Journal of Statistical Research *ISSN 0256 - 422 X*
2011, Vol. 45, No. 2, pp. 139-154
Bangladesh

SIMULTANEOUS PREDICTIONS UNDER EXACT RESTRICTIONS IN ULTRASTRUCTURAL MODEL

GAURAV GARG
Indian Institute of Management Lucknow, Lucknow - 226 013, Uttar Pradesh, INDIA
Email: ggarg@iiml.ac.in, ggarg31@gmail.com

SHALABH
Department of Mathematics and Statistics, Indian Institute of Technology Kanpur, Kanpur-208016, India
Email: shalab@iitk.ac.in

SUMMARY

Present article discusses the problem of simultaneous prediction of actual and average values of study variable in an ultrastructural measurement error model. Some prior information is also available on regression coefficients of the model in terms of exact linear restrictions. Some predictors are obtained and their properties are analyzed. The effect of departure from normality of the distributions of measurement errors is also studied.

Keywords and phrases: Exact linear restrictions, instrumental variables, measurement errors, ultrastructural model.

AMS Classification: 62J05, 62H12, 62P20.

1 Introduction

In regression analysis, we usually predict either the actual value of study variable or the average value of study variable. It depends on the situation what we want to predict - the actual value or the average value. In some situations, we may like to predict the both simultaneously. For example, a long term investor may like to predict the average price of a particular stock in a long run, while a short term investor may like to predict the actual price of the stock, say, next year. Shalabh (1995) proposed a target function for the simultaneous prediction of actual and average values of the study variable. We use this target function for simultaneous prediction of actual and average value of study variable in our model.

We consider a multiple linear regression model where the covariates are observed with measurement errors. Presence of measurement errors in the observations is very common and obvious. However, most of the times this fact is ignored and the statistical results which are obtained for no measurement error situations are used. This leads to some wrong conclusions and may affect the proceeds of analysts badly. The model, we consider here is called as ultrastructural measurement error model and is a generalized measurement error model first proposed by Dolby (1976).

In measurement error models, some additional information is required for consistent estimation of parameters. Cheng and Van Ness (1999) and Fuller (1987) provide a detailed discussion on these additional information about unknown parameters. In the present article, we use the common variance of measurement errors for consistent estimation.

Sometimes, we are provided with some prior information on unknown regression coefficients. This prior information may be coming from some similar analysis conducted in past or from the experience of the analyst. We consider here that this prior information is available in the form of exact linear restrictions on regression coefficients. It is also common to assume normality of the random terms in the model. The observed data may or may not have normal distribution. When the observations are not normally distributed, the conclusions may be incorrect. We do not assume any specific form of the distributions of the random terms in the model. We only assume the finite existence of first four moments. Under this setup, Shalabh, Garg, and Misra (2007, 2009) provides the consistent of regression coefficients. In the present article, we take the topic farther and obtain the predictors to predict the actual and average values of study variable simultaneously. The effect of departure from the normal distribution on the efficiency properties of obtained predictors is studied. We obtain one unrestricted predictor and two restricted predictors and compare their efficiencies through simulation. Asymptotic mean squared errors of the proposed predictors are obtained and analyzed. A Monte-Carlo simulation experiment is conducted to study the sample properties of the estimators and the effect of departure from normality of the measurement errors on them is also studied.

In Section 2 of this paper, we present the ultrastructural measurement error model and the required assumptions along with the target functions and linear restrictions on regression coefficients. In Section 3, we obtain the predictors of target function and their asymptotic properties are discussed in Section 4. Results of simulation study are presented in Section 5 followed by some concluding remarks in Section 6.

2 Model and Target Function

We consider that the study variable y and the covariates $\xi_1, \xi_2, \ldots, \xi_p$ have the following relationship:

$$y = \xi_1\beta_1 + \xi_2\beta_2 + \cdots + \xi_p\beta_p + \epsilon,$$

where $\beta_1, \beta_2, \ldots, \beta_p$, are regression coefficients. For a sample of size n, we write

$$\boldsymbol{y} = \Xi\boldsymbol{\beta} + \boldsymbol{\epsilon}, \tag{2.1}$$

where $\boldsymbol{y}$ is the $n \times 1$ vector of n observations on study variable y and $\Xi = (\xi_{ij})$ is the $n \times p$ matrix of n observations on true covariates $\xi_1, \xi_2, \ldots, \xi_p$ and $\boldsymbol{\beta}$ is $p \times 1$ vector of regression coefficients, $\boldsymbol{\epsilon} = (\epsilon_1, \epsilon_2, \ldots, \epsilon_n)'$ is the vector of equation error. Due to the presence of measurement errors, we can not observe Ξ. Instead, we observe X as

$$X = \Xi + \Delta, \tag{2.2}$$

where $\boldsymbol{y}$ is a $n \times 1$ vector of observed values on true study variable, X is a $n \times p$ matrix of observed covariates, and $\Delta = (\boldsymbol{\delta}_{ij})$ is a $n \times p$ matrix of measurement errors. We further consider that $\xi_{ij} = \mu_{ij} + \phi_{ij}$ $(i = 1, 2, \ldots, n, j = 1, 2, \ldots, p.)$, where μ_{ij} are unknown parameters where ϕ_{ij} are independent and identically distributed (i.i.d.) random variables with mean zero. Thus, we can express,

$$\Xi = M + \Phi, \tag{2.3}$$

where $M = (\mu_{ij})$ and $\Phi = (\phi_{ij})$ are $n \times p$ matrices. Equations (2.1)-(2.3) specify a multivariate ultrastructural model, Shalabh, Garg and Misra (2007, 2009). The ultrastructural model is a synthesis of functional and structural forms of measurement error model. When all the row vectors of M are identical, then this implies that rows of X are i.i.d. random vectors and we get the specification of structural form of measurement error model. When Φ is identically equal to a null matrix implying that the matrix X is fixed but is measured with error, then we obtain the specification of functional form of measurement error model. When both Δ and Φ are identically equal to a null matrix, we get the specification of classical regression model without any measurement error.

ϵ_i, $(i = 1, 2, \ldots, n)$ are assumed to be i.i.d. random variables with mean 0 and variance σ_ϵ^2. Similarly, δ_{ij}, $(i = 1, 2, \ldots, n;\ j = 1, 2, \ldots, p)$ are assumed to be i.i.d. random variables with mean 0, variance σ_δ^2, third moment $\gamma_{1\delta}\sigma_\delta^3$ and fourth moment $(\gamma_{2\delta} + 3)\sigma_\delta^4$. Also, ϕ_{ij}, $(i = 1, 2, \ldots, n;\ j = 1, 2, \ldots, p)$ are assumed to be i.i.d. random variables with mean 0, variance σ_ϕ^2, third moment $\gamma_{1\phi}\sigma_\phi^3$ and fourth moment $(\gamma_{2\phi} + 3)\sigma_\phi^4$. Here, γ_1 and γ_2 represent the coefficients of skewness and kurtosis, respectively. We also assume that ϵ_i, δ_{ij} and ϕ_{ij} are independent of each other for all i and j and nth row of matrix M converges to $\boldsymbol{\sigma}'_\mu$.

