1759 | 0.9572 | 0.6040 | 0.6051 | 0.6214 |
| $t$ | MedAB | 0.4552 | 0.3600 | 0.3612 | 0.4499 | 0.0850 | 0.0616 | 0.0635 | 0.0750 |
| | tr(MedSEM) | 9.1887 | 5.1926 | 5.1990 | 5.6467 | 1.0582 | 0.6811 | 0.6829 | 0.6954 |
| Gamma | MedAB | 0.4370 | 0.3536 | 0.3525 | 0.4266 | 0.0495 | 0.0391 | 0.0395 | 0.0521 |
| | tr(MedSEM) | 9.0793 | 5.2467 | 5.2428 | 5.6679 | 1.0636 | 0.6746 | 0.6740 | 0.6876 |




(a) Normal-Dist. (b) $t$-Dist. (c) Gamma-Dist.

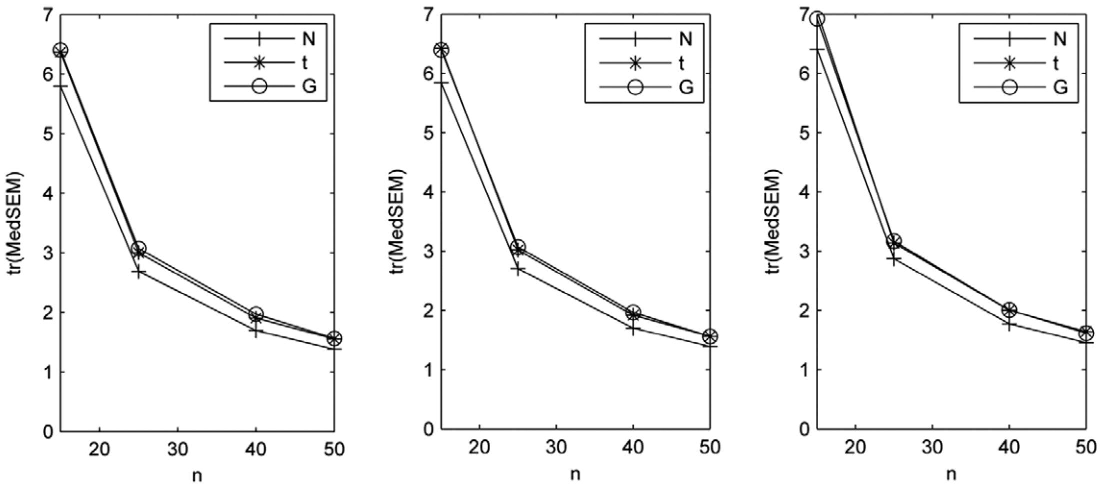
**Fig. 3.** tr(MedSEM) of all estimators.

From Tables 1–8 and Figs. 3 and 4, it is concluded that tr(MedSEM) and MedAB for restricted estimators are less than that for unrestricted estimator for all distributions and sample sizes. This means that use of prior information improves the




(a) Normal-Dist. (b) $t$-Dist. (c) Gamma-Dist.

**Fig. 4.** MedAB of all estimators.




(a) $\widehat{B}_2$. (b) $\widehat{B}_3$. (c) $\widehat{B}_4$.

**Fig. 5.** tr(MedSEM) of estimators for different distributions.

estimators in terms of the variance-covariance matrix and bias. Since

$$\operatorname{tr}\left(\operatorname{MedSEM}\left(\widehat{B}_2\right)\right) \cong \operatorname{tr}\left(\operatorname{MedSEM}\left(\widehat{B}_3\right)\right) < \operatorname{tr}\left(\operatorname{MedSEM}\left(\widehat{B}_4\right)\right), \tag{5.1}$$

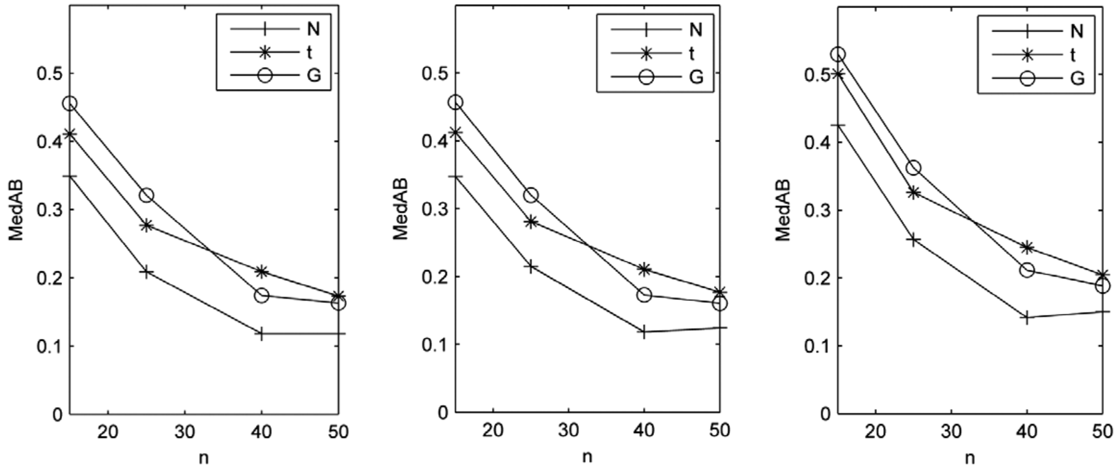
and

$$\operatorname{MedAB}\left(\widehat{B}_2\right) \cong \operatorname{MedAB}\left(\widehat{B}_3\right) < \operatorname{MedAB}\left(\widehat{B}_4\right), \tag{5.2}$$

hence weighted restricted estimators $\hat{B}_2$ and $\hat{B}_3$ are equivalent in terms of variance-covariance matrix and bias but both dominate the unweighted restricted estimator $\hat{B}_4$. This happens for small as well as large sample sizes. The same ordering is observed with respect to tr(MedSEM) and MedAB for all parametric combinations.

To assess the effect of non-normality of random error component and measurement error, we plot tr(MedSEM) and MedAB of estimators for Normal, $t$ and Gamma distributions in Figs. 5 and 6. This is done for $\left(\rho, \sigma_\epsilon^2, \sigma_\psi^2, \sigma_\delta^2\right) = (0.8, 0.5, 0.5, 1.0)$.

It is observed that tr(MedSEM) and MedAB are smallest for Normal distribution. From the tables, Figs. 5 and 6, it is noticed that the variance and bias of the estimators gets inflated due to non-normality. In case of Gamma and $t$-distributions, the variances of estimators are close to each other. However, the bias is greater in the case of Gamma-distribution as compared to $t$-distribution when the sample size is small. As the sample size increases, the trend gets reversed.




(a) $\widehat{B}_2$. (b) $\widehat{B}_3$. (c) $\widehat{B}_4$.

