JIRSS (2022)
Vol. 21, No. 01, pp 105-125
DOI: 10.22034/JIRSS.2022.704615

Corrected Likelihood Estimation in Semiparametric Linear Mixed Measurement Error Models: Asymptotic Results

Hadi Emami [1], **Shaho Zarei** [1]

[1] Department of statistics, University of Kurdistan, Sanandaj, Iran.

Received: 03/08/2021, Accepted: 30/12/2022, Published online: 03/05/2023

Abstract. This paper is concerned with the estimation problem in semiparametric linear mixed models when some of the covariates are measured with errors. The authors proposed the corrected score function estimators for the parametric and non parametric components. The resulting estimators are shown to be consistent and asymptotically normal. An iterative algorithm is proposed for estimating the parameters. Asymptotic normality of the estimators is also derived. Finite sample performance of the proposed estimators is assessed by Monte Carlo simulation studies. We further illustrate the proposed procedures by an application.

Keywords. Corrected Score Method, Smoothing Spline, Semiparametric Linear Models, Linear Measurement Error Models.

MSC: 62J05, 62G05, 62G08, 62G20.

Corresponding Author: Hadi Emami (h.emami@uok.ac.ir)
Shaho Zarei (sh.zarei@uok.ac.ir)

1 Introduction

The parametric linear mixed model has been frequently used in repeated measures and grouped and longitudinal data. Applications have been reported in different areas such as agriculture, biology, economy, geophysics and social sciences, Diggle et al. (2002). But parametric linear mixed models are sometimes inappropriate and likely to introduce modeling biases when underlying models are complicated. To relax the parametric assumptions, various semiparametric models have been developed for longitudinal data. See, Chen and Jin (2006), Ruppert et al (2003), Fan and Li (2001). Semiparametric mixed models (Diggle et al. (2002), Zhang et al. (1998), Fuller (1987), Fung et al. (2002) and Emami and Mansoori (2018)) are useful extensions to linear mixed models, which use parametric fixed effects to represent the covariate effects and an arbitrary smooth function to model the time effect, and accounts for the within-subject correlation using random effects.

On the other hand, in practice there are many cases under which the covariates are unobservable, or say that they are observed with measurement error. For instance, it has been well documented in the literature that the covariates such as blood pressure, urinary sodium chloride level and exposure to pollutants are often subject to measurement errors. For this reason, there has been extensive research in the linear measurement error problem, as can be seen, for example, in the books by Fuller (1987) and Wu (2010).

If the covariates are measured with errors are not properly accounted, it can lead us to incorrect statistical inferences. For example, a significant covariate can be considered not significant, as can be seen in Wu (2010). If the measurement errors are not accounted for naive approach, then the parameter estimates are biased and not consistent (see Carrol et al. (1995)). Recently, the combination of random effects and measurement errors in parametric linear mixed-effects models is investigated by some authors. For example, Zhong et al. (2002), Zare et al. (2011) and Riquelmea et al (2015) studied the estimation problem when the fixed effect has measurement errors. They applied the corrected score approach of Nakamura (1990) to obtain the estimators of the regression parameters and proved the asymptotic normality.

Measurement error problems in semiparametric context are less well studied than their parametric counterparts, probably due to the difficulty of handling multiple infinite-dimensional parameters. The combination of measurement error with semiparametric linear mixed models is worth investigating in linear models. It is often the case in practice that covariate values collected on individuals are measured with nonnegligible errors and the inference of these models is less developed. For example in this area, Emami and Mansoori (2018) developed the influence diagnostics approach in semiprametric linear mixed models with measurement error, Yalaz and Kuran (2020) applied the profile kernel method and used the weighted least squares to estimate the parameters. As the most recently study, Kuran and Yalaz (2022) extended ridge regression technique to combat multicollinearity in partially linear mixed measurement error models. Although in such studies the parametric term has received much attention, the asymptotic distribution of estimators and specially non parametric estimators has not been discussed.

In this paper, instead of the kernel method and least squares estimation, we derive the estimator of parametric and non parametric terms by employing the spline method and corrected score function of Nakamura (1990). We obtain the asymptotic distribution of the estimators of both parametric and nonparametric component with details. The plan of the paper is as follows: In Section 2, we introduce the corrected score function for semiparametric linear mixed models with measurement error. A specialized simple algorithm is developed to estimate the parameters and variance component. The asymptotic properties of the estimators are given in Section 3. To illustrate the proposed estimation methods, a simulation study is considered in Section 4 and an example of real data is presented in Section 5. Finally, concluding remarks are given in Section 6.

2 The Methodology

2.1 Model and Notations

Consider the following semiparametric linear mixed models with errors in variables

$$Y = \mathbf{Z}\beta + \mathbf{U}\mathbf{b} + f(\mathbf{t}) + \epsilon, \tag{1.a}$$

$$\mathbf{X} = \mathbf{Z} + \Delta, \tag{1.b}$$

where Y is a $n \times 1$ vector of observations, $\mathbf{Z}$ and $\mathbf{U} = [\mathbf{U}_1, \mathbf{U}_2, \ldots, \mathbf{U}_c]$ are matrices of "regressors" with dimensions $n \times p$ and $n \times q$ respectively. $\mathbf{U}_i$ is an $n \times q_i$ known design matrix of the random effect factor i and $b^\tau = (b_1^\tau, \ldots, b_c^\tau)$ where b_i is a $q_i \times 1$ vector of unobservable random effects from $\mathcal{N}(0, \sigma_i^2\mathbf{I})$, where $sigma_i^2$, $i = 1, \ldots, c$ are called variance components. β is a p-vector of parameters, f is a twice differentiable smooth function on some finite interval and $\mathbf{t} = (t_1, , \ldots, t_n)$ where t_i is a scalar $(a \leq t_1, \ldots, t_n \leq b)$, t_i's are not all identical. ϵ is an $n \times 1$ vector of unobservable random errors. Covariate

$\mathbf{Z}$ is unobservable for all study subjects which can be observed from random matrix $\mathbf{X}$. Following Zhong et al. (2002), the random errors follow $\mathcal{N}(0, \sigma^2\mathbf{I})$. $\mathbf{X}$ is the observed value of $\mathbf{Z}$ with the measurement error Δ, where Δ is random matrix from $\mathcal{N}(0, \mathbf{I} \otimes \mathbf{\Lambda})$ and $\mathbf{\Lambda}$ is positive definite matrix.