Suppose that we wish to predict the value of study variable y_0 for observed values on covariates given by $\boldsymbol{x}_0$. Clearly, $\boldsymbol{x}_0$ is a $p \times 1$ vector of observed values with measurement errors $\boldsymbol{\delta}_0$ for true covariates given by $\boldsymbol{\xi}_0$, i.e., $\boldsymbol{x}_0 = \boldsymbol{\xi}_0 + \boldsymbol{\delta}_0$. Elements of $\boldsymbol{\delta}_0$ have the same properties as possessed by the elements of Δ. Also, $\boldsymbol{\xi}_0 = \boldsymbol{\mu}_0 + \boldsymbol{\phi}_0$, where $\boldsymbol{\mu}_0$ is unknown constant vector and elements of $\boldsymbol{\phi}_0$ have the same properties as possessed by the elements of Φ. We have the relationship $y_0 = \boldsymbol{x}_0'\boldsymbol{\beta} + (\epsilon_0 - \boldsymbol{\delta}_0'\boldsymbol{\beta})$. Here ϵ_0 is the equation error distributed with mean 0 and variance σ_ϵ^2. For some estimate $\hat{\boldsymbol{\beta}}$ of unknown $\boldsymbol{\beta}$, the prediction of actual value y_0 is given by $\hat{y}_0 = \boldsymbol{x}_0'\hat{\boldsymbol{\beta}}$. In some situations, we like to predict the average value of y_0 in place of actual value. The average value of y_0 is given by $E(y_0|\boldsymbol{x}_0) = E(\boldsymbol{x}_0'\boldsymbol{\beta} + (\epsilon_0 - \boldsymbol{\delta}_0'\boldsymbol{\beta})|\boldsymbol{x}_0) = \boldsymbol{x}_0'\boldsymbol{\beta}$. Clearly, the average value $E(y_0|\boldsymbol{x}_0)$ is also estimated by $\boldsymbol{x}_0'\hat{\boldsymbol{\beta}}$.

For the situations when it is required to predict both the actual value and average value of unknown y_0 for given $\boldsymbol{x}_0$, we define the target function:

$$T \equiv T(y_0) = \lambda y_0 + (1 - \lambda)E(y_0|\boldsymbol{x}_0), \tag{2.4}$$

where $0 \leq \lambda \leq 1$ is a real number specifying the weight assigned to the prediction of actual and average values of y_0. For details, see Shalabh (1995) and Toutenburg and Shalabh (1997). Clearly, for some estimate $\hat{\boldsymbol{\beta}}$ of unknown $\boldsymbol{\beta}$, the estimate of $T(y_0)$ is given by $\hat{T} = \boldsymbol{x}_0'\hat{\boldsymbol{\beta}}$.

Further, we consider that some prior information about unknown regression coefficients $\beta_1, \beta_2, \ldots, \beta_p$, is provided in terms of exact linear restrictions. This information is represented as

$$\boldsymbol{r} = R\boldsymbol{\beta}, \tag{2.5}$$

where $\boldsymbol{r}$ is a $J \times 1$ known vector and R is a $J \times p$ known full row rank matrix.

We aim to obtain predictor of target function under the exact linear restrictions. In the next section we see the inadequacy of the available predictors and provide the more suitable predictors.

3 Prediction

The ordinary least squares estimator (OLSE) $\boldsymbol{b}$ of $\boldsymbol{\beta}$ and the restricted least squares estimator (RLSE) $\boldsymbol{b}_R$ of $\boldsymbol{\beta}$ under restrictions (2.5) in a classical regression model without measurement errors are given by

$$\boldsymbol{b} = (X'X)^{-1}X'\boldsymbol{y} \tag{3.1}$$

and

$$\boldsymbol{b}_R = \boldsymbol{b} - \{I_p - f_R(X'X)\}(\boldsymbol{b} - \boldsymbol{\beta}) \tag{3.2}$$

respectively, where the function $f_R : \mathbb{R}^{p\times p} \to \mathbb{R}^{p\times p}$ is defined as

$$f_R(U) = I_p - U^{-1}R'(RU^{-1}R')^{-1}R, \qquad U \in \mathbb{R}^{p\times p}. \tag{3.3}$$

It can be proved that $\operatorname*{plim}_{n\to\infty} \boldsymbol{b} \neq \boldsymbol{\beta}$ and $\operatorname*{plim}_{n\to\infty} \boldsymbol{b}_R \neq \boldsymbol{\beta}$ as well. For proof, see Shalabh, Garg and Misra (2007, 2009). Thus both $\boldsymbol{b}$ and $\boldsymbol{b}_R$ are inconsistent for $\boldsymbol{\beta}$ under the model (2.1)–(2.3). The OLSE $\boldsymbol{b}$ does not satisfy the restrictions (2.5), i.e., $R\boldsymbol{b} \neq \boldsymbol{r}$ while RLSE $\boldsymbol{b}_R$ satisfies the given restrictions, i.e., $R\boldsymbol{b}_R = \boldsymbol{r}$. Therefore the predictors $\boldsymbol{x}_0'\boldsymbol{b}$ as well as $\boldsymbol{x}_0'\boldsymbol{b}_R$ are not suitable to predict $T(y_0)$.

It is well known that measurement error models are unidentifiable and some additional information about unknown parameters is required for consistent estimation of regression coefficients. See Cheng and Van-Ness (1999) and Fuller (1987) for details. Here we assume that common variance σ_δ^2 of measurement errors δ_{ij} is known. Using σ_δ^2, Shalabh, Garg and Misra (2007) obtained some consistent estimators of $\boldsymbol{\beta}$ under the exact linear restrictions given by (2.5). In the present article, we use these estimators to obtain the predictors of target function $T(y_0)$.

A consistent estimator of $\boldsymbol{\beta}$ in the model (2.1) - (2.3) for known σ_δ^2 is given by

$$\boldsymbol{b}_\delta^{(1)} = (I_p - n\sigma_\delta^2 S^{-1})^{-1}\boldsymbol{b}; \tag{3.4}$$

see Shalabh (2003). Although, the estimator $\boldsymbol{b}_\delta^{(1)}$ is consistent for estimating $\boldsymbol{\beta}$, it does not satisfy the given linear restrictions (2.5). Estimate of target function using $\boldsymbol{b}_\delta^{(1)}$ is

$$\hat{T}_1 = \boldsymbol{x}_0'\boldsymbol{b}_\delta^{(1)}. \tag{3.5}$$

We wish to study the effect of linear restrictions on the predictors. Therefore, we wish to compare the restricted predictor with unrestricted predictors obtained later in this section.

In order to obtain an estimator of $\boldsymbol{\beta}$, that is consistent as well as satisfies the given linear restrictions (2.5), inconsistent estimator $\boldsymbol{b}$ in $\boldsymbol{b}_R$ is replaced by the consistent estimator $\boldsymbol{b}_\delta^{(1)}$. This yields the following estimator

$$\boldsymbol{b}_\delta^{(2)} \quad = \quad \boldsymbol{b}_\delta^{(1)} + S^{-1}R'(RS^{-1}R')^{-1}(\boldsymbol{r} - R\boldsymbol{b}_\delta^{(1)}); \tag{3.6}$$

see Shalabh, Garg and Misra (2007, 2009). Since this estimator is consistent for estimating $\boldsymbol{\beta}$ and also satisfies the restrictions (2.5), it is suitable for predicting the target function $T(y_0)$. Such predictor of target function is

$$\hat{T}_2 = \boldsymbol{x}_0' \boldsymbol{b}_\delta^{(2)}. \tag{3.7}$$

Another restricted consistent estimator of $\boldsymbol{\beta}$ obtained by Shalbh, Garg and Misra (2007, 2009) is

$$\boldsymbol{b}_\delta^{(3)} \quad = \quad \{I_p - n\sigma_\delta^2 f_R(S) S^{-1}\}^{-1} \boldsymbol{b}_R. \tag{3.8}$$

This estimator is obtained by adjusting the inconsistency in $\boldsymbol{b}_R$. We obtain another predictor of the target function using $\boldsymbol{b}_\delta^{(3)}$, which is

$$\hat{T}_3 = \boldsymbol{x}_0' \boldsymbol{b}_\delta^{(3)}. \tag{3.9}$$

Thus, we have three predictors $\hat{T}_1$, $\hat{T}_2$ and $\hat{T}_3$ of the target function (2.4). In the next section, we analyze the efficiencies of these predictors.