**Fig. 6.** MedAB of estimators for different distributions.

## 6. Conclusion

Using the knowledge of the reliability matrix associated with predictor variables, three consistent and restricted estimators of regression coefficient matrix are proposed. No distributional form is imposed on measurement errors and random error components. Asymptotic properties of the estimators are derived. The effect of departure from Normality of measurement error and the small sample properties of the estimators are studied using a Monte-Carlo simulation study. It is observed that the incorporation of additional information in the form of exact linear restrictions provides better estimators.

## Acknowledgments

The authors are grateful to the referees for their valuable suggestions.

## Appendix

**Definition A.1** (*[4]*)**.** Let $\{A_n : n = 1, 2, \ldots\}$ be a sequence of random matrices and $\{b_n : n = 1, 2, \ldots\}$ be a sequence of real numbers. For all $i$ and $j$,

(i) $A_n = \left((A_n)_{ij}\right) = O_p(b_n)$ if every element of the random matrix $A_n$ is $O_P(b_n)$, that is, $\forall n, \varepsilon > 0$ and a positive real number $k$ (depending on $\varepsilon$),

$$P\left(\left|\left((A_n)_{ij}\right)\right| \leq kb_n\right) \geq 1 - \varepsilon;$$

(ii) $A_n = o_P(b_n)$ if every element of the random matrix $A_n$ is $o_P(b_n)$, that is, $\mathrm{p}\lim \frac{(A_n)_{ij}}{b_n} = 0$;

(iii) $\mathrm{p}\lim A_n = A$ if $\mathrm{p}\lim (A_n)_{ij} = A_{ij}$.

**Definition A.2** (Vec(.) *Operator*)**.** For a $(p \times q)$ matrix $A = \left(a_1, \ldots, a_q\right)$ where $a_i$; $i = 1, \ldots, q$ each of order $(p \times 1)$, are columns of $A$, vec(.) is given by

$$\text{vec}(A) = \begin{bmatrix} a_1' & \cdots & a_q' \end{bmatrix}', \quad \text{a column of order } (pq \times 1).$$

The following definition is extracted from [8].

**Definition A.3.** The random matrix $X_{p\times n}$ is said to follow a matrix-variate Normal distribution with the mean matrix $M_{p\times n}$ and the covariance matrix $(W \otimes \Omega)$ if $\text{vec}\left(X'\right) \sim N_{pn}\left(\text{vec}\left(M'\right), W \otimes \Omega\right)$ where $W_{p\times p} > 0$ and $\Omega_{n\times n} > 0$.

The notation used is $X \sim MN_{p,n}\left(M, W \otimes \Omega\right)$. More details about matrix-variate Normal distribution can be found in [2,12].

The following lemmas state a few results that are used in the derivations that follow.

**Lemma A.1.** *As* $n \to \infty$

(i) $n^{-\frac{1}{2}}M'E = n^{-\frac{1}{2}}M'\Psi = n^{-\frac{1}{2}}M'\Delta = O_P(1)$;

(ii) $n^{-\frac{1}{2}}\Psi'\Delta = n^{-\frac{1}{2}}\Delta'E = n^{-\frac{1}{2}}\Psi'E = O_P(1)$;

(iii) $n^{-\frac{1}{2}}\Delta'\Delta - n^{\frac{1}{2}}\sigma_\delta^2 I_p = O_P(1)$, $n^{-\frac{1}{2}}\Psi'\Psi - n^{\frac{1}{2}}\sigma_\psi^2 I_p = O_P(1)$;
(iv) $\text{p}\lim(n^{-1}\Delta'\Delta) = \sigma_\delta^2 I_p$, $\text{p}\lim(n^{-1}\Psi'\Psi) = \sigma_\psi^2 I_p$;
(v) $\text{p}\lim(n^{-1}M'E) = \text{p}\lim(n^{-1}M'\Psi) = \text{p}\lim(n^{-1}M'\Delta) = \text{p}\lim(n^{-1}\Psi'\Delta) = \text{p}\lim(n^{-1}\Delta'E) = \text{p}\lim(n^{-1}\Psi'E) = 0$;
(vi) $\text{p}\lim(n^{-1}X'X) = \Sigma$, $\text{p}\lim(n^{-1}X'Z) = \left(\Sigma - \sigma_\delta^2 I_p\right)B$ *where* $\Sigma = [\sigma_M\sigma_M' + \sigma_\psi^2 I_p + \sigma_\delta^2 I_p]$.

The proofs follow using the Definition A.1 and assumptions 1–5.

**Lemma A.2.** *Let* $C = \left(c_{ij}\right)$ *be a* $(m \times m)$ *matrix. Let* $\|C\|_1 = \max_{1\le i\le m}\sum_{j=1}^{m}\left|c_{ij}\right|$ *and* $\|C\|_2 = \max_{1\le j\le m}\sum_{i=1}^{m}\left|c_{ij}\right|$ *be the maximum column sum and maximum row sum matrix norms respectively. If* $\|C\|_1 < 1$ *and/or* $\|C\|_2 < 1$*, then* $(I_m - C)$ *is invertible and* $(I_m - C)^{-1} = \sum_{i=0}^{\infty} C^i$*, where* $C^o = I_m$.

For the proof, one can refer to [14].

**Lemma A.3.** *Let* $V_n = \sum_{j=1}^{n} U_{jn}X_j$ *where* $X_1, \ldots, X_n$ *are* $(p \times 1)$ *independent and identically distributed random vectors with* $E\left(X_j\right) = 0$*, and* $U_{1n}, \ldots, U_{nn}$ *are* $(q \times p)$ *non-stochastic matrices. Suppose that* $\lim_{n\to\infty}\text{cov}\left(V_n\right) = \Lambda$*; where* $\left|\Lambda_{ij}\right| < \infty$*, for each* $i, j$ *and* $\Lambda$ *is* $+$ *ve definite. If there exists a function* $\omega(n)$ *such that* $\lim_{n\to\infty}\omega(n) = \infty$*, and if elements of* $\omega(n)U_{jn}$ *are bounded, then* $V_n \xrightarrow{d} N_q(0, \Lambda)$ *as* $n \to \infty$.

The above result, known as the *Central Limit Theorem* is due to Malinvaud [11].

**Lemma A.4.** $\text{vec}(ABC) = \left(C' \otimes A\right)\text{vec}(B)$*, where* $A, B$ *and* $C$ *are three matrices such that the product* $(ABC)$ *is defined.*

A few expressions, helpful in the proof of Theorem 4.1 are derived below.

Define $H = n^{-\frac{1}{2}}S - n^{\frac{1}{2}}\Sigma_X$, where $S = X'X$ and let $h = n^{-\frac{1}{2}}\left(X'\left[E - \Delta B\right]\right) + n^{\frac{1}{2}}\sigma_\delta^2 B$.