We assume that b_i, ϵ and Δ are mutually independent. In addition, we can write $b \sim \mathcal{N}(0, \sigma^2\mathbf{D})$ where $\mathbf{D}$ is a block diagonal matrix with the ith block being $\theta_i \mathbf{I}_{ci}$, with $\theta_i = \sigma_i^2/\sigma^2$, so that Y has a multivariate normal distribution with $\mathcal{E}(Y) = \mathbf{Z}\beta + f(t)$ and $var(Y) = \sigma^2\mathbf{V}$, in which $\mathbf{V} = \mathbf{I} + \mathbf{U}\mathbf{D}\mathbf{U}^\tau = \mathbf{I} + \sum_{i=1}^{c} \theta_i \mathbf{U}_i^\tau \mathbf{U}_i$. This is underlying model for response vector Y in term of covariates. We denote the ordered distinct values among $t_1, \ldots, t_n$ by $s_1, \ldots, s_q$. Then the connection between $t_1, \ldots, t_n$ and $s_1, \ldots, s_r$ is captured by means of the $n \times r$ incidence matrix $\mathbf{N}$, with entries $N_{ij} = 1$ if $t_i = s_j$ and 0 otherwise. Let $\mathbf{f}$ be the vector of values $a_i = f(s_i)$. The term in (1.a) can then be written as

$$\begin{aligned} Y &= \mathbf{X}\beta + \mathbf{U}b + \mathbf{N}\mathbf{f} + \epsilon, \\ \mathbf{X} &= \mathbf{Z} + \Delta, \end{aligned} \tag{2.2}$$

Let $\mathcal{L}(\beta, b, \mathbf{f}; \mathbf{Z}, Y) = logl(Y, b; \beta, \mathbf{f}, \mathbf{Z})$, where $l(Y, b; \beta, \mathbf{f}, \mathbf{Z})$ denote the joint probability density of Y and b, then from (2.2) the joint penalized log-likelihood is defined as

$$\begin{aligned} \mathcal{L}(\beta, b, \mathbf{f}; \mathbf{Z}, Y) = \kappa_{\sigma^2} &- \frac{1}{2\sigma^2}(Y - \mathbf{Z}\beta - \mathbf{U}b - \mathbf{N}\mathbf{f})^\tau (Y - \mathbf{Z}\beta - \mathbf{U}b - \mathbf{N}\mathbf{f}) \\ &- \frac{1}{2\sigma^2} b^\tau \mathbf{D}^{-1} b - \frac{\lambda}{2\sigma^2} \int f''(t)^2 dt, \end{aligned} \tag{2.3}$$

where $\kappa_{\sigma^2} = -(1/2)log(2\pi\sigma^2)^{n+q} - (1/2)log|\mathbf{D}|$. Following the approach of Harvill (1977) by solving equation $\frac{\partial \mathcal{L}}{\partial b} = 0$, we get $\tilde{b}_{\beta,\mathbf{f}}(\mathbf{Z}) = (\mathbf{U}^\tau\mathbf{U} + \mathbf{D}^{-1})^{-1}\mathbf{U}^\tau(Y - \mathbf{Z}\beta - \mathbf{N}\mathbf{f})$. Substituting this formula in equation (2.3) and after some simplifications, we have

$$\begin{aligned} \mathcal{L}_p(\beta, \mathbf{f}; \mathbf{Z}, Y) = \kappa_{\sigma^2} &- \frac{1}{2\sigma^2}(Y - \mathbf{Z}\beta - \mathbf{N}\mathbf{f})^\tau \mathbf{V}^{-1} (Y - \mathbf{Z}\beta - \mathbf{N}\mathbf{f}) \\ &- \frac{\lambda}{2\sigma^2} \mathbf{f}^\tau \mathbf{K} \mathbf{f}, \end{aligned} \tag{2.4}$$

where λ is a smoothing parameter and $\mathbf{K}$ is the non-negative definite smoothing matrix, see Emami and Mansoori (2018). In this paper, the choice of the smoothing parameter λ is accomplished by minimizing the generalized cross validation citerion $GCV(\lambda)$.

2.2 Estimation of Parametric and non Parametric Components

As we have mentioned before, the covariate $\mathbf{Z}$ is measured with error and the correlated structure arises from the random effects. Since the covariate $\mathbf{Z}$ is measured with error, if we simply replace $\mathbf{Z}$ by $\mathbf{X}$, then the estimates obtained from the score functions are not consistent in general. Various ways are proposed in dealing with measurement error models. In this paper, we use corrected score method proposed by Nakamura (1990) which is a common approach in measurement error models. In this method, we have to find a corrected score function whose expectation with respect to the measurement error distribution coincides with the usual score function based on the unknown true independent variables. Let $\mathcal{E}^*$ denotes the conditional mean with respect to $\mathbf{X}$ given Y. The corrected penalized loglikelihood $\mathcal{L}^*(\beta, b, \mathbf{f}; \mathbf{X})$ for our model should satisfy

$$\mathcal{E}^*[\partial\mathcal{L}^*(\beta, b, \mathbf{f}; \mathbf{X}, Y)/\partial b] = \partial\mathcal{L}(\beta, b, \mathbf{f}; \mathbf{Z}, Y)/\partial b, \tag{2.5}$$

$$\mathcal{E}^*[\partial\mathcal{L}_p^*(\beta, \mathbf{f}; \mathbf{X}, Y)/\partial\beta] = \partial\mathcal{L}_p(\beta, \mathbf{f}; \mathbf{Z}, Y)/\partial\beta. \tag{2.6}$$

Given $\mathbf{\Lambda}$, from (2.4) and (2.5), $\mathcal{L}^*$ is obtained as

$$\begin{aligned}\mathcal{L}^*(\beta, b, \mathbf{f}; \mathbf{X}, Y) = \kappa_{\sigma^2} - \frac{1}{2\sigma^2}\{(Y - \mathbf{X}\beta - \mathbf{U}b - \mathbf{N}\mathbf{f})^\tau(Y - \mathbf{X}\beta - \mathbf{U}b - \mathbf{N}\mathbf{f}) - tr(\mathbf{V}^{-1})\beta^\tau\mathbf{\Lambda}\beta\} \\ -\frac{1}{2\sigma^2}b^\tau\mathbf{D}^{-1}b - \frac{\lambda}{2\sigma^2}\mathbf{f}^\tau\mathbf{K}\mathbf{f}.\end{aligned} \tag{2.7}$$

An explicit expression for b is derived from the corrected penalized likelihood equation $\frac{\partial\mathcal{L}^*}{\partial b} = 0$ as

$$\begin{aligned}\tilde{b}_{\beta,\mathbf{f}}(\mathbf{X}) &= (\mathbf{U}^\tau\mathbf{U} + \mathbf{D}^{-1})^{-1}\mathbf{U}^\tau(Y - \mathbf{X}\beta - \mathbf{N}\mathbf{f}) \\ &= \mathbf{D}\mathbf{U}^\tau\mathbf{V}^{-1}(Y - \mathbf{X}\beta - \mathbf{N}\mathbf{f}).\end{aligned} \tag{2.8}$$