4 Asymptotic Properties of Predictors

Following leamma is helpful in studying the asymptotic efficiencies of the predictors $\hat{T}_1$, $\hat{T}_2$ and $\hat{T}_3$.

Lemma 4.1. *For $l = 1, 2, 3$, the asymptotic distributions of $\sqrt{n}(\boldsymbol{b}_\delta^{(l)} - \boldsymbol{\beta})$ are p-variate normal with mean vector* **0** *and covariance matrices $A_l \Omega A_l{}'$ where*

$$\left.\begin{aligned}
A_1 &= (\boldsymbol{\sigma}_\mu \boldsymbol{\sigma}_\mu' - \sigma_\delta^2 I_p)^{-1} \\
A_2 &= f_R(\boldsymbol{\sigma}_\mu \boldsymbol{\sigma}_\mu')(\boldsymbol{\sigma}_\mu \boldsymbol{\sigma}_\mu' - \sigma_\delta^2 I_p)^{-1} \\
A_3 &= f_R(\boldsymbol{\sigma}_\mu \boldsymbol{\sigma}_\mu' - \sigma_\delta^2 I_p)(\boldsymbol{\sigma}_\mu \boldsymbol{\sigma}_\mu' - \sigma_\delta^2 I_p)^{-1} \\
\Omega &= (\sigma_\epsilon^2 + \sigma_\delta^2(\boldsymbol{\beta}'\boldsymbol{\beta}))\Sigma + \sigma_\delta^4 \boldsymbol{\beta}\boldsymbol{\beta}' + \gamma_{1\delta}\sigma_\delta^3\{f(\boldsymbol{\sigma}_\mu \boldsymbol{e}_p', \boldsymbol{\beta}\boldsymbol{\beta}') \\
&\qquad + \big(f(\boldsymbol{\sigma}_\mu \boldsymbol{e}_p', \boldsymbol{\beta}\boldsymbol{\beta}')\big)'\} + \gamma_{2\delta}\sigma_\delta^4 f(I_p, \boldsymbol{\beta}\boldsymbol{\beta}'),
\end{aligned}\right\} , \tag{4.1}$$

$\boldsymbol{e}_p$ is a $p \times 1$ vector of 1's and $f_R(\cdot)$ is defined in (3.3).

Proof See Shalabh, Garg and Misra (2007). □

Following theorem gives the asymptotic mean squared errors of the predictors $\hat{T}_1$, $\hat{T}_2$ and $\hat{T}_3$.

Theorem 1. *Conditional asymptotic mean squared errors of $\hat{T}_l$ $(l = 1, 2, 3)$ for known $\boldsymbol{x}_0$ are*

$$\lim_{n\to\infty} \text{MSE}(\hat{T}_l | \boldsymbol{x}_0) = \lambda^2 \left(\sigma_\epsilon^2 + \sigma_\delta^2 \boldsymbol{\beta}'\boldsymbol{\beta}\right). \tag{4.2}$$

Proof We have from (2.4),

$$\begin{aligned}
T &= \lambda y_0 + (1-\lambda)\boldsymbol{x}_0'\boldsymbol{\beta} \\
&= \lambda\left(\boldsymbol{x}_0'\boldsymbol{\beta} + (\epsilon_0 - \boldsymbol{\delta}_0'\boldsymbol{\beta})\right) + (1-\lambda)\boldsymbol{x}_0'\boldsymbol{\beta} \\
&= \boldsymbol{x}_0'\boldsymbol{\beta} + \lambda\left(\epsilon_0 - \boldsymbol{\delta}_0'\boldsymbol{\beta}\right)
\end{aligned}$$

For $l = 1, 2, 3$, we write

$$\hat{T}_l - T = \boldsymbol{x}_0'(\boldsymbol{b}_\delta^{(l)} - \boldsymbol{\beta}) + \lambda(\epsilon_0 - \boldsymbol{\delta}_0'\boldsymbol{\beta})$$

So,

$$\begin{aligned}(\hat{T}_l - T)^2 &= [\boldsymbol{x}_0'(\boldsymbol{b}_\delta^{(l)} - \boldsymbol{\beta})]^2 + \lambda^2(\epsilon_0 - \boldsymbol{\delta}_0'\boldsymbol{\beta})^2 \\ &= \boldsymbol{x}_0'(\boldsymbol{b}_\delta^{(l)} - \boldsymbol{\beta})(\boldsymbol{b}_\delta^{(l)} - \boldsymbol{\beta})'\boldsymbol{x}_0 + \lambda^2(\epsilon_0 - \boldsymbol{\delta}_0'\boldsymbol{\beta})^2 \\ &= \text{trace}[\boldsymbol{x}_0'(\boldsymbol{b}_\delta^{(l)} - \boldsymbol{\beta})(\boldsymbol{b}_\delta^{(l)} - \boldsymbol{\beta})'\boldsymbol{x}_0] + \lambda^2(\epsilon_0 - \boldsymbol{\delta}_0'\boldsymbol{\beta})^2 \\ &= \text{trace}[(\boldsymbol{x}_0\boldsymbol{x}_0')(\boldsymbol{b}_\delta^{(l)} - \boldsymbol{\beta})(\boldsymbol{b}_\delta^{(l)} - \boldsymbol{\beta})'] + \lambda^2(\epsilon_0 - \boldsymbol{\delta}_0'\boldsymbol{\beta})^2\end{aligned}$$

Taking conditional expectation for known $\boldsymbol{x}_0$, we have

$$\begin{aligned}E[(\hat{T}_l - T)^2|\boldsymbol{x}_0] &= \text{trace}[E\{(\boldsymbol{x}_0\boldsymbol{x}_0')(\boldsymbol{b}_\delta^{(l)} - \boldsymbol{\beta})(\boldsymbol{b}_\delta^{(l)} - \boldsymbol{\beta})'|\boldsymbol{x}_0\}] + \lambda^2\left(\sigma_\epsilon^2 + \sigma_\delta^2\boldsymbol{\beta}'\boldsymbol{\beta}\right) \\ &= \text{trace}[(\boldsymbol{x}_0\boldsymbol{x}_0')E\{(\boldsymbol{b}_\delta^{(l)} - \boldsymbol{\beta})(\boldsymbol{b}_\delta^{(l)} - \boldsymbol{\beta})'\}] + \lambda^2\left(\sigma_\epsilon^2 + \sigma_\delta^2\boldsymbol{\beta}'\boldsymbol{\beta}\right) \\ &= \text{trace}[\frac{1}{n}(\boldsymbol{x}_0\boldsymbol{x}_0')E\{n(\boldsymbol{b}_\delta^{(l)} - \boldsymbol{\beta})(\boldsymbol{b}_\delta^{(l)} - \boldsymbol{\beta})'\}] + \lambda^2\left(\sigma_\epsilon^2 + \sigma_\delta^2\boldsymbol{\beta}'\boldsymbol{\beta}\right),\end{aligned}$$

for $l = 1, 2, 3$. Conditional asymptotic mean squared errors of $\hat{T}_1$, $\hat{T}_2$ and $\hat{T}_3$ for known $\boldsymbol{x}_0$ are given by

$$\begin{aligned}\lim_{n\to\infty} \text{MSE}(\hat{T}_l|\boldsymbol{x}_0) &= \lim_{n\to\infty} E[(\hat{T}_l - T)^2|\boldsymbol{x}_0] \\ &= \lim_{n\to\infty} \text{trace}[\frac{1}{n}(\boldsymbol{x}_0\boldsymbol{x}_0')E\{n(\boldsymbol{b}_\delta^{(l)} - \boldsymbol{\beta})(\boldsymbol{b}_\delta^{(l)} - \boldsymbol{\beta})'\}] + \lambda^2\left(\sigma_\epsilon^2 + \sigma_\delta^2\boldsymbol{\beta}'\boldsymbol{\beta}\right).\end{aligned}$$