Using Lemma A.1, it can be easily seen that $h = O_P(1)$, $H = O_P(1)$, $\Sigma_X = O(1)$ and $\Sigma_D = O(1)$, where $\Sigma_X$ and $\Sigma_D$ are defined in Section 3. Using definitions of $H$ and $h$, we have

$$\begin{aligned} S^{-1} &= \left[n\Sigma_X + n^{\frac{1}{2}}H\right]^{-1} = \left[n\Sigma_X\left[I_p + n^{-\frac{1}{2}}\Sigma_X^{-1}H\right]\right]^{-1} \\ &= n^{-1}\left[I_p + n^{-\frac{1}{2}}\Sigma_X^{-1}H\right]^{-1}\Sigma_X^{-1} = n^{-1}\left[I_p - n^{-\frac{1}{2}}\Sigma_X^{-1}H + O_P\left(n^{-1}\right)\right]\Sigma_X^{-1}, \quad \text{using Lemma A.2} \\ &= n^{-1}\left[I_p - n^{-\frac{1}{2}}\Sigma_X^{-1}H\right]\Sigma_X^{-1} + O_P\left(n^{-2}\right) \end{aligned} \tag{A.1}$$

and $X'\left[E - \Delta B\right] = n^{\frac{1}{2}}h - n\sigma_\delta^2 B$. Now using (2.1), (2.2) and (3.3), we have

$$\begin{aligned} \widehat{B}_1 &= K_X^{-1}\widehat{B} = \Sigma_D^{-1}\Sigma_X\left[B + S^{-1}X'\left(E - \Delta B\right)\right] \\ &= \Sigma_D^{-1}\Sigma_X\left[B + \left\{n^{-1}\left[I_p - n^{-\frac{1}{2}}\Sigma_X^{-1}H\right]\Sigma_X^{-1} + O_P\left(n^{-2}\right)\right\}\left\{n^{\frac{1}{2}}h - n\sigma_\delta^2 B\right\}\right] \\ &= n^{-\frac{1}{2}}\Sigma_D^{-1}h + n^{-\frac{1}{2}}\Sigma_D^{-1}H\left(\sigma_\delta^2\Sigma_X^{-1}\right)B + \Sigma_D^{-1}\left[\Sigma_X - \sigma_\delta^2 I_p\right]B + O_P\left(n^{-1}\right) \\ &= n^{-\frac{1}{2}}\Sigma_D^{-1}h + n^{-\frac{1}{2}}\Sigma_D^{-1}H\overline{K}_X B + \Sigma_D^{-1}\Sigma_D B + O_P\left(n^{-1}\right) \quad \text{where } \overline{K}_X = \left[I_p - K_X\right] = \sigma_\delta^2\Sigma_X^{-1}. \end{aligned}$$

$$\text{Hence } n^{\frac{1}{2}}\left(\widehat{B}_1 - B\right) = \Sigma_D^{-1}\left[h + H\overline{K}_X B\right] + O_P\left(n^{-\frac{1}{2}}\right) \tag{A.2}$$

$$\text{and} \quad n^{\frac{1}{2}}\left(\widehat{B}_1 - B\right)' = \left[h' + B'\overline{K}_X H\right]\Sigma_D^{-1} + O_P\left(n^{-\frac{1}{2}}\right). \tag{A.3}$$

Applying the vec( ) operator on both sides of (A.3) and using Lemma A.4, we get

$$\text{vec}\left[n^{\frac{1}{2}}\left(\widehat{B}_1 - B\right)'\right] = \left(\Sigma_D^{-1} \otimes I_q\right)\left[\text{vec}\left(h'\right) + \text{vec}\left(B'\overline{K}_X H\right)\right] + O_P\left(n^{-\frac{1}{2}}\right). \tag{A.4}$$

For the estimator $\widehat{B}_2$, using (3.16)

$$\left(\widehat{B}_2 - B\right) = \left(\widehat{B}_1 - B\right) + K_X^{-1}S^{-1}R_1'\left[R_1K_X^{-1}S^{-1}R_1'\right]^{-1}\left(\theta - R_1\widehat{B}_1R_2\right)(R_2'R_2)^{-1}R_2'. \tag{A.5}$$

Eq. (A.1) gives

$$\begin{aligned} (SK_X)^{-1} &= K_X^{-1}S^{-1} = \Sigma_D^{-1}\Sigma_X S^{-1} = \Sigma_D^{-1}\Sigma_X\left[n^{-1}\left(I_p - n^{-\frac{1}{2}}\Sigma_X^{-1}H\right)\Sigma_X^{-1} + O_P\left(n^{-2}\right)\right] \\ &= n^{-1}\left[\Sigma_D^{-1}\Sigma_X\left(I_p - n^{-\frac{1}{2}}\Sigma_X^{-1}H\right)\Sigma_X^{-1} + O_P\left(n^{-1}\right)\right]. \end{aligned} \tag{A.6}$$

Using (A.6), we write

$$\begin{aligned} n^{-1}\left[R_1K_X^{-1}S^{-1}R_1'\right]^{-1} &= \left[R_1\left(nK_X^{-1}S^{-1}\right)R_1'\right]^{-1} \\ &= \left[R_1\left\{\Sigma_D^{-1}\Sigma_X\left[I_p - n^{-\frac{1}{2}}\Sigma_X^{-1}H\right]\Sigma_X^{-1} + O_P\left(n^{-1}\right)\right\}R_1'\right]^{-1} \\ &= \left[R_{1D} - n^{-\frac{1}{2}}R_1\Sigma_D^{-1}H\Sigma_X^{-1}R_1' + O_P\left(n^{-1}\right)\right]^{-1} \quad \text{where } R_{1D} = R_1\Sigma_D^{-1}R_1' \\ &= \left[R_{1D}\left\{I_{r_1} - \left(n^{-\frac{1}{2}}R_{1D}^{-1}R_1\Sigma_D^{-1}H\Sigma_X^{-1}R_1' - O_P\left(n^{-1}\right)\right)\right\}\right]^{-1} \\ &= \left[I_{r_1} - \left(n^{-\frac{1}{2}}R_{1D}^{-1}R_1\Sigma_D^{-1}H\Sigma_X^{-1}R_1' - O_P\left(n^{-1}\right)\right)\right]^{-1}R_{1D}^{-1} \\ &= \left[I_{r_1} + n^{-\frac{1}{2}}R_{1D}^{-1}R_1\Sigma_D^{-1}H\Sigma_X^{-1}R_1'\right]R_{1D}^{-1} + O_P\left(n^{-1}\right), \quad \text{using Lemma A.2.} \end{aligned} \tag{A.7}$$

Using (A.2), $\widehat{B}_1 = n^{-\frac{1}{2}}\Sigma_D^{-1}\left[h + H\overline{K}_XB\right] + B + O_P\left(n^{-1}\right)$.

Hence $R_1\widehat{B}_1R_2 = n^{-\frac{1}{2}}R_1\Sigma_D^{-1}\left[h + H\overline{K}_XB\right]R_2 + R_1BR_2 + O_P\left(n^{-1}\right)$.