Analogous to (2.7), use (2.4) to verify that

$$\begin{aligned}\mathcal{L}_p^*(\beta, \mathbf{f}; \mathbf{X}, Y) = \mathcal{L}^*(\beta, \tilde{b}_{\beta,\mathbf{f}}(\mathbf{X}), \mathbf{f}; \mathbf{X}, Y) = \kappa_{\sigma^2} - \frac{1}{2\sigma^2}(Y - \mathbf{X}\beta - \mathbf{N}\mathbf{f})^\tau\mathbf{V}^{-1}(Y - \mathbf{X}\beta - \mathbf{N}\mathbf{f}) \\ +\frac{1}{2\sigma^2}tr(\mathbf{V}^{-1})\beta^\tau\mathbf{\Lambda}\beta - \frac{\lambda}{2\sigma^2}\mathbf{f}^\tau\mathbf{K}\mathbf{f}.\end{aligned} \tag{2.9}$$

Hence, maximizing (2.9) with respect to β and $\mathbf{f}$ the maximum penalized corrected likelihood estimators (MPCLE's) of β and $\mathbf{f}$ are obtained as

$$\hat{\beta} = (\mathbf{X}^\tau\mathbf{W}\mathbf{X} - tr(\mathbf{V}^{-1})\mathbf{\Lambda})^{-1}\mathbf{X}^\tau\mathbf{W}Y, \tag{2.10}$$

$$\hat{\mathbf{f}} = (\mathbf{N}^\tau \mathbf{W}_x \mathbf{N} + \lambda \mathbf{K})^{-1} \mathbf{N}^\tau \mathbf{W}_x Y, \tag{2.11}$$

where

$$\mathbf{W} = \mathbf{V}^{-1} - \mathbf{V}^{-1}\mathbf{S}\mathbf{V}^{-1}, \quad \mathbf{S} = \mathbf{N}(\mathbf{N}^\tau \mathbf{V}^{-1}\mathbf{N} + \lambda\mathbf{K})^{-1}\mathbf{N}^\tau,$$

and

$$\mathbf{W}_x = \mathbf{V}^{-1} - \mathbf{V}^{-1}\mathbf{X}(\mathbf{X}^\tau \mathbf{V}^{-1}\mathbf{X} - tr(\mathbf{V}^{-1})\mathbf{\Lambda})^{-1}\mathbf{X}^\tau \mathbf{V}^{-1}. \tag{2.12}$$

From (2.8) , the MPCLE of the random effect b is given by

$$\begin{aligned} \tilde{b} &= (\mathbf{U}^\tau \mathbf{U} + \mathbf{D}^{-1})^{-1} \mathbf{U}^\tau (Y - \mathbf{X}\hat{\beta} - \mathbf{N}\hat{\mathbf{f}}) \\ &= \mathbf{D}\mathbf{U}^\tau \mathbf{V}^{-1}(Y - \mathbf{X}\hat{\beta} - \mathbf{N}\hat{\mathbf{f}}). \end{aligned} \tag{2.13}$$

Using Equations (2.10)-(2.13), the fitted values $\hat{Y} = \mathbf{X}\hat{\beta} + \mathbf{U}\hat{b} + \mathbf{N}\hat{\mathbf{f}} = \mathbf{H}Y$ where $\mathbf{H}$ is obtained, where

$$\mathbf{H} = \mathbf{I} - \mathbf{V}^{-1} + \mathbf{V}^{-1}\mathbf{H}^*,$$

and

$$\mathbf{H}^* = (\mathbf{X} \quad \mathbf{N}) \begin{pmatrix} \mathbf{X}^\tau \mathbf{V}^{-1}\mathbf{X} - tr(\mathbf{V}^{-1})\mathbf{\Lambda} & \mathbf{X}^\tau \mathbf{V}^{-1}\mathbf{N} \\ \mathbf{N}^\tau \mathbf{V}^{-1}\mathbf{X} & \mathbf{N}^\tau \mathbf{V}^{-1}\mathbf{N} + \lambda\mathbf{K} \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{X}^\tau \\ \mathbf{N}^\tau \end{pmatrix} \mathbf{V}^{-1},$$

see Emami and Mansoori (2018).

2.3 Specialized Algorithm

Consider $\mathcal{L}^*(\beta, b, \mathbf{f}; \mathbf{X}, Y)$ in (2.7). Since $\hat{\beta}$, $\hat{\mathbf{f}}$ and $\hat{b}$ are the solutions of simultaneous equations $\frac{\partial \mathcal{L}^*}{\partial \beta} = 0$, $\frac{\partial \mathcal{L}^*}{\partial \mathbf{f}} = 0$ and $\frac{\partial \mathcal{L}^*}{\partial b} = 0$, we have

$$\begin{aligned} \begin{pmatrix} \mathbf{X}^\tau \mathbf{X} - tr(\mathbf{V}^{-1})\mathbf{\Lambda} & \mathbf{X}^\tau \mathbf{N} & \mathbf{X}^\tau \mathbf{U} \\ \mathbf{N}^\tau \mathbf{X} & \mathbf{N}^\tau \mathbf{N} + \lambda\mathbf{K} & \mathbf{N}^\tau \mathbf{U} \\ \mathbf{U}^\tau \mathbf{X} & \mathbf{U}^\tau \mathbf{N} & \mathbf{U}^\tau \mathbf{U} + \mathbf{D}^{-1} \end{pmatrix} \begin{pmatrix} \hat{\beta} \\ \hat{\mathbf{f}} \\ \hat{b} \end{pmatrix} &\triangleq \Phi \begin{pmatrix} \hat{\beta} \\ \hat{\mathbf{f}} \\ \tilde{b} \end{pmatrix} \\ &= \begin{pmatrix} \mathbf{X}^\tau Y \\ \mathbf{N}^\tau Y \\ \mathbf{U}^\tau Y \end{pmatrix}. \end{aligned} \tag{2.14}$$

These equations form the basis for the following algorithm, which as an extension of Harvill (1977) , is selected to deal with the extra measurement errors in semiparametric mixed models:

Step 0: Set $m = 0$ and choose starting values $\sigma^{2(0)}$ and $\sigma_i^{2(0)}\quad i = 1\ldots, c$.

Step 1: Calculate estimates $\hat{\beta}_c$, $\hat{\mathbf{f}}$ and $\tilde{b}_1, \ldots, \tilde{b}_c$ as the solutions to the linear equations (2.14).