From Lemma (4.1), it is clear that $\lim_{n\to 0} E\{n(\boldsymbol{b}_\delta^{(l)} - \boldsymbol{\beta})(\boldsymbol{b}_\delta^{(l)} - \boldsymbol{\beta})'\} = A_l\Omega A_l'$ for $l = 1, 2, 3$. Thus for known $\boldsymbol{x}_0$, we get

$$\lim_{n\to\infty} \text{MSE}(\hat{T}_l|\boldsymbol{x}_0) = \lambda^2\left(\sigma_\epsilon^2 + \sigma_\delta^2\boldsymbol{\beta}'\boldsymbol{\beta}\right),$$

for $l = 1, 2, 3$. □

Thus, the asymptotic mean squared errors (MSE) of all three predictors $\hat{T}_1$, $\hat{T}_2$ and $\hat{T}_3$ for given $\boldsymbol{x}_0$ are the same. However, it could be noted that the MSE is affected by the constant λ. Recall that $0 \leq \lambda \leq 1$ is a real number specifying the weight assigned to the prediction of actual and average values of y_0. Clearly, for $\lambda = 0$, this MSE is zero. When $\lambda = 0$, predicting target function is same as predicting average value of study variable. Also note that the asymptotic MSE of these predictors is not affected by the skewness and kurtosis of the distribution of measurement errors. Some more insight of the behavior of MSE is studied through Monte-Carlo simulation in the next section.

Also, we have not obtained asymptotic bias of these predictors. Cheng and Kukush (2006) proved that the first moment of $\boldsymbol{b}_\delta^{(1)}$ does not exist. Because of the relationship of $\boldsymbol{b}_\delta^{(2)}$ and $\boldsymbol{b}_\delta^{(3)}$ on $\boldsymbol{b}_\delta^{(1)}$, it looks that the first moment of later estimators also do not exist. Therefore, we have doubt on the existence of the first moment of out predictors. Simulation study conducted in the next section confirms this doubt.

5 Simulation Study

In order to study efficiencies of the predictors $\hat{T}_1$, $\hat{T}_2$ and $\hat{T}_3$, we conducted a Monte-Carlo simulation study. We adopted various combinations of values of $\boldsymbol{\beta}$, σ_ϵ^2, σ_ϕ^2, σ_δ^2. In order to observe the effect of departure from normality in terms of skewness and kurtosis, we adopted Normal distribution (having zero skewness and zero kurtosis), Student's t distribution with degree of freedom 12 (having zero skewness but non-zero kurtosis), and Gamma Distribution (having both non-zero skewness and non-zero kurtosis).

The random observations ϵ_i, ϕ_{ij}, and δ_{ij}, $i = 1, 2, \ldots, n, j = 1, 2, \ldots, 5$, are generated and transformed suitably to have zero mean and specified variances. A new vector-valued observed covariate $\boldsymbol{x}_0$ and corresponding value of study variable y_0 is fixed in advance. For various values of λ, the target function T is calculated. The value of predictors $\hat{T}_1$, $\hat{T}_2$ and $\hat{T}_3$ are computed for various combinations of parametric values.

We obtained absolute bias (AB)of the predictors $\hat{T}_1$, $\hat{T}_2$ and $\hat{T}_3$ empirically based on $100,000$ repetitions under aforesaid distributional assumptions and for various combinations of parametric values. For an estimator $\hat{T}$ of the regression coefficients T and for a fixed parametric value, the absolute bias of $\hat{T}$ is defined as

$$\text{AB}(\hat{T}) \quad = \quad \sqrt{\big(E(\hat{T}-T)\big)'\big(E(\hat{T}-T)\big)}. \tag{5.1}$$

We observe that these AB do not converge to anywhere and are vague in nature. Cheng and Kukush (2006) proved that the moments of the estimator $\boldsymbol{b}_\delta^{(1)}$ do not exist. Since the estimators $\boldsymbol{b}_\delta^{(2)}$, and $\boldsymbol{b}_\delta^{(3)}$ are related to $\boldsymbol{b}_\delta^{(1)}$, there is a doubt on the existence of moments of $\boldsymbol{b}_\delta^{(2)}$ and $\boldsymbol{b}_\delta^{(3)}$. The simulation findings support this doubt. Therefore, to study the small sample properties of these estimators we consider the criteria of absolute median bias which is defined as

$$\text{AMdB}(\hat{T}) \quad = \quad \sqrt{\big(\text{Median}(\hat{T})-T\big)'\big(\text{Median}(\hat{T})-T\big)}. \tag{5.2}$$

In order to study the efficiencies of the predictors $\hat{T}_1$, $\hat{T}_2$ and $\hat{T}_3$, we obtain mean squared error (MSE) empirically based on $100,000$ repetitions under aforesaid distributional assumptions and for various combinations of parametric values. For an estimator $\hat{T}$ of the regression coefficients T and for a fixed parametric value, the MSE of $\hat{T}$ is defined as

$$\text{MSE}(\hat{T}) \quad = \quad E\big((\hat{T}-T)'(\hat{T}-T)\big). \tag{5.3}$$

In order to save space, here we present only few important outcomes of simulation in Tables 1–6. From out simulation study, we observe that absolute median bias of restricted predictors $\hat{T}_2$ and $\hat{T}_3$ are approximately the same. It is not clear, which predictor is having the least bias. For sample size 50, all three absolute bias are approximately the same. AMdB are minimum under normal distribution and much larger in case of Gamma and t distribution. The variance σ_δ^2 affects the AMdB in the positive direction in case of Gamma and T distributions and not in case of normal distribution. The values of λ have significant effect on AMdB of these predictors.

We also observe that mean squared error (MSE) is the highest for unrestricted predictor $\hat{T}_1$, while it is approximately the same for restricted predictors $\hat{T}_2$ and $\hat{T}_3$. For sample size 50, all three

MSE are approximately the same. The variance σ_δ^2 affects the MSE in the positive direction. The values of MSE are minimum under normal distribution and much larger in case of Gamma and t distribution. The values of λ have significant effect on MSE of these predictors. MSE are highest when $\lambda = 1$. From Theorem 1, it is clear that asymptotic MSE are zero for all three predictors when $\lambda = 0$. However, for sample size 20 and 50, simulation does not suggest that. It means that the small sample behavior of these predictors are very different than that in large sample case and larger sample is required for the convergence of MSE.