Using (2.4), this gives $\theta - R_1\widehat{B}_1R_2 = -n^{-\frac{1}{2}}R_1\Sigma_D^{-1}\left[h + H\overline{K}_XB\right]R_2 + O_P\left(n^{-1}\right)$

$$\Rightarrow \left(\theta - R_1\widehat{B}_1R_2\right)\left(R_2'R_2\right)^{-1}R_2' = -n^{-\frac{1}{2}}R_1\Sigma_D^{-1}\left[h + H\overline{K}_XB\right]R_2\left(R_2'R_2\right)^{-1}R_2' + O_P\left(n^{-1}\right). \tag{A.8}$$

Using (A.2) and (A.6)–(A.8) in (A.5), we have

$$\begin{aligned} \left(\widehat{B}_2 - B\right) &= n^{-\frac{1}{2}}\Sigma_D^{-1}\left[h + H\overline{K}_XB\right] + O_P\left(n^{-1}\right) - \left\{\left[\Sigma_D^{-1}\Sigma_X\left(I_p - n^{-\frac{1}{2}}\Sigma_X^{-1}H\right)\Sigma_X^{-1} + O_P\left(n^{-1}\right)\right]\right. \\ &\quad \times R_1'\left[\left(I_{r_1} + n^{-\frac{1}{2}}R_{1D}^{-1}R_1\Sigma_D^{-1}H\Sigma_X^{-1}R_1'\right)R_{1D}^{-1} + O_P\left(n^{-1}\right)\right] \\ &\quad \left.\times\left[n^{-\frac{1}{2}}R_1\Sigma_D^{-1}\left[h + H\overline{K}_XB\right]R_2\left(R_2'R_2\right)^{-1}R_2' + O_P\left(n^{-1}\right)\right]\right\} \\ &= n^{-\frac{1}{2}}\Sigma_D^{-1}\left[h + H\overline{K}_XB\right] + O_P\left(n^{-1}\right) - \left\{\left[\Sigma_D^{-1}R_1' - n^{-\frac{1}{2}}\Sigma_D^{-1}H\Sigma_X^{-1}R_1' + O_P\left(n^{-1}\right)\right]\right. \\ &\quad \times\left[R_{1D}^{-1} + n^{-\frac{1}{2}}R_{1D}^{-1}R_1\Sigma_D^{-1}H\Sigma_X^{-1}R_1'R_{1D}^{-1} + O_P\left(n^{-1}\right)\right] \\ &\quad \left.\times\left[n^{-\frac{1}{2}}R_1\Sigma_D^{-1}\left[h + H\overline{K}_XB\right]R_2\left(R_2'R_2\right)^{-1}R_2' + O_P\left(n^{-1}\right)\right]\right\} \\ &= n^{-\frac{1}{2}}\Sigma_D^{-1}\left[h + H\overline{K}_XB\right] + O_P\left(n^{-1}\right) - \left\{\left[\Sigma_D^{-1}R_1'R_{1D}^{-1} + n^{-\frac{1}{2}}\Sigma_D^{-1}R_1'R_{1D}^{-1}R_1\Sigma_D^{-1}H\Sigma_X^{-1}R_1'R_{1D}^{-1}\right.\right. \\ &\quad \left. - n^{-\frac{1}{2}}\Sigma_D^{-1}H\Sigma_X^{-1}R_1'R_{1D}^{-1} - n^{-1}\Sigma_D^{-1}H\Sigma_X^{-1}R_1'R_{1D}^{-1}R_1\Sigma_D^{-1}H\Sigma_X^{-1}R_1'R_{1D}^{-1} + O_P\left(n^{-1}\right)\right] \\ &\quad \left.\times\left[n^{-\frac{1}{2}}R_1\Sigma_D^{-1}\left(h + H\overline{K}_XB\right)R_2\left(R_2'R_2\right)^{-1}R_2' + O_P\left(n^{-1}\right)\right]\right\}. \end{aligned}$$

After simplification, we get

$$n^{\frac{1}{2}}\left(\widehat{B}_2 - B\right) = \Sigma_D^{-1}\left[h + H\overline{K}_XB\right] - \Sigma_D^{-1}R_1'R_{1D}^{-1}R_1\Sigma_D^{-1}\left[h + H\overline{K}_XB\right]R_2\left(R_2'R_2\right)^{-1}R_2' + O_P\left(n^{-\frac{1}{2}}\right). \tag{A.9}$$

Using (A.3), (A.4) and Lemma A.4, we get

$$\begin{aligned} \operatorname{vec}\left[n^{\frac{1}{2}}\left(\widehat{B}_2 - B\right)'\right] &= \left\{\left[\left(\Sigma_D^{-1}\otimes I_q\right) - \left(\left(\Sigma_D^{-1}R_1'R_{1D}^{-1}R_1\Sigma_D^{-1}\right)\otimes\left(R_2\left(R_2'R_2\right)^{-1}R_2'\right)\right)\right]\right. \\ &\quad \left.\times\left[\operatorname{vec}\left(h'\right) + \operatorname{vec}\left(B'\overline{K}_XH\right)\right]\right\} + O_P\left(n^{-\frac{1}{2}}\right). \end{aligned} \tag{A.10}$$

Using (3.19), we can write

$$\left(\widehat{B}_3 - B\right) = \left(\widehat{B}_1 - B\right) + S^{-1}R_1'\left[R_1S^{-1}R_1'\right]^{-1}\left(\theta - R_1\widehat{B}_1R_2\right)(R_2'R_2)^{-1}R_2'. \tag{A.11}$$

Consider

$$\begin{aligned}
n^{-1}\left[R_1S^{-1}R_1'\right]^{-1} &= \left[R_1\left(nS^{-1}\right)R_1'\right]^{-1} = \left[R_1\left\{\left[I_p - n^{-\frac{1}{2}}\Sigma_X^{-1}H\right]\Sigma_X^{-1} + O_P\left(n^{-1}\right)\right\}R_1'\right]^{-1}, \quad \text{using (A.1)}\\
&= \left[R_{1X} - n^{-\frac{1}{2}}R_1\Sigma_X^{-1}H\Sigma_X^{-1}R_1' + O_P\left(n^{-1}\right)\right]^{-1} \quad \text{where } R_{1X} = R_1\Sigma_X^{-1}R_1'\\
&= \left[R_{1X}\left\{I_{r_1} - \left(n^{-\frac{1}{2}}R_{1X}^{-1}R_1\Sigma_X^{-1}H\Sigma_X^{-1}R_1' - O_P\left(n^{-1}\right)\right)\right\}\right]^{-1}\\
&= \left[I_{r_1} - \left(n^{-\frac{1}{2}}R_{1X}^{-1}R_1\Sigma_X^{-1}H\Sigma_X^{-1}R_1' - O_P\left(n^{-1}\right)\right)\right]^{-1}R_{1X}^{-1}\\
&= \left[I_{r_1} + n^{-\frac{1}{2}}R_{1X}^{-1}R_1\Sigma_X^{-1}H\Sigma_X^{-1}R_1'\right]R_{1X}^{-1} + O_P\left(n^{-1}\right), \quad \text{using Lemma A.2.} \qquad (A.12)
\end{aligned}$$