Step 2: Let $\mathcal{D}_i = \theta_i \mathbf{U}_i^\tau \mathbf{V}^{-1}$, then from (2.13) we have $\tilde{b}_i = \hat{\mathcal{D}}_i(Y - \mathbf{X}\hat{\beta} - \mathbf{N}\hat{\mathbf{f}})$. Let $\mathbf{T}^*$ be the matrix formed by the last q rows and columns of Φ^{-1}, partitioned conformably with $\mathbf{D}$ as

$$\mathbf{T}^* = \Phi^{-1} = \begin{pmatrix} T^*_{11} & \cdots & T^*_{1c} \\ \vdots & \ddots & \vdots \\ T^*_{c1} & \cdots & T^*_{cc} \end{pmatrix}, \tag{2.15}$$

then calculate

$$\begin{aligned}\hat{\sigma}^{2(m+1)} = \frac{1}{n}[(Y - \mathbf{X}\hat{\beta}^{(m)} - \mathbf{N}\hat{\mathbf{f}}^{(m)})^\tau \hat{\mathbf{V}}^{-1}(Y - \mathbf{X}\hat{\beta} - \mathbf{N}\hat{\mathbf{f}}^{(m)}) \\ -tr(\hat{\mathbf{V}}^{-1})\hat{\beta}^{(m)\tau}\hat{\mathbf{\Lambda}}\hat{\beta}^{(m)} + \lambda \hat{\mathbf{f}}^{(m)\tau}\mathbf{K}\hat{\mathbf{f}}^{(m)}],\end{aligned} \tag{2.16}$$

and

$$\hat{\sigma}_{ic}^{2(m+1)} = \frac{1}{q_i - \eta_i}[\hat{b}_{ic}^{(m)\tau}\hat{b}_{ic}^{(m)} - tr(\hat{\mathcal{D}}_i^{(m)\tau}\hat{\mathcal{D}}_i^{(m)})\hat{\beta}^{(m)\tau}\hat{\mathbf{\Lambda}}\hat{\beta}^{(m)} + \lambda \hat{\mathbf{f}}^{(m)\tau}\mathbf{K}\hat{\mathbf{f}}^{(m)}], \tag{2.17}$$

where $\eta_i = tr(T^*_{ii})$ evaluated at the current estimates are penalties under the general linear mixed models and the term $tr(\hat{\mathbf{V}}^{-1})\hat{\beta}^{(m)\tau}\hat{\mathbf{\Lambda}}\hat{\beta}^{(m)}$ and $tr(\hat{\mathcal{D}}_i^{(m)\tau}\hat{\mathcal{D}}_i^{(m)})$ are the correction for the extra measurement error.

Step 3: If convergence is reached, set $\hat{\sigma}^2 = \hat{\sigma}^{2(m+1)}$ and $\hat{\sigma}_{ic}^2 = \hat{\sigma}_{ic}^{2(m+1)}$ and repeat step 1 and quit; otherwise increase m by 1 and return to step 1.

3 Asymptotic Results

In this section, the asymptotic results for the estimates are derived. It should be noted that the components of Y are not mutually independent. We assume that all the derivatives related to the likelihood exist and the parameters are identifiable.

Assumption 1: We assume that as $n \to \infty$, the following limits exist: $n^{-1}tr(\mathbf{V}^{-1})$, $n^{-1}\mathbf{Z}^\tau\mathbf{W}^2\mathbf{Z}$, $n^{-1}tr(\mathbf{V}^{-2}\mathbf{S})$ and $n^{-1}tr(\mathbf{V}^{-4}\mathbf{S}^2)$.

Lemma 3.1. *Under Assumption 1, we have*

$$\mathbf{X}^\tau\mathbf{W}\mathbf{X} = \mathbf{Z}^\tau\mathbf{W}\mathbf{Z} + tr(\mathbf{V}^{-1})\mathbf{\Lambda} + O_p(n^{\frac{1}{2}}). \tag{3.1}$$

Proof. Using (1.b), we get

$$n^{-1}(\mathbf{X}^\tau\mathbf{W}\mathbf{X} - \mathbf{Z}^\tau\mathbf{W}\mathbf{Z} - tr(\mathbf{V}^{-1})\mathbf{\Lambda}) = n^{-1}(\mathbf{Z}^\tau\mathbf{W}\Delta + \Delta^\tau\mathbf{W}\mathbf{Z} + C),$$

where $C = \Delta^\tau\mathbf{W}\Delta - tr(\mathbf{V}^{-1})\mathbf{\Lambda}$. We have also $n^{-1/2}\mathbf{Z}^\tau\mathbf{W}\Delta \sim N(0, n^{-1}\mathbf{Z}^\tau\mathbf{W}^2\mathbf{Z}\otimes\mathbf{\Lambda})$. By assumption, as $n\to\infty$, $n^{-1}\mathbf{Z}^\tau\mathbf{W}^2\mathbf{Z}$ exist. So, we have $n^{-1/2}\mathbf{Z}^\tau\mathbf{W}\Delta = O_p(n^{-1/2})$. Similarly, $n^{-1}\Delta^\tau\mathbf{W}\mathbf{Z} = O_p(n^{-1/2})$ holds. Denote the (m,n)th element of C by C_{mn}. Then

$$C_{mn} = \sum_{i=1}^{n}\sum_{j=1}^{n}\Delta_{im}\mathbf{W}_{ij}\Delta_{jn} - \sum_{i=1}^{n}\mathbf{V}^{ii}\mathbf{\Lambda}_{mn},$$

where $\Delta = (\Delta_{ij})$, $\mathbf{V} = (\mathbf{V}^{ij})$, $\mathbf{\Lambda} = (\mathbf{\Lambda}_{mn})$, $i,j = 1,2,\ldots,n$ and $m,n = 1,2,\ldots,p$. Since $\mathbf{W} = \mathbf{V}^{-1} - \mathbf{V}^{-1}\mathbf{S}\mathbf{V}^{-1}$ and $E(\Delta_{im}\Delta_{in}) = \mathbf{\Lambda}_{mn}$, we get $E(C_{mn}) = tr(\mathbf{V}^{-2}\mathbf{S})$. Further, we have

$$\begin{aligned}\mathcal{E}(C_{mn}^2) &= \sum_{i,j}\sum_{k,l}\mathcal{E}(\Delta_{im}\Delta_{jn}\Delta_{km}\Delta_{ln})\mathbf{W}_{ij}\mathbf{W}_{kl} - \mathbf{\Lambda}_{mn}^2\{tr(\mathbf{V}^{-1})\}^2\\ &= \mathbf{\Lambda}_{mm}\mathbf{\Lambda}_{nn}[tr(\mathbf{V}^{-2}) + tr(\mathbf{V}^4\mathbf{S}^2)] + \mathbf{\Lambda}_{mn}tr(\mathbf{V}^{-2}).\end{aligned}$$