In order to provide a straightaway idea of the properties of obtained estimators, the AMdB of the estimators are shown in Figures 1 – 6 and MSE of the estimators are shown in Figures 7 – 12.

Table 1: Absolute Median bias (AMdB) of $\hat{T}_l,\ l = 1, 2, 3$, when $(\boldsymbol{\epsilon}, \boldsymbol{\phi}, \boldsymbol{\delta})$ have normal distribution

$\sigma_\epsilon^2 = 0.5, \sigma_\phi^2 = 0.5, \sigma_\delta^2 = 0.5$						
	$n = 20$			$n = 50$		
λ	$\hat{T}_1$	$\hat{T}_2$	$\hat{T}_3$	$\hat{T}_1$	$\hat{T}_2$	$\hat{T}_3$
0.000	0.691	0.004	0.011	0.108	0.001	0.002
0.250	2.219	2.904	2.914	2.851	2.917	2.910
0.500	4.421	3.397	3.391	3.577	3.384	3.390
0.750	0.026	0.820	0.824	0.675	0.826	0.823
1.000	0.232	0.824	0.824	0.713	0.824	0.825
$\sigma_\epsilon^2 = 0.5, \sigma_\phi^2 = 0.5, \sigma_\delta^2 = 1.25$						
	$n = 20$			$n = 50$		
λ	$\hat{T}_1$	$\hat{T}_2$	$\hat{T}_3$	$\hat{T}_1$	$\hat{T}_2$	$\hat{T}_3$
0.000	1.103	0.011	0.007	0.274	0.010	0.000
0.250	1.583	0.423	0.421	0.218	0.420	0.421
0.500	4.586	3.214	3.156	3.479	3.119	3.141
0.750	1.634	0.491	0.490	0.134	0.489	0.490
1.000	1.525	0.348	0.329	0.556	0.321	0.328

Table 2: Absolute Median bias (AMdB) of $\hat{T}_l,\ l = 1, 2, 3$, when $(\boldsymbol{\epsilon}, \boldsymbol{\phi}, \boldsymbol{\delta})$ have Gamma distribution

$\sigma^2_\epsilon = 0.5, \sigma^2_\phi = 0.5, \sigma^2_\delta = 0.5$						
	$n = 20$			$n = 50$		
λ	$\hat{T}_1$	$\hat{T}_2$	$\hat{T}_3$	$\hat{T}_1$	$\hat{T}_2$	$\hat{T}_3$
0.000	0.834	0.001	0.003	0.017	0.278	0.273
0.250	1.033	0.285	0.293	0.168	0.003	0.004
0.500	0.209	0.871	0.877	0.686	0.868	0.865
0.750	2.395	1.375	1.356	1.492	1.347	1.354
1.000	4.176	3.171	3.168	3.405	3.174	3.175
$\sigma^2_\epsilon = 0.5, \sigma^2_\phi = 0.5, \sigma^2_\delta = 1.25$						
	$n = 20$			$n = 50$		
λ	$\hat{T}_1$	$\hat{T}_2$	$\hat{T}_3$	$\hat{T}_1$	$\hat{T}_2$	$\hat{T}_3$
0.000	2.364	0.242	0.075	0.479	0.047	0.026
0.250	0.391	1.217	1.088	0.486	1.046	1.079
0.500	4.430	0.778	0.782	1.892	0.783	0.782
0.750	6.268	7.965	7.981	7.487	7.985	7.980
1.000	8.435	9.204	9.317	9.198	9.333	9.310

6 Concluding Remarks

We defined a target function given by T in (2.4) for simultaneous prediction of actual and average value of the study variable in an ultrastructural measurement error model. We also assumed that some prior information on regression coefficients is available in terms of exact liner restrictions given by (2.5). We obtained one unrestricted predictor given by $\hat{T}_1$ and two restricted predictors $\hat{T}_2$ and $\hat{T}_3$ of the target function T. Asymptotic mean squared errors of these predictors are obtained in Theorem 1 and are the same. However, the simulation study suggest they are not the same even for a sample size of 50. The outcome of simulation study clearly suggest that restricted predictors $\hat{T}_2$ and $\hat{T}_3$ are far efficient than unrestricted predictor $\hat{T}_1$. The distributions of measurement errors also play an important role in the efficiency of these predictors. Although, the effect of departure from normality is not present in asymptotic mean square errors, it can clearly be noticed in simulation results. The variance of measurement errors σ^2_δ affects the efficiencies of these predictors; larger the variance, lesser the efficiency. The constant λ also affect the efficiency of the predictors. The direction of the effect is not very clear for as large sample size as 50. However, the change is in positive direction asymptotically.

Thus, in the situations when some prior information about regression coefficients is available in terms of exact linear restrictions and we wish to predict the actual value and average values simultaneously, we recommend the usage of $\hat{T}_2$ or $\hat{T}_3$. The efficiency properties of the restricted estimators, $\hat{T}_2$ or $\hat{T}_3$ are almost the same.

Table 3: Absolute Median bias (AMdB) of $\hat{T}_l,\ l = 1, 2, 3$, when $(\boldsymbol{\epsilon}, \boldsymbol{\phi}, \boldsymbol{\delta})$ have Student's t distribution with 12 d.f.

$\sigma_\epsilon^2 = 0.5, \sigma_\phi^2 = 0.5, \sigma_\delta^2 = 0.5$						
	$n = 20$			$n = 50$		
λ	$\hat{T}_1$	$\hat{T}_2$	$\hat{T}_3$	$\hat{T}_1$	$\hat{T}_2$	$\hat{T}_3$
0.000	0.792	0.005	0.001	0.123	0.003	0.001
0.250	0.524	1.205	1.213	1.092	1.219	1.213
0.500	0.636	1.454	1.460	1.344	1.458	1.452
0.750	3.927	4.709	4.713	4.633	4.734	4.725
1.000	2.212	1.490	1.489	1.576	1.481	1.486
$\sigma_\epsilon^2 = 0.5, \sigma_\phi^2 = 0.5, \sigma_\delta^2 = 1.25$						
	$n = 20$			$n = 50$		
λ	$\hat{T}_1$	$\hat{T}_2$	$\hat{T}_3$	$\hat{T}_1$	$\hat{T}_2$	$\hat{T}_3$
0.000	1.598	0.029	0.002	0.523	0.012	0.003
0.250	3.748	2.099	2.050	2.517	2.037	2.051
0.500	1.237	1.271	1.423	0.823	1.495	1.441
0.750	7.677	5.971	5.899	6.403	5.880	5.898
1.000	9.404	7.367	7.290	7.795	7.267	7.292

Table 4: Mean squared error (MSE) of $\hat{T}_l,\ l = 1, 2, 3$, when $(\boldsymbol{\epsilon}, \boldsymbol{\phi}, \boldsymbol{\delta})$ have normal distribution