Therefore using (A.1), (A.2), (A.8) and (A.12) in (A.11), we write

$$\begin{aligned}
\left(\widehat{B}_3 - B\right) &= n^{-\frac{1}{2}}\Sigma_D^{-1}\left[h + H\overline{K}_XB\right] + O_P\left(n^{-1}\right) - \left\{\left[\left(I_p - n^{-\frac{1}{2}}\Sigma_X^{-1}H\right)\Sigma_X^{-1} + O_P\left(n^{-1}\right)\right]\right.\\
&\quad \times R_1'\left[\left(I_{r_1} + n^{-\frac{1}{2}}R_{1X}^{-1}R_1\Sigma_X^{-1}H\Sigma_X^{-1}R_1'\right)R_{1X}^{-1} + O_P\left(n^{-1}\right)\right]\\
&\quad \left.\times\left[n^{-\frac{1}{2}}R_1\Sigma_D^{-1}\left[h + H\overline{K}_XB\right]R_2\left(R_2'R_2\right)^{-1}R_2' + O_P\left(n^{-1}\right)\right]\right\}\\
&= n^{-\frac{1}{2}}\Sigma_D^{-1}\left[h + H\overline{K}_XB\right] + O_P\left(n^{-1}\right) - \left\{\left[\Sigma_X^{-1}R_1' - n^{-\frac{1}{2}}\Sigma_X^{-1}H\Sigma_X^{-1}R_1' + O_P\left(n^{-1}\right)\right]\right.\\
&\quad \times\left[R_{1X}^{-1} + n^{-\frac{1}{2}}R_{1X}^{-1}R_1\Sigma_X^{-1}H\Sigma_X^{-1}R_1'R_{1X}^{-1} + O_P\left(n^{-1}\right)\right]\\
&\quad \left.\times\left[n^{-\frac{1}{2}}R_1\Sigma_D^{-1}\left[h + H\overline{K}_XB\right]R_2\left(R_2'R_2\right)^{-1}R_2' + O_P\left(n^{-1}\right)\right]\right\}\\
&= n^{-\frac{1}{2}}\Sigma_D^{-1}\left[h + H\overline{K}_XB\right] + O_P\left(n^{-1}\right) - \left\{\left[\Sigma_X^{-1}R_1'R_{1X}^{-1} + n^{-\frac{1}{2}}\Sigma_X^{-1}R_1'R_{1X}^{-1}R_1\Sigma_X^{-1}H\Sigma_X^{-1}R_1'R_{1X}^{-1}\right.\right.\\
&\quad \left. - n^{-\frac{1}{2}}\Sigma_X^{-1}H\Sigma_X^{-1}R_1'R_{1X}^{-1} - n^{-1}\Sigma_X^{-1}H\Sigma_X^{-1}R_1'R_{1X}^{-1}R_1\Sigma_X^{-1}H\Sigma_X^{-1}R_1'R_{1X}^{-1} + O_P\left(n^{-1}\right)\right]\\
&\quad \left.\times\left[n^{-\frac{1}{2}}R_1\Sigma_D^{-1}\left(h + H\overline{K}_XB\right)R_2\left(R_2'R_2\right)^{-1}R_2' + O_P\left(n^{-1}\right)\right]\right\}.
\end{aligned}$$

After simplification, we get

$$n^{\frac{1}{2}}\left(\widehat{B}_3 - B\right) = \Sigma_D^{-1}\left[h + H\overline{K}_XB\right] - \Sigma_X^{-1}R_1'R_{1X}^{-1}R_1\Sigma_D^{-1}\left[h + H\overline{K}_XB\right]R_2\left(R_2'R_2\right)^{-1}R_2' + O_P\left(n^{-\frac{1}{2}}\right). \qquad (A.13)$$

Using (A.3), (A.4) and Lemma A.4, this leads to

$$\begin{aligned}
\operatorname{vec}\left[n^{\frac{1}{2}}\left(\widehat{B}_3 - B\right)'\right] &= \left\{\left[\left(\Sigma_D^{-1}\otimes I_q\right) - \left(\left(\Sigma_X^{-1}R_1'R_{1X}^{-1}R_1\Sigma_D^{-1}\right)\otimes\left(R_2\left[R_2'R_2\right]^{-1}R_2'\right)\right)\right]\right.\\
&\qquad \left.\times\left[\operatorname{vec}\left(h'\right) + \operatorname{vec}\left(B'\overline{K}_XH\right)\right]\right\} + O_P\left(n^{-\frac{1}{2}}\right). \qquad (A.14)
\end{aligned}$$

Using (3.20), (A.2) and (A.8), we write

$$\begin{aligned}
\left(\widehat{B}_4 - B\right) &= \left(\widehat{B}_1 - B\right) + R_1'\left[R_1R_1'\right]^{-1}\left(\theta - R_1\widehat{B}_1R_2\right)\left(R_2'R_2\right)^{-1}R_2'\\
&= n^{-\frac{1}{2}}\Sigma_D^{-1}\left[h + H\overline{K}_XB\right] + O_P\left(n^{-1}\right) + R_1'\left[R_1R_1'\right]^{-1}\left[n^{-\frac{1}{2}}R_1\Sigma_D^{-1}\left[h + H\overline{K}_XB\right]R_2(R_2'R_2)^{-1}R_2' + O_P\left(n^{-1}\right)\right].
\end{aligned}$$

After simplification, we get

$$n^{\frac{1}{2}}\left(\widehat{B}_4 - B\right) = \Sigma_D^{-1}\left[h + H\overline{K}_XB\right] - R_1'\left(R_1R_1'\right)^{-1}R_1\Sigma_D^{-1}\left[h + H\overline{K}_XB\right]R_2\left(R_2'R_2\right)^{-1}R_2' + O_P\left(n^{-\frac{1}{2}}\right). \qquad (A.15)$$

Using (A.3), (A.4) and Lemma A.4, we get

$$\begin{aligned}
\operatorname{vec}\left[n^{\frac{1}{2}}\left(\widehat{B}_4 - B\right)'\right] &= \left\{\left[\left(\Sigma_D^{-1}\otimes I_q\right) - \left(\left(R_1'\left(R_1R_1'\right)^{-1}R_1\Sigma_D^{-1}\right)\otimes\left(R_2\left(R_2'R_2\right)^{-1}R_2'\right)\right)\right]\right.\\
&\qquad \left.\times\left[\operatorname{vec}\left(h'\right) + \operatorname{vec}\left(B'\overline{K}_XH\right)\right]\right\} + O_P\left(n^{-\frac{1}{2}}\right). \qquad (A.16)
\end{aligned}$$

**Proof of Theorem 4.1.** Let $M'_{(j)}$, $\Psi'_{(j)}$, $\Delta'_{(j)}$ and $E'_{(j)}$ be $j$th rows of $M$, $\Psi$, $\Delta$ and $E$ respectively.