By assumption, $n^{-1}tr(\mathbf{V}^{-1})$ and $tr(\mathbf{V}^{-4}\mathbf{S}^2)$ exist, so we have $\mathcal{E}(n^{-1/2}C_{mn})^2 = O(1)$ as $n\to\infty$ and $n^{-1}C = O_p(n^{-1/2})$. Combining all the above results, we get Lemma 3.1. □

Assumption 2: Assume that the following limits exist; $n^{-1}\sigma^2\mathbf{M}^{-1}tr(\mathbf{W}\mathbf{V})\mathbf{M}^{-1}$, $n^{-1}\mathbf{M}^{-1}(\mathbf{Z}\beta+\mathbf{N}\mathbf{f})^\tau\mathbf{W}(\mathbf{Z}\beta, \mathbf{N}\mathbf{f})\mathbf{\Lambda}\mathbf{M}^{-1}$ and $n^{-1}\sigma^2\mathbf{M}^{-1}(\mathbf{Z}^\tau\mathbf{W}\mathbf{V}\mathbf{W}\mathbf{Z})\mathbf{M}^{-1}$, where $\mathbf{M} = n^{-1}\mathbf{Z}^\tau\mathbf{W}\mathbf{Z}$.

Theorem 3.1. *Suppose that Assumptions 1 and 2 hold. Then $\hat{\beta}$ is an asymptotically normal estimator; that is*

$$\sqrt{n}(\hat{\beta}-\beta) \xrightarrow{d} \mathcal{N}(0, C) \qquad as \quad n\to\infty, \tag{3.2}$$

where $C = n^{-1}\sigma^2\mathbf{M}^{-1}tr(\mathbf{W}\mathbf{V})\mathbf{M}^{-1} + n^{-1}\mathbf{M}^{-1}(\mathbf{Z}\beta+\mathbf{N}\mathbf{f})^\tau\mathbf{W}(\mathbf{Z}\beta+\mathbf{N}\mathbf{f})\mathbf{\Lambda}\mathbf{M}^{-1} + n^{-1}\sigma^2\mathbf{M}^{-1}(\mathbf{Z}^\tau\mathbf{W}\mathbf{V}\mathbf{W}\mathbf{Z})\mathbf{M}^{-1}$.

Proof. Using (2.10) and (3.1), we get

$$\begin{aligned}&\sqrt{n}\hat{\beta} = \{\mathbf{M} + O_p(n^{-1/2})\}n^{-1/2}\mathbf{X}^\tau\mathbf{W}Y = \{\mathbf{I}_p + O_p(n^{-1/2})\}^{-1}\mathbf{M}^{-1}n^{-1/2}\mathbf{X}^\tau\mathbf{W}Y\\ &= \{\mathbf{I}_p + O_p(n^{-1/2})\}\mathbf{M}^{-1}n^{-1/2}\mathbf{X}^\tau\mathbf{W}Y,\end{aligned} \tag{3.3}$$

the last equality holds since $\{\mathbf{I}_p + O_p(n^{-1/2})\}^{-1} = \{\mathbf{I}_p + O_p(n^{-1/2})\}$ can be obtained from Taylor series expansion. Let $\mathbf{W}^{1/2} = \zeta\Omega\zeta^{\tau}$ denote the spectral decomposition of $\mathbf{W}^{1/2}$, where $\boldsymbol{\zeta\zeta^{\tau}} = \mathbf{I}_n$ and $\Omega = diag(\omega_1^{1/2}, \ldots, \omega_n^{1/2})$ and ω_i's are the eigenvalues of $\mathbf{W}$. Then for asymptotic properties of $\xi = n^{-1/2}\mathbf{X}^{\tau}\mathbf{W}Y$, we have

$$\xi = n^{-1/2}\mathbf{X}^{\tau}\mathbf{W}Y = n^{-1/2}\mathbf{X}^{\tau}\zeta\Omega\zeta^{\tau}\mathbf{W}^{1/2}Y = n^{-1/2}\mathbf{X}^{*\tau}\Omega Y^{*}, \tag{3.4}$$

where

$$\begin{aligned} \mathbf{X}^{*} &= \zeta^{\tau}\mathbf{X} \sim N(\boldsymbol{\zeta}^{\boldsymbol{\tau}}\mathbf{Z}, \mathbf{I}_n \otimes \boldsymbol{\Lambda}), \\ Y^{*} &= \boldsymbol{\zeta}^{\tau}\mathbf{W}^{1/2}Y \sim N(\boldsymbol{\zeta}^{\tau}\mathbf{W}^{1/2}(\mathbf{Z}\beta + \mathbf{Nf}), \sigma^2\mathbf{I}_n), \end{aligned} \tag{3.5}$$

and the jth element of ξ is given by

$$\xi_j = n^{-1/2}\sum_{i=1}^{n}\mathbf{X}_{ij}^{*}\omega^{1/2}Y_i^{*} = n^{1/2}\sum_{i=1}^{n}\vartheta_i.$$

Since ϑ_i's are independent and the limit of $var(\xi_j)$ exist as $n \to \infty$, by the central limit theorem, ξ_j are asymptotically normal. From (3.3), we have $\sqrt{n}\hat{\beta} = \mathbf{M}^{-1}\xi + O_p(n^{-1/2})$, since $\mathcal{E}(\xi) = n^{-1/2}\mathbf{Z}^{\tau}\mathbf{W}(\mathbf{Z}\beta + \mathbf{Nf}) = \mathbf{M}\sqrt{n}\beta + O_p(n^{-1/2})$. Then it follows that $\sqrt{n}(\hat{\beta} - \beta)$ is asymptotically normal with mean 0. For finding asymptotic variance of $\hat{\beta}$, we can write

$$\begin{aligned} \sqrt{n}(\hat{\beta} - \beta) &= \mathbf{M}^{-1}\xi - \mathbf{M}^{-1}\mathbf{M}\sqrt{n}\beta + O_p(n^{-1/2}) \\ &= \mathbf{M}^{-1}(\xi - \mathcal{E}(\xi)) + O_p(n^{-1/2}). \end{aligned} \tag{3.6}$$