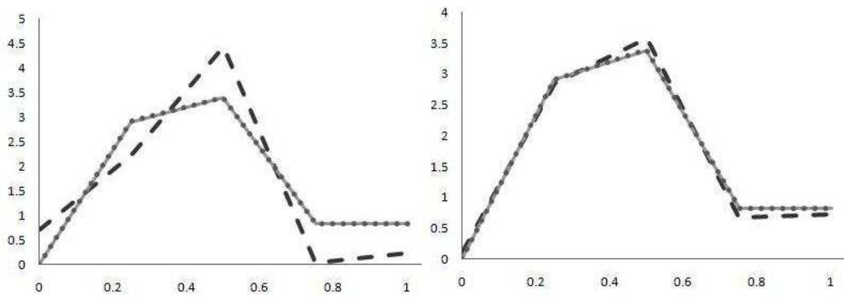
$\sigma_\epsilon^2 = 0.5, \sigma_\phi^2 = 0.5, \sigma_\delta^2 = 0.5$						
	$n = 20$			$n = 50$		
λ	$\hat{T}_1$	$\hat{T}_2$	$\hat{T}_3$	$\hat{T}_1$	$\hat{T}_2$	$\hat{T}_3$
0.000	7.937	0.489	0.468	1.418	0.174	0.171
0.250	19.072	9.064	9.015	8.884	8.514	8.469
0.500	21.058	11.562	11.508	12.797	11.454	11.494
0.750	14.153	0.755	0.744	2.742	0.683	0.679
1.000	10.049	0.680	0.680	2.155	0.679	0.680
$\sigma_\epsilon^2 = 0.5, \sigma_\phi^2 = 0.5, \sigma_\delta^2 = 1.25$						
	$n = 20$			$n = 50$		
λ	$\hat{T}_1$	$\hat{T}_2$	$\hat{T}_3$	$\hat{T}_1$	$\hat{T}_2$	$\hat{T}_3$
0.000	127.514	0.780	0.462	12.514	0.174	0.158
0.250	81.947	0.183	0.177	7.559	0.176	0.178
0.500	4.586	3.214	3.156	13.921	9.744	9.868
0.750	117.297	0.242	0.240	11.650	0.239	0.240
1.000	56.050	0.691	0.432	6.167	0.203	0.189

Table 5: Mean squared error (MSE) of $\hat{T}_l,\ l = 1, 2, 3$, when $(\boldsymbol{\epsilon}, \boldsymbol{\phi}, \boldsymbol{\delta})$ have Gamma distribution

$\sigma^2_\epsilon = 0.5, \sigma^2_\phi = 0.5, \sigma^2_\delta = 0.5$						
	$n = 20$			$n = 50$		
λ	$\hat{T}_1$	$\hat{T}_2$	$\hat{T}_3$	$\hat{T}_1$	$\hat{T}_2$	$\hat{T}_3$
0.000	10.953	0.079	0.076	2.078	0.285	0.282
0.250	11.319	0.745	0.708	1.957	0.027	0.027
0.500	14.954	0.805	0.867	2.848	0.753	0.749
0.750	23.824	3.279	3.125	5.526	1.985	1.994
1.000	23.397	10.052	10.038	11.674	10.074	10.792
$\sigma^2_\epsilon = 0.5, \sigma^2_\phi = 0.5, \sigma^2_\delta = 1.25$						
	$n = 20$			$n = 50$		
λ	$\hat{T}_1$	$\hat{T}_2$	$\hat{T}_3$	$\hat{T}_1$	$\hat{T}_2$	$\hat{T}_3$
0.000	100.626	28.790	17.315	12.368	6.664	5.975
0.250	112.347	7.343	4.458	11.227	2.064	1.956
0.500	169.533	0.608	0.611	18.057	0.612	0.611
0.750	148.767	63.494	63.706	57.191	63.752	63.683
1.000	122.693	87.073	86.887	84.718	87.109	86.677

Table 6: Mean squared error (MSE) of $\hat{T}_l,\ l = 1, 2, 3$, when $(\boldsymbol{\epsilon}, \boldsymbol{\phi}, \boldsymbol{\delta})$ have Student's t distribution with 12 d.f.

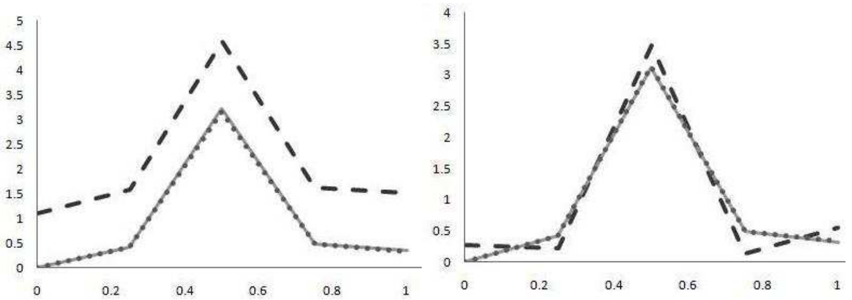
$\sigma^2_\epsilon = 0.5, \sigma^2_\phi = 0.5, \sigma^2_\delta = 0.5$						
	$n = 20$			$n = 50$		
λ	$\hat{T}_1$	$\hat{T}_2$	$\hat{T}_3$	$\hat{T}_1$	$\hat{T}_2$	$\hat{T}_3$
0.000	14.712	0.803	0.765	2.658	0.281	0.275
0.250	6.587	1.839	1.791	1.822	1.494	1.482
0.500	15.500	2.807	2.739	3.939	2.155	2.140
0.750	21.592	22.185	22.211	21.462	22.414	22.324
1.000	11.541	2.395	2.353	3.399	2.193	2.209
$\sigma^2_\epsilon = 0.5, \sigma^2_\phi = 0.5, \sigma^2_\delta = 1.25$						
	$n = 20$			$n = 50$		
λ	$\hat{T}_1$	$\hat{T}_2$	$\hat{T}_3$	$\hat{T}_1$	$\hat{T}_2$	$\hat{T}_3$
0.000	78.427	0.948	0.595	6.613	0.224	0.204
0.250	111.859	5.510	4.424	12.249	4.161	4.210
0.500	82.781	20.503	12.840	8.525	5.670	5.175
0.750	93.317	36.913	34.872	41.047	34.582	34.792
1.000	192.283	56.791	53.266	61.289	52.816	53.166



(a) $n = 20$ (b) $n = 50$

AMdB($\hat{T}_1$) AMdB($\hat{T}_2$) AMdB($\hat{T}_3$)

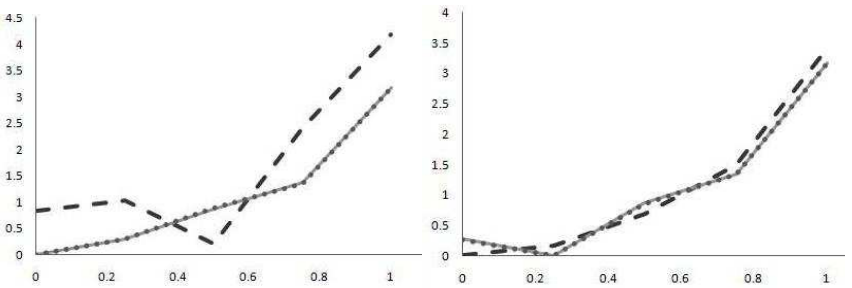
Figure 1: AMdB of $\hat{T}_1, \hat{T}_2, \hat{T}_3$ in case of normal distribution with $\sigma_\epsilon^2 = 0.5, \sigma_\phi^2 = 0.5, \sigma_\delta^2 = 0.5$



(a) $n = 20$ (b) $n = 50$

AMdB($\hat{T}_1$) AMdB($\hat{T}_2$) AMdB($\hat{T}_3$)