Since,

$$\begin{aligned} h &= \left[n^{-\frac{1}{2}}\left[X'(E-\Delta B)\right]+n^{\frac{1}{2}}\sigma_\delta^2 B\right] \\ &= n^{-\frac{1}{2}}\left(X_{(1)},\ldots,X_{(n)}\right)\begin{bmatrix}E'_{(1)}\\ \vdots\\ E'_{(n)}\end{bmatrix} - n^{-\frac{1}{2}}\left(\left(X_{(1)},\ldots,X_{(n)}\right)\begin{bmatrix}\Delta'_{(1)}\\ \vdots\\ \Delta'_{(n)}\end{bmatrix} - n\sigma_\delta^2 I_p\right)B \\ &= n^{-\frac{1}{2}}\sum_{j=1}^{n}\left\{X_{(j)}E'_{(j)} - \left(X_{(j)}\Delta'_{(j)} - \sigma_\delta^2 I_p\right)B\right\}. \end{aligned}$$

After taking a transpose, we get

$$h' = n^{-\frac{1}{2}}\sum_{j=1}^{n}\left\{E_{(j)}M'_{(j)} + E_{(j)}\Psi'_{(j)} + E_{(j)}\Delta'_{(j)} - B'\Delta_{(j)}M'_{(j)} - B'\Delta_{(j)}\Psi'_{(j)} - B'\left(\Delta_{(j)}\Delta'_{(j)} - \sigma_\delta^2 I_p\right)\right\}.$$

Applying vec( ) operator and using Lemma A.4, this leads to

$$\begin{aligned} \mathrm{vec}\left(h'\right) &= n^{-\frac{1}{2}}\sum_{j=1}^{n}\left[M_{(j)}\otimes I_q, I_{pq}, I_{pq}, -M_{(j)}\otimes B', -I_p\otimes B', -I_p\otimes B'\right]\begin{bmatrix}\mathrm{vec}\left(E_{(j)}\right)\\ \mathrm{vec}\left(E_{(j)}\Psi'_{(j)}\right)\\ \mathrm{vec}\left(E_{(j)}\Delta'_{(j)}\right)\\ \mathrm{vec}\left(\Delta_{(j)}\right)\\ \mathrm{vec}\left(\Delta_{(j)}\Psi'_{(j)}\right)\\ \mathrm{vec}\left(\Delta_{(j)}\Delta'_{(j)} - \sigma_\delta^2 I_p\right)\end{bmatrix} \\ &= \sum_{j=1}^{n} C_{1j}W_{1j} \end{aligned} \tag{A.17}$$

where for $j = 1,\ldots,n$, the $C_{1j}$ of order $qp \times \left(2qp + 2p^2 + p + q\right)$ are matrices of constants and $W_{1j}$ of order $\left(2qp + 2p^2 + p + q\right) \times 1$ are independent and identically distributed random vectors.

Now, we write

$$\begin{aligned} H &= \left(n^{-\frac{1}{2}}\left(X'X\right) - n^{\frac{1}{2}}\Sigma_X\right) \\ &= n^{-\frac{1}{2}}\left[M'\Psi + M'\Delta + \Psi'M + \Psi'\Delta + \Delta'M + \Delta'\Psi\right] + \left[n^{-\frac{1}{2}}\Psi'\Psi - n^{\frac{1}{2}}\sigma_\psi^2 I_p\right] + \left[n^{-\frac{1}{2}}\Delta'\Delta - n^{\frac{1}{2}}\sigma_\delta^2 I_p\right]. \end{aligned} \tag{A.18}$$

Consider,

$$\begin{aligned} & n^{-\frac{1}{2}}\left[M'\Psi + M'\Delta + \Psi'M + \Psi'\Delta + \Delta'M + \Delta'\Psi\right] \\ &= n^{-\frac{1}{2}}\left[\left(M_{(1)},\ldots,M_{(n)}\right)\left\{\begin{bmatrix}\Psi'_{(1)}\\ \vdots\\ \Psi'_{(n)}\end{bmatrix} + \begin{bmatrix}\Delta'_{(1)}\\ \vdots\\ \Delta'_{(n)}\end{bmatrix}\right\} + \left\{\left[\Psi_{(1)},\ldots,\Psi_{(n)}\right] + \left[\Delta_{(1)},\ldots,\Delta_{(n)}\right]\right\}\begin{bmatrix}M'_{(1)}\\ \vdots\\ M'_{(n)}\end{bmatrix}\right. \\ &\quad \left. + \left[\Psi_{(1)},\ldots,\Psi_{(n)}\right]\begin{bmatrix}\Delta'_{(1)}\\ \vdots\\ \Delta'_{(n)}\end{bmatrix} + \left[\Delta_{(1)},\ldots,\Delta_{(n)}\right]\begin{bmatrix}\Psi'_{(1)}\\ \vdots\\ \Psi'_{(n)}\end{bmatrix}\right] \\ &= n^{-\frac{1}{2}}\sum_{j=1}^{n}\left\{M_{(j)}\left[\Psi'_{(j)} + \Delta'_{(j)}\right] + \left[\Psi_{(j)} + \Delta_{(j)}\right]M'_{(j)} + \left[\Psi_{(j)}\Delta'_{(j)} + \Delta_{(j)}\Psi'_{(j)}\right]\right\}. \end{aligned} \tag{A.19}$$

Now

$$\begin{aligned} & \left\{\left[n^{-\frac{1}{2}}\Psi'\Psi - n^{\frac{1}{2}}\sigma_\psi^2 I_p\right] + \left[n^{-\frac{1}{2}}\Delta'\Delta - n^{\frac{1}{2}}\sigma_\delta^2 I_p\right]\right\} \\ &= n^{-\frac{1}{2}}\left[\Psi_{(1)},\ldots,\Psi_{(n)}\right]\begin{bmatrix}\Psi'_{(1)}\\ \vdots\\ \Psi'_{(n)}\end{bmatrix} - n^{\frac{1}{2}}\sigma_\psi^2 I_p + n^{-\frac{1}{2}}\left[\Delta_{(1)},\ldots,\Delta_{(n)}\right]\begin{bmatrix}\Delta'_{(1)}\\ \vdots\\ \Delta'_{(n)}\end{bmatrix} - n^{\frac{1}{2}}\sigma_\delta^2 I_p \\ &= n^{-\frac{1}{2}}\sum_{j=1}^{n}\left[\left(\Psi_{(j)}\Psi'_{(j)} - \sigma_\psi^2 I_p\right) + \left(\Delta_{(j)}\Delta'_{(j)} - \sigma_\delta^2 I_p\right)\right]. \end{aligned} \tag{A.20}$$