So, we get $var(\sqrt{n}\hat{\beta}) = \mathbf{M}^{-1}var(\xi)\mathbf{M}^{-1}$. On the other hand, we have

$$\begin{aligned} var(\xi) &= \mathcal{E}^{+}[var^{*}(\xi)] + var^{+}[\mathcal{E}^{*}(\xi)] \\ &= n^{-1}\mathcal{E}^{+}(Y^{\tau}\mathbf{W}^2Y\boldsymbol{\Lambda}) + n^{-1}var^{+}(\mathbf{Z}^{\tau}\mathbf{W}Y) \\ &= n^{-1}\mathcal{E}^{+}(Y^{\tau}\mathbf{W}^2Y\boldsymbol{\Lambda}) + n^{-1}\sigma^2(\mathbf{Z}^{\tau}\mathbf{WVW}^{\tau}\mathbf{Z}) \\ &= n^{-1}\sigma^2 tr(\mathbf{W}^2\mathbf{V}) + n^{-1}(\mathbf{Z}\beta + \mathbf{Nf})^{\tau}\mathbf{W}(\mathbf{Z}\beta + \mathbf{Nf}) + n^{-1}\sigma^2\mathbf{Z}^{\tau}\mathbf{WVWZ}, \end{aligned} \tag{3.7}$$

where $\mathcal{E}^{+}$ and var^{+} are the expectation and variance with respect to Y and b. By the assumptions, all of above limits exist as $n \to \infty$ and then the proof is completed. □

Theorem 3.1 implies that MPCLE of parametric term β is consistent.

Corollary 3.1. *$\hat{\beta}$ is consistent in probability and* $\sqrt{n}(\hat{\beta} - \beta) = O_p(1)$.

Furthermore, $\hat{\beta}$ is weakly consistent with the same order of convergence in our model, since it has an extra correlated structure arising from the random effects, whereas it is strongly consistent for fixed effect models (see Nakamura (1990)).

Assumption 3: Assume that the limits $n^{-1}tr(\mathbf{Z}^{\tau}\mathbf{V}^{-1}\mathbf{Z})^{-2}$, $n^{-1}(\mathbf{Z}^{\tau}\mathbf{V}^{-1}\mathbf{Z})^{-2}$ and $n^{-1}[\mathbf{V}^{-1}\mathbf{Z}(\mathbf{Z}^{\tau}\mathbf{V}^{-1}\mathbf{Z})^{-1}\mathbf{V}^{-2}]^{\otimes 2}$, exists as $n \to \infty$, where $\mathbf{A}^{\otimes 2}$ stands for $\mathbf{A}^{\tau}\mathbf{A}$.

Theorem 3.1. *Under Assumption 3, we have*

$$\mathbf{W}_x = \mathbf{W}_z + O_p(n^{1/2}), \tag{3.8}$$

where $\mathbf{W}_z = \mathbf{V}^{-1} - \mathbf{V}^{-1}\mathbf{Z}(\mathbf{Z}^{\tau}\mathbf{V}^{-1}\mathbf{Z})^{-1}\mathbf{Z}^{\tau}\mathbf{V}^{-1}$.

Proof. From Zhong et al. (2002), we have $\mathbf{X}^{\tau}\mathbf{V}^{-1}\mathbf{X} = \mathbf{Z}^{\tau}\mathbf{V}^{-1}\mathbf{Z} + tr(\mathbf{V}^{-1})\mathbf{\Lambda} + O_p(n^{1/2})$, thus

$$\begin{aligned} & n^{-1}\{\mathbf{V}^{-1}\mathbf{X}(\mathbf{X}^{\tau}\mathbf{V}^{-1}\mathbf{X} - tr(\mathbf{V}^{-1})\mathbf{\Lambda})^{-1}\mathbf{X}^{\tau}\mathbf{V}^{-1} - \mathbf{V}^{-1}\mathbf{Z}(\mathbf{Z}^{\tau}\mathbf{V}^{-1}\mathbf{Z})^{-1}\mathbf{Z}^{\tau}\mathbf{V}^{-1}\} \\ &= n^{-1}\mathbf{V}^{-1}\mathbf{Z}(\mathbf{Z}^{\tau}\mathbf{V}^{-1}\mathbf{Z})^{-1}\tilde{\Delta} + n^{-1}\tilde{\Delta}^{\tau}(\mathbf{Z}^{\tau}\mathbf{V}^{-1}\mathbf{Z})^{-1}\mathbf{Z}^{\tau}\mathbf{V}^{-1} \\ &+ n^{-1}\tilde{\Delta}(\mathbf{Z}^{\tau}\mathbf{V}^{-1}\mathbf{Z})^{-1}\tilde{\Delta}^{\tau} + O_p(n^{-1/2}), \end{aligned} \tag{3.9}$$

where $\tilde{\Delta} = \mathbf{V}^{-1}\Delta$ follows $N(0, \mathbf{V}^{-2} \otimes \mathbf{\Lambda})$. Now since $n^{-1/2}\mathbf{V}^{-1}\mathbf{Z}(\mathbf{Z}^{\tau}\mathbf{V}^{-1}\mathbf{Z})^{-1}$ $\tilde{\Delta} \sim N(0, n^{-1}[\mathbf{V}^{-1}\mathbf{Z}(\mathbf{Z}^{\tau}\mathbf{V}^{-1}\mathbf{Z})^{-1}\mathbf{V}^{-2}]^{\otimes 2} \otimes \mathbf{\Lambda})$, by the assumption, as $n \to \infty$, $n^{-1}[\mathbf{V}^{-1}\mathbf{Z}(\mathbf{Z}^{\tau}\mathbf{V}^{-1}\mathbf{Z})^{-1}\mathbf{V}^{-2}]^{\otimes 2}$ exist. So we have $n^{-1}\mathbf{V}^{-1}\mathbf{Z}(\mathbf{Z}^{\tau}\mathbf{V}^{-1}\mathbf{Z})^{-1}\tilde{\Delta} = O_p(n^{-1/2})$.