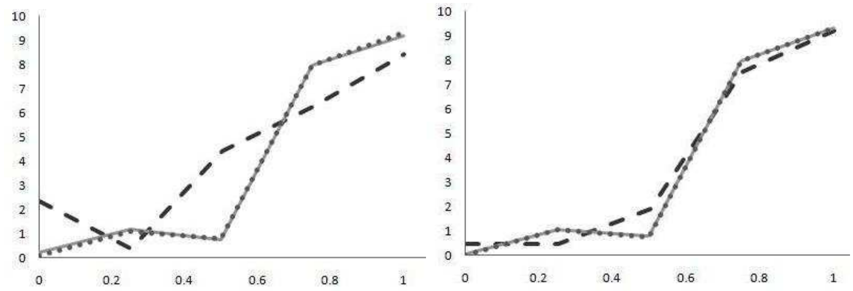
Figure 2: AMdB of $\hat{T}_1, \hat{T}_2, \hat{T}_3$ in case of normal distribution with $\sigma_\epsilon^2 = 0.5, \sigma_\phi^2 = 0.5, \sigma_\delta^2 = 1.25$



(a) $n = 20$ (b) $n = 50$

AMdB($\hat{T}_1$) AMdB($\hat{T}_2$) AMdB($\hat{T}_3$)

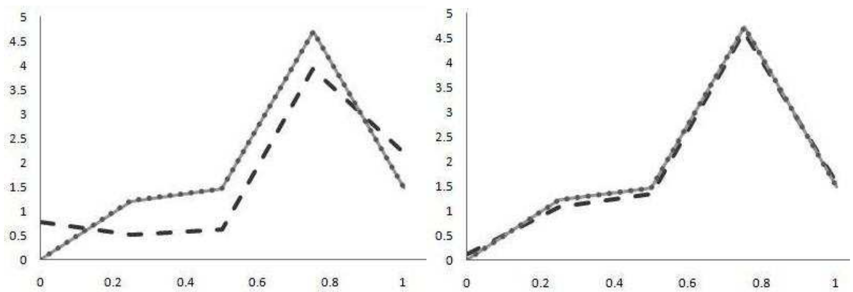
Figure 3: AMdB of $\hat{T}_1, \hat{T}_2, \hat{T}_3$ in case of Gamma distribution with $\sigma_\epsilon^2 = 0.5, \sigma_\phi^2 = 0.5, \sigma_\delta^2 = 0.5$



(a) $n = 20$ (b) $n = 50$

— — - AMdB($\hat{T}_1$) ········ AMdB($\hat{T}_2$) ——— AMdB($\hat{T}_3$)

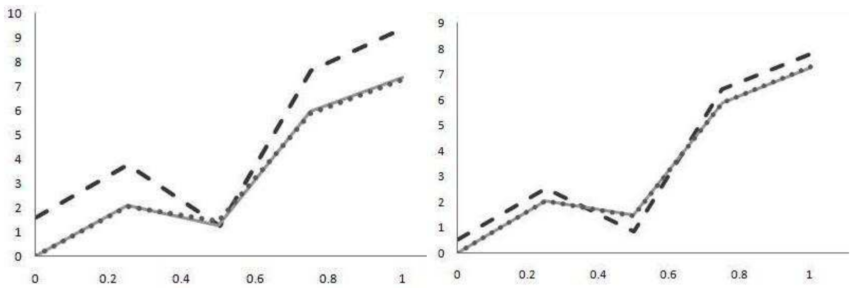
Figure 4: AMdB of $\hat{T}_1, \hat{T}_2, \hat{T}_3$ in case of Gamma distribution with $\sigma_\epsilon^2 = 0.5, \sigma_\phi^2 = 0.5, \sigma_\delta^2 = 1.25$



(a) $n = 20$ (b) $n = 50$

— — - AMdB($\hat{T}_1$) ········ AMdB($\hat{T}_2$) ——— AMdB($\hat{T}_3$)

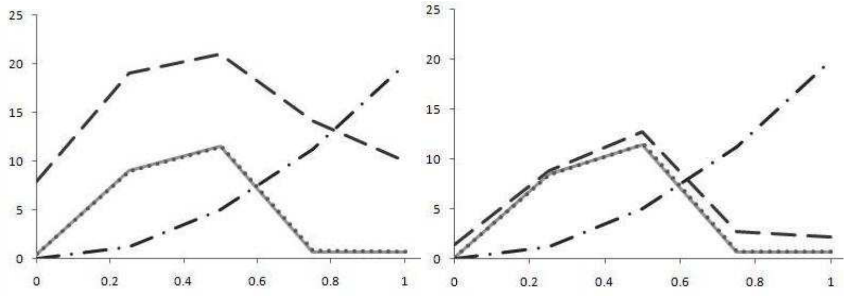
Figure 5: AMdB of $\hat{T}_1, \hat{T}_2, \hat{T}_3$ in case of $t_{(12)}$ distribution with $\sigma_\epsilon^2 = 0.5, \sigma_\phi^2 = 0.5, \sigma_\delta^2 = 0.5$



(a) $n = 20$ (b) $n = 50$

— — - AMdB($\hat{T}_1$) ········ AMdB($\hat{T}_2$) ——— AMdB($\hat{T}_3$)

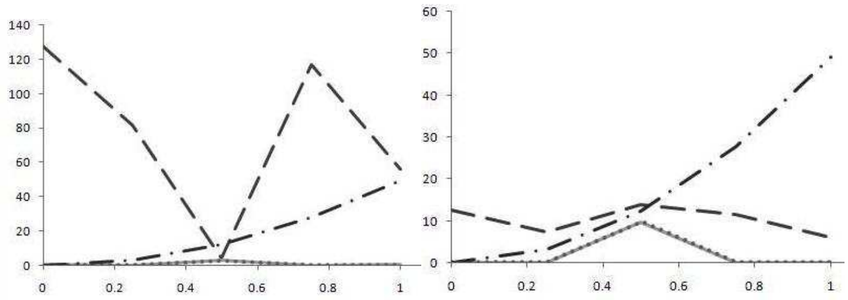
Figure 6: AMdB of $\hat{T}_1, \hat{T}_2, \hat{T}_3$ in case of $t_{(12)}$ distribution with $\sigma_\epsilon^2 = 0.5, \sigma_\phi^2 = 0.5, \sigma_\delta^2 = 1.25$



(a) $n = 20$ (b) $n = 50$

MSE($\hat{T}_1$) MSE($\hat{T}_1$) MSE($\hat{T}_3$) $\lim_{n\to\infty}$MSE($\hat{T}_i \mid x_0$)

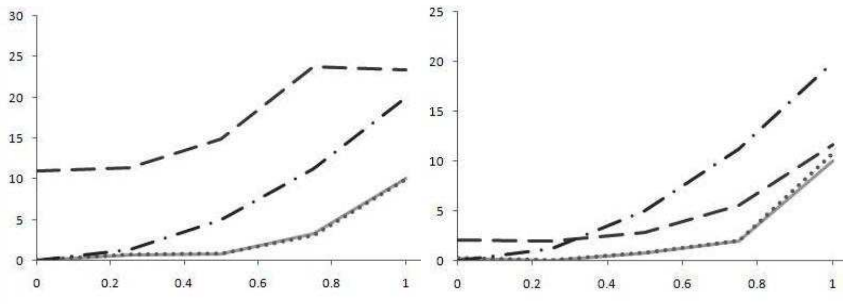
Figure 7: MSE of $\hat{T}_1, \hat{T}_2, \hat{T}_3$ in case of normal distribution with $\sigma_\epsilon^2 = 0.5, \sigma_\phi^2 = 0.5, \sigma_\delta^2 = 0.5$



(a) $n = 20$ (b) $n = 50$

MSE($\hat{T}_1$) MSE($\hat{T}_1$) MSE($\hat{T}_3$) $\lim_{n\to\infty}$MSE($\hat{T}_i \mid x_0$)