Using (A.19) and (A.20) in (A.18) and Lemma A.4 with $C = I_p$, we get

$$\begin{aligned}\operatorname{vec}\left(B'\overline{K}_X H\right) &= \left(I_p \otimes B'\overline{K}_X\right)\operatorname{vec}(H)\\ &= \sum_{j=1}^{n} n^{-\frac{1}{2}}\left(I_p \otimes B'\overline{K}_X\right)\left[\left(I_p \otimes M_{(j)}\right), \left(M_{(j)} \otimes I_p\right), I_{p^2}, I_{p^2}, I_{p^2}\right] \times \begin{bmatrix}\operatorname{vec}\left(\Psi'_{(j)} + \Delta'_{(j)}\right)\\ \left(\Psi_{(j)} + \Delta_{(j)}\right)\\ \operatorname{vec}\left(\Delta_{(j)}\Psi'_{(j)} + \Psi_{(j)}\Delta'_{(j)}\right)\\ \operatorname{vec}\left(\Psi_{(j)}\Psi'_{(j)} - \sigma_\psi^2 I_p\right)\\ \operatorname{vec}\left(\Delta_{(j)}\Delta'_{(j)} - \sigma_\delta^2 I_p\right)\end{bmatrix}\\ &= \sum_{j=1}^{n} C_{2j}W_{2j} \qquad (A.21)\end{aligned}$$

where for $j = 1, \ldots, n$, the $C_{2j}$ of order $qp \times \left(3p^2 + 2p\right)$ are matrices of constants and $W_{2j}$ of order $\left(3p^2 + 2p\right) \times 1$ are independent and identically distributed random vectors respectively.

Thus using (A.21) and (A.17), we can write

$$\left[\operatorname{vec}\left(h'\right) + \operatorname{vec}\left(B'\overline{K}_X H\right)\right] = \sum_{j=1}^{n}\left[C_{1j}, C_{2j}\right]\begin{bmatrix}W_{1j}\\ W_{2j}\end{bmatrix} = \sum_{j=1}^{n} U_j W_j. \qquad (A.22)$$

Since assumptions 1–5 imply that, $\omega(n)U_j = n^{\frac{1}{2}}\left[C_{1j}, C_{2j}\right] \to L$ (constant) as $n \to \infty$ and $W_j$ for $j = 1, \ldots, n$ are independent and identically distributed random vectors with mean zero. Hence the Central limit theorem (Lemma A.3) gives that

$$\left[\operatorname{vec}\left(h'\right) + \operatorname{vec}\left(B'\overline{K}_X H\right)\right] \xrightarrow{d} N_{qp}\left(O_{qp\times 1}, \Lambda_{qp\times qp}\right) \qquad (A.23)$$

where $\Lambda = \lim_{n\to\infty} E\left\{\left[\operatorname{vec}\left(h'\right) + \operatorname{vec}\left(B'\overline{K}_X H\right)\right]\left[\operatorname{vec}\left(h'\right) + \operatorname{vec}\left(B'\overline{K}_X H\right)\right]'\right\}$ is the variance-covariance matrix. It is easy to see that

$$\Lambda = \lim_{n\to\infty} E\left\{\operatorname{vec}\left(h'\right)\left(\operatorname{vec}\left(h'\right)\right)' + \operatorname{vec}\left(h'\right)\left(\operatorname{vec}\left(B'\overline{K}_X H\right)\right)' + \operatorname{vec}\left(B'\overline{K}_X H\right)\left(\operatorname{vec}\left(h'\right)\right)' + \operatorname{vec}\left(B'\overline{K}_X H\right)\left(\operatorname{vec}\left(B'\overline{K}_X H\right)\right)'\right\}.$$

Let $e_j$ be a $(p \times 1)$ vector with $j$th element as one and all other elements as zero. So for any matrix $A_{q\times p}$, $Ae_j$ gives the $j$th column of $A$. Thus we can write

$$\operatorname{vec}(A)\left(\operatorname{vec}(A)\right)' = \begin{bmatrix}Ae_1\\ \vdots\\ Ae_p\end{bmatrix}\begin{bmatrix}e_1'A' & \cdots & e_p'A'\end{bmatrix}$$

where the $(s, t)$th partitioned matrix is given by $Ae_s e_t' A'$, for $s, t = 1, \ldots, p$. Hence

$$\Lambda_{st} = \lim_{n\to\infty} E\left\{h'e_s e_t' h + h'e_s e_t' H\overline{K}_X B + B'\overline{K}_X H e_s e_t' h + B'\overline{K}_X H e_s e_t' H\overline{K}_X B\right\} \qquad (A.24)$$

is the $(s, t)$th partitioned matrix of $\Lambda$.

Using (2.2), (2.3) and definition of $h$, we write

$$\begin{aligned}E\left\{h'e_s e_t' h\right\} &= n^{-1}E\Big\{\left(E'\left(M + \Psi + \Delta\right) - B'\Delta'\left(M + \Psi + \Delta\right) + n\sigma_\delta^2 B'\right) \times e_s e_t'\\ &\quad \times \left(\left(M' + \Psi' + \Delta'\right)E - \left(M' + \Psi' + \Delta'\right)\Delta B + n\sigma_\delta^2 B\right)\Big\}.\end{aligned}$$

Using assumptions 1–5 and methods for obtaining expectations of product of stochastic matrices (Ref. [21]), we get

$$\begin{aligned}E\left\{h'e_s e_t' h\right\} &= \left\{\operatorname{tr}\left(n^{-1}Me_s e_t' M'\right) + \left(\sigma_\delta^2 + \sigma_\psi^2\right)\operatorname{tr}\left(e_s e_t'\right)\right\}\left[\Sigma_o + \sigma_\delta^2 B'B\right] + \sigma_\delta^4 B'e_t e_s' B\\ &\quad + \gamma_{1\delta}\sigma_\delta^3 B'\left[\left(a_p a_n' n^{-1}Me_s e_t' * I_p\right) + \left(a_p a_n' n^{-1}Me_s e_t' * I_p\right)'\right]B + \gamma_{2\delta}\sigma_\delta^4 B'\left(e_s e_t' * I_p\right)B.\end{aligned}$$

Writing the expressions for other expectations on similar lines, we get from (A.24)

$$\begin{aligned}\Lambda_{st} &= \left(\sigma_\delta^2 + \sigma_\psi^2\right)\left\{\operatorname{tr}\left(e_s e_t'\right)\left[\Sigma_o + \sigma_\delta^2 B'KB\right] + B'\overline{K}e_t e_s'\sigma_M\sigma_M'\overline{K}B\right\} + \sigma_\delta^2 B'e_t e_s'\left(\sigma_\psi^2\overline{K} - \sigma_\delta^2 K\right)B\\ &\quad + e_t'\sigma_M\sigma_M' e_s\left(\Sigma_o + \sigma_\delta^2 B'K^2 B + \sigma_\psi^2 B'\overline{K}^2 B\right) + N\left(\gamma\right).\end{aligned}$$