Similarly, $n^{-1}\tilde{\Delta}^{\tau}(\mathbf{Z}^{\tau}\mathbf{V}^{-1}\mathbf{Z})^{-1}\mathbf{Z}^{\tau}\mathbf{V}^{-1} = O_p(n^{-1/2})$ holds. For the third term of (3.9), if $\mathcal{G}_{tu}$ be the (t, u)th element of $\mathcal{G} = \tilde{\Delta}(\mathbf{Z}^{\tau}\mathbf{V}^{-1}\mathbf{Z})^{-1}\tilde{\Delta}^{\tau}$., then by some calculation we have $\mathcal{E}(\mathcal{G}_{tu}) = \mathbf{\Lambda}_{t,u}tr((\mathbf{Z}^{\tau}\mathbf{V}^{-1}\mathbf{Z})^{-2})$ and $\mathcal{E}(\mathcal{G}_{tu}) = \mathbf{\Lambda}_{t,t}\mathbf{\Lambda}_{u,u}tr((\mathbf{Z}^{\tau}\mathbf{V}^{-1}\mathbf{Z})^{-2})$. By assumption, $n^{-1}(\mathbf{Z}^{\tau}\mathbf{V}^{-1}\mathbf{Z})^{-2}$ exist, so we have $n^{-1}\mathcal{G} = O_p(n^{-1/2})$. Combining all the above results, we get Theorem 3.1. □

Assumption 4: Assume that the limit $n^{-1}\mathbf{R}^{-1}\mathbf{N}^{\tau}\mathbf{W}_z\mathbf{V}\mathbf{W}_z^{\tau}\mathbf{N}\mathbf{R}^{-1}$ exists, as $n \to \infty$, where $\mathbf{R} = n^{-1}(\mathbf{N}^{\tau}\mathbf{W}_z\mathbf{N} + \lambda\mathbf{K})$.

Theorem 3.2. *Suppose that Assumptions 1 to 4 hold. The mean and variance of* $\sqrt{n}(\hat{\mathbf{f}} - \mathbf{f})$ *tends to* $\mathbf{R}^{-1}\mathbf{N}^{\tau}\mathbf{W}_z(\mathbf{Z}\beta + \mathbf{N}\mathbf{f})/\sqrt{n}$ *and* $O(1)$, *respectively. Hence,* $\hat{\mathbf{f}}$ *is a consistent estimator of* $\mathbf{f}$.

Proof. From lemma 3.1, we have

$$\begin{aligned} \sqrt{n}\hat{\mathbf{f}} &= \{n^{-1}(\mathbf{N}^{\tau}[\mathbf{W}_z + O(n^{1/2})]\mathbf{N} + \lambda\mathbf{K})\}^{-1}n^{-1/2}\mathbf{N}^{\tau}[\mathbf{W}_z + O(n^{1/2})]Y \\ &= \{n^{-1}(\mathbf{N}^{\tau}\mathbf{W}_z\mathbf{N} + \lambda\mathbf{K}) + O(n^{-1/2})\}^{-1}[n^{-1/2}\mathbf{N}^{\tau}\mathbf{W}_z + O(1)]Y \\ &= \{\mathbf{I}_q + O(n^{-1/2})\}\{n^{-1}(\mathbf{N}^{\tau}\mathbf{W}_z\mathbf{N} + \lambda\mathbf{K})\}[n^{-1/2}\mathbf{N}^{\tau}\mathbf{W}_z + O(1)]Y \\ &= \mathbf{R}^{-1}\zeta + O(n^{-1/2}), \end{aligned} \tag{3.10}$$

where $\zeta = n^{-1/2}\mathbf{N}^\tau\mathbf{W}_z Y$. Since $\mathcal{E}(\zeta) = n^{-1/2}\mathbf{N}^\tau\mathbf{W}_z(\mathbf{Z}\beta + \mathbf{N}\mathbf{f})$, it follows that $\sqrt{n}(\hat{\mathbf{f}} - \mathbf{f})$ is asymptotically normal with mean $\mathbf{R}^{-1}\mathcal{E}(\zeta)$. For the asymptotic variance, we have $var(\sqrt{n}\hat{\mathbf{f}}) = \mathbf{R}^{-1}var(\zeta)\mathbf{R}^{-1}$, where $var(\zeta) = n^{-1}\mathbf{N}^\tau\mathbf{W}_z\mathbf{V}\mathbf{W}_z^\tau\mathbf{N}\sigma^2$. □

Assumption 5: Assume that the following limits exist; $\Gamma_1 = n^{-1}(\mathbf{U}^\tau\mathbf{U} + \mathbf{D}^{-1})$, $\Gamma_2 = n^{-1}\mathbf{U}^\tau\mathbf{X}$, $\Gamma_3 = n^{-1}\mathbf{U}^\tau\mathbf{N}$ and $n^{-1}\mathbf{M}^{-1}\mathbf{N}^\tau\mathbf{W}_z\mathbf{V}\mathbf{W}^\tau\mathbf{Z}\mathbf{R}^{-1}$.

Theorem 3.3. *Under Assumptions 1-5, $\sqrt{n}(\hat{b} - b_t) = O_p(1)$ is asymptotically normally distributed with the asymptotic variance*

$$\begin{aligned} var(\sqrt{n}(\hat{b} - b_t)) = \Gamma_1^{-1}\Gamma_2 var(\sqrt{n}\hat{\beta})\Gamma_2^\tau\Gamma_1^{-1} + \Gamma_1^{-1}\Gamma_3 var(\sqrt{n}\hat{\mathbf{f}})\Gamma_3^\tau\Gamma_1^{-1} \\ +2\Gamma_1^{-1}\Gamma_2 cov(\sqrt{n}\hat{\beta}, \sqrt{n}\hat{\mathbf{f}})\Gamma_3^\tau\Gamma_1^{-1}, \end{aligned}$$

where $cov(\sqrt{n}\hat{\beta}, \sqrt{n}\hat{\mathbf{f}}) \approx n^{-1}\mathbf{M}^{-1}\mathbf{N}^\tau\mathbf{W}_z\mathbf{V}\mathbf{W}^\tau\mathbf{Z}\mathbf{R}^{-1}\sigma^2$.

Proof. From (2.7) and (2.8), we have

$$\begin{aligned} \sqrt{n}(\hat{b} - b_t) &= -(\mathbf{U}^\tau\mathbf{U} + \mathbf{D}^{-1})^{-1}\mathbf{U}^\tau\{\mathbf{X}\sqrt{n}(\hat{\beta} - \beta) + \mathbf{N}\sqrt{n}(\hat{\mathbf{f}} - \mathbf{f})\} \\ &= -\{n^{-1}(\mathbf{U}^\tau\mathbf{U} + \mathbf{D}^{-1})\}^{-1}\{n^{-1}\mathbf{U}^\tau\mathbf{Z} + O_p(n^{-1/2})\}\sqrt{n}(\hat{\beta} - \beta) \\ &\quad -\{n^{-1/2}(\mathbf{U}^\tau\mathbf{U} + \mathbf{D}^{-1})\}^{-1}n^{-1}\mathbf{U}^\tau\mathbf{N}\sqrt{n}(\hat{\mathbf{f}} - \mathbf{f}) \\ &= -\Gamma_1^{-1}\Gamma_2\sqrt{n}(\hat{\beta} - \beta) - \Gamma_1^{-1}\Gamma_3\sqrt{n}(\hat{\mathbf{f}} - \mathbf{f}) + O_p(n^{-1/2}). \end{aligned} \tag{3.11}$$

Here, we used the relation $\mathbf{U}^\tau\mathbf{X} = \mathbf{U}^\tau\mathbf{Z} + O_p(n^{1/2})$ whose proof is similar to Lemmas (3.1). Since $\hat{\beta} = \mathbf{M}^{-1}n^{-1/2}\xi + O_p(n^{-1/2})$ and $\sqrt{n}\hat{\mathbf{f}} = \mathbf{R}^{-1}\zeta + O(n^{-1/2})$ and using conditional rule, we get

$$cov(\xi, \zeta) = n^{-1}\mathbf{N}^\tau\mathbf{W}_z\mathbf{V}\mathbf{W}^\tau\mathbf{Z}.$$

So we have

$$\sqrt{n}cov(\hat{\beta}, \hat{\mathbf{f}}) \approx cov(\mathbf{M}^{-1}\xi, \mathbf{R}^{-1}\zeta) = n^{-1}\mathbf{M}^{-1}\mathbf{N}^\tau\mathbf{W}_z\mathbf{V}\mathbf{W}^\tau\mathbf{Z}\mathbf{R}^{-1}\sigma^2. \tag{3.12}$$

Using (3.11) and (3.12), the proof will be completed. □

4 Numerical Experiments

4.1 Example 1

We illustrate the performance of the estimators in the following. The response Y_{ij} is simulated from the model $Y_{ij} = z_{ij}^{(1)}\beta_1 + z_{ij}^{(2)}\beta_2 + b_{1j} + f(t_{ij}) + \epsilon_{ij}$ for $i = 1, \ldots, m$

and $j = 1, \ldots, q$. Expressing this model in the matrix form of (1.a), we write $c = 1$, $b_1 = (b_{11}, \ldots, b_{1q})^\tau$ and $\mathbf{Z}$, $\mathbf{f}$ and ϵ would be rewritten in concordance with

$$Y = (Y_{11}, \ldots, Y_{1q}, Y_{21}, \ldots, Y_{2q}, \ldots, Y_{m1}, \ldots, Y_{mq})^\tau.$$

Then for simulation, we take the following combinations: $q = 50$, $m = 3$ or 8, which are common sample sizes in longitudinal studies. We also set $\beta_1 = 1$, $\beta_2 = 2$, $z_{ij} \sim \mathcal{N}(0, 1)$,$f(t_{ij}) = sin2\pi t_{ij}$, $t_{ij} \sim \mathcal{U}(0, 1)$, $b_{1j} \sim \mathcal{N}(0, \sigma_1^2)$, $\epsilon_{ij} \sim \mathcal{N}(0, \sigma^2)$, $\sigma_1^2 = 0.5^2$, $\sigma^2 = 0.6^2$. We consider two cases: (I) with known $\mathbf{\Lambda} = diag(0.25^2, 0.25^2)$ (II): with unknown $\mathbf{\Lambda}$. To estimate $\mathbf{\Lambda}$, since there are replicated observations $\mathbf{X}_{is} = \mathbf{Z}_i + \Delta_{is}$, $s = 1, 2, \ldots, m_i$, we use the consistent unbiased estimator of $\mathbf{\Lambda}$ as

$$\hat{\mathbf{\Lambda}} = \frac{\sum_{i=1}^{n} \sum_{s=1}^{m_i} (\mathbf{X}_{is} - \bar{\mathbf{X}}_i)(\mathbf{X}_{is} - \bar{\mathbf{X}}_i)^\tau}{\sum_{i=1}^{n} (m_i - 1)}, \tag{4.1}$$

in which $\bar{\mathbf{X}}_i$ is the sample mean of the replicates. For each combination of parameters, 1000 repetitions were performed. The results of maximum penelized likelihood (MPLE) based on true value of $\mathbf{Z}$, MPCLEs and usual naive estimators ignoring measurement errors in $\mathbf{X}$, are presente in Tables 1 and 2. It can be seen from Table 2 that the estimator based on the MPCLE is obviously consistent. Besides, it can be observed that the performance of the MPCLEs for the small sample size with measurement errors are very well. The mean of estimators are close to their true values. By contrast, the naive estimators are bias. The biases cause, in turn, the error variance estimate $\hat{\sigma}^2$ to be overestimated to a rather large extent. Similarly, for the larger sample size the penalized corrected score function estimators give the nearly unbiased estimates for parameters and likewise the estimates of naive methods are biased. The normal QQ plots in Fig. 1 show that, even for small sample sizes, the distribution of the MPCLEs is close to the normal distribution. The MPCLE and naive estimators of the nonparametric component $f(.)$ of case (II) for $n = 150$ and $n = 400$ are plotted in Fig. 2. We see that the estimates of nonparametric component outperforms for naive estimator.

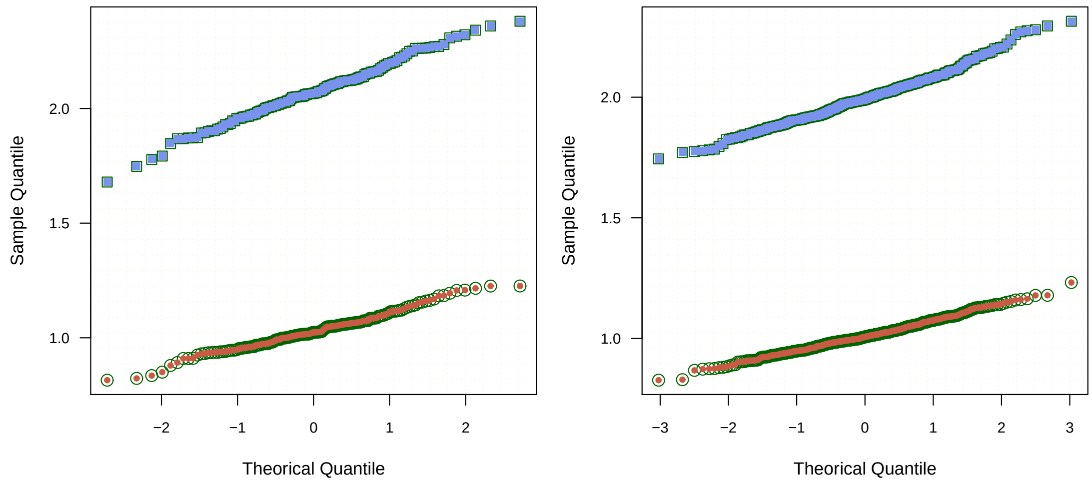

Figure 1: Normal QQ plot of