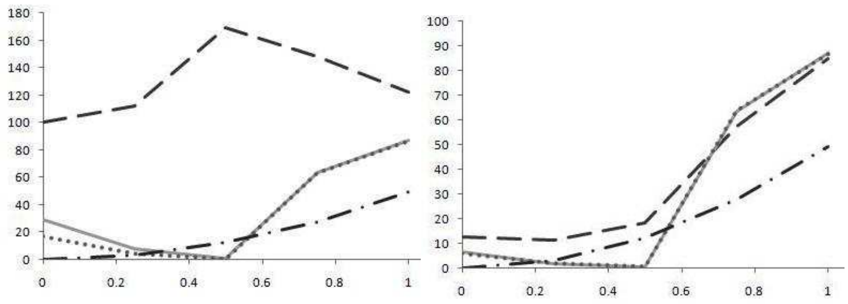
Figure 8: MSE of $\hat{T}_1, \hat{T}_2, \hat{T}_3$ in case of normal distribution with $\sigma_\epsilon^2 = 0.5, \sigma_\phi^2 = 0.5, \sigma_\delta^2 = 1.25$



(a) $n = 20$ (b) $n = 50$

MSE($\hat{T}_1$) MSE($\hat{T}_1$) MSE($\hat{T}_3$) $\lim_{n\to\infty}$MSE($\hat{T}_i \mid x_0$)

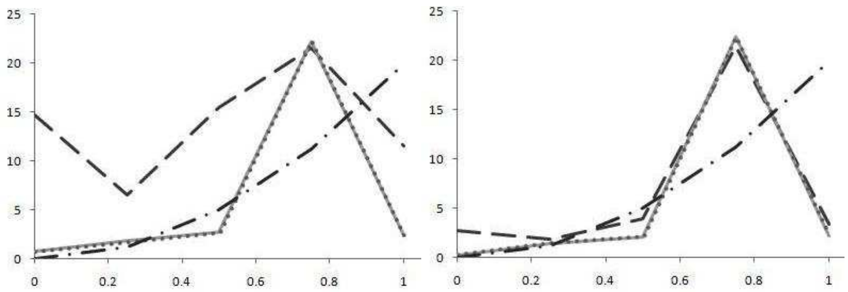
Figure 9: MSE of $\hat{T}_1, \hat{T}_2, \hat{T}_3$ in case of Gamma distribution with $\sigma_\epsilon^2 = 0.5, \sigma_\phi^2 = 0.5, \sigma_\delta^2 = 0.5$



(a) $n = 20$ (b) $n = 50$

MSE($\hat{T}_1$) MSE($\hat{T}_1$) MSE($\hat{T}_3$) $\lim_{n\to\infty}$MSE($\hat{T}_i \mid \mathbf{x}_0$)

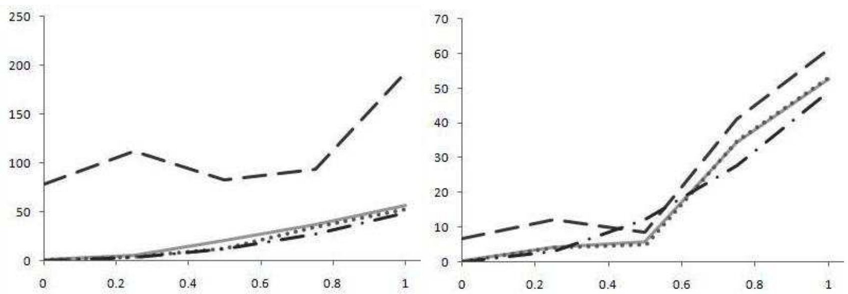
Figure 10: MSE of $\hat{T}_1, \hat{T}_2, \hat{T}_3$ in case of Gamma distribution with $\sigma^2_\epsilon = 0.5, \sigma^2_\phi = 0.5, \sigma^2_\delta = 1.25$



(a) $n = 20$ (b) $n = 50$

MSE($\hat{T}_1$) MSE($\hat{T}_1$) MSE($\hat{T}_3$) $\lim_{n\to\infty}$MSE($\hat{T}_i \mid \mathbf{x}_0$)

Figure 11: MSE of $\hat{T}_1, \hat{T}_2, \hat{T}_3$ in case of $t_{(12)}$ distribution with $\sigma^2_\epsilon = 0.5, \sigma^2_\phi = 0.5, \sigma^2_\delta = 0.5$



(a) $n = 20$ (b) $n = 50$

MSE($\hat{T}_1$) MSE($\hat{T}_1$) MSE($\hat{T}_3$) $\lim_{n\to\infty}$MSE($\hat{T}_i \mid \mathbf{x}_0$)

Figure 12: MSE of $\hat{T}_1, \hat{T}_2, \hat{T}_3$ in case of $t_{(12)}$ distribution with $\sigma^2_\epsilon = 0.5, \sigma^2_\phi = 0.5, \sigma^2_\delta = 1.25$

References

[1] Cheng, C. L. and Van Ness, J. W. (1991). On the unreplicated ultrastructural model, *Biometrika*, **78**, 442–445.

[2] Cheng, C. L. and Van Ness, J. W. (1999). *Statistical Regression with Measurement Errors*. Arnold Publishers.

[3] Cheng, C. L. and Kukush, A. (2006). Non-existence of the first moment of the adjusted least squares estimator in multivariate errors-in-variables model. *Metrika* **64**, N1, 41–46.

[4] Fuller, W. A. (1987). *Measurement Error Models*. John Willey.

[5] Dolby, G. R. (1976). The ultrastructural relation: a synthesis of the functional and structural relations. *Biometrika*, **63**, 39–50.

[6] Gleser, L. J. (1984). A note on G.R. Dolby's ultrastructural model. *Biometrika* **72**, 117–124.

[7] Rao, C. R. and Toutenburg, H. Shalabh, Heumann, C. (2008). *Linear models: least squares and alternatives,* 3^{rd} *edition*. Springer.

[8] Shalabh (1995): Performance of Stein-Rule procedure for simultaneous prediction of actual and average values of study variable in linear regression model. *Bulletin of the International Statistical Institute* **56**, 1375–1390.

[9] Shalabh, Garg, G. and Misra, N. (2007). Restricted regression estimation in measurement error models, *Computational Statistics and Data Analysis* **52**(2), 1149–1166.

[10] Shalabh, Gaurav, G. and Misra, N. (2009). Use of prior information in the consistent estimation of regression coefficients in measurement error models. *Journal of Multivariate Analysis* **100**(7), 1498–1520.

[11] Shalabh, Gleser J. L. and Rosen, O. (2004). On the usefulness of the knowledge of error variances in the consistent estimation of an unreplicated ultrastructural model. *Journal of Statistical Computation & Simulation,* **74**(6), 391–417.

[12] Srivastava, A. K. and Shalabh (1997). Consistent estimation for the non-normal ultrastructural model. *Stat. & Prob. Letters* **34**, 67–73.

[13] Srivastava, A. K. and Shalabh (1997). Asymptotic efficiency properties of least squares in an ultrastructural model. *Sociedad Espanola de Estadistica e Investigacion Operativa* **6**(2), 419–431.

[14] Toutenburg, H. (1982). *Prior information in linear models*. John Wiley.

[15] Toutenberg, H. and Shalabh (1996). Predictive performance of the methods of restricted and mixed regression estimators. *Biometrical Journal* **38**(8), 951–959.