$N(\gamma)$, a function of coefficient of skewness and kurtosis is given by

$$\begin{aligned}N(\gamma) = {} & B'K\left\{\gamma_{1\delta}\sigma_\delta^3\left(\left(a_p\sigma_M'e_se_t' * I_p\right) + \left(a_p\sigma_M'e_se_t' * I_p\right)'\right) + \gamma_{2\delta}\sigma_\delta^4\left(e_se_t' * I_p\right)\right\}KB \\ & + B'\overline{K}\left\{\gamma_{1\psi}\sigma_\psi^3\left(\left(a_p\sigma_M'e_se_t' * I_p\right) + \left(a_p\sigma_M'e_se_t' * I_p\right)'\right) + \gamma_{2\psi}\sigma_\psi^4\left(e_se_t' * I_p\right)\right\}\overline{K}B \\ & + B'\left\{\left(\gamma_{1\psi}\sigma_\psi^3\overline{K} - \gamma_{1\delta}\sigma_\delta^3 K\right)\left(e_se_t' * I_p\right)a_p\sigma_M'\overline{K} + \overline{K}\sigma_M a_p'\left(e_se_t' * I_p\right)\left(\gamma_{1\psi}\sigma_\psi^3\overline{K} - \gamma_{1\delta}\sigma_\delta^3 K\right)\right\}B.\end{aligned}$$

From Eqs. (A.4), (A.10), (A.14) and (A.16), it follows that the asymptotic distribution of vec $\left[n^{\frac{1}{2}}\left(\widehat{B}_j - B\right)'\right]$, for $j = 1, 2, 3$ and 4 is same as the asymptotic distribution of $\left[\text{vec}\left(h'\right) + \text{vec}\left(B'\overline{K}_X H\right)\right]$. Thus (A.23) implies that

$$\text{vec}\left[n^{\frac{1}{2}}\left(\widehat{B}_j - B\right)'\right] \xrightarrow{d} N_{qp}\left(O_{qp\times 1}, A_j\left(\Lambda_{qp\times qp}\right)A_j'\right) \tag{A.25}$$

where

$$\begin{aligned}A_1 &= \left((\Sigma K)^{-1}\otimes I_q\right); \\ A_2 &= A_1 - \left(\left((\Sigma K)^{-1}R_1'\left[R_1(\Sigma K)^{-1}R_1'\right]^{-1}R_1(\Sigma K)^{-1}\right)\otimes\left(R_2\left[R_2'R_2\right]^{-1}R_2'\right)\right); \\ A_3 &= A_1 - \left(\left(\Sigma^{-1}R_1'\left[R_1\Sigma^{-1}R_1'\right]^{-1}R_1(\Sigma K)^{-1}\right)\otimes\left(R_2\left[R_2'R_2\right]^{-1}R_2'\right)\right); \\ A_4 &= A_1 - \left(\left(R_1'\left[R_1R_1'\right]^{-1}R_1(\Sigma K)^{-1}\right)\otimes\left(R_2\left[R_2'R_2\right]^{-1}R_2'\right)\right).\end{aligned}$$

So now using definitions (A.3) and (A.25), it can be concluded that the random matrices $\left[n^{\frac{1}{2}}\left(\widehat{B}_j - B\right)\right]$, $j = 1, 2, 3, 4$ are distributed as Matrix-variate Normal, that is,

$$\left[n^{\frac{1}{2}}\left(\widehat{B}_j - B\right)\right] \xrightarrow{d} MN_{p,q}\left(O_{p\times q}, A_j\left(\Lambda_{qp\times qp}\right)A_j'\right). \quad \square$$

## References

[1] C.L. Cheng, J.W. Van Ness, Statistical Regression with Measurement Errors, Arnold Publishers, 1999.
[2] K.T. Dinh, T.T. Nguyen, A characterization of matrix variate normal distribution, Int. J. Math. Math. Sci. 17 (2) (1994) 341–346.
[3] G.R. Dolby, The ultrastructural relation: a synthesis of the functional and structural relations, Biometrika 63 (1976) 39–50.
[4] W.A. Fuller, Introduction to Statistical Time Series, John Wiley, 1976.
[5] W.A. Fuller, Measurement Error Models, John Wiley, 1987.
[6] L.J. Gleser, The importance of assessing measurement reliability in multivariate regression, J. Amer. Statist. Assoc. 87 (419) (1992) 696–707.
[7] L.J. Gleser, Estimators of slopes in linear errors-in-variables regression models when the predictors have known reliability matrix, Statist. Probab. Lett. 17 (1993) 113–121.
[8] A.K. Gupta, D.K. Nagar, Matrix Variate Distributions, CRC Press, 2000.
[9] N.L. Johnson, S. Kotz, Distributions in Statistics: Continuous Multivariate Distributions, John Wiley, 1972.
[10] K. Kim, N.H. Timm, Univariate and Multivariate General Linear Models: Theory and Applications with SAS, 2nd ed., CRC Press, 2006.
[11] E. Malinvaud, Statistical Methods of Econometrics, North-Holland Publishing Co., Amsterdam, 1966.
[12] T.T. Nguyen, A note on matrix variate normal distribution, J. Multivariate Anal. 60 (1997) 148–153.
[13] C.R. Rao, Linear Statistical Inference and its Applications, 2nd ed., Wiley, 1973.
[14] C.R. Rao, M.B. Rao, Matrix Algebra and its Applications to Statistics and Econometrics, World Scientific, 1998.
[15] C.R. Rao, H. Toutenburg, Shalabh, C. Heumann, Linear Models and Generalizations: Least Squares and Alternatives, 3rd ed., Springer, 2007.
[16] D.A.J. Ryan, J.J. Hubert, E.M. Carter, J.B. Sprague, J. Parrott, Reduced rank multivariate regression approach to aquatic joint toxicity experiments, Biometrics 48 (1992) 155–162.
[17] H. Schneeweiss, Consistent estimation of a regression with errors in the variables, Metrika 23 (1976) 101–115.
[18] H. Schneeweiss, Note on a linear model with errors in the variables and with trend, Statist. Papers 32 (1991) 261–264.
[19] Shalabh, Consistent estimation of coefficients in measurement error models under non-normality, J. Multivariate Anal. 86 (2) (2003) 227–241.
[20] Shalabh, G. Garg, N. Misra, Restricted regression estimation in measurement error models, Comput. Statist. Data Anal. 52 (2007) 1149–1166.
[21] V.K. Srivastava, R. Tiwari, Evaluation of expectation of product of stochastic matrices, Scand. J. Statist. 3 (1976) 135–138.
[22] N.H. Timm, Applied Multivariate Analysis, Springer, 2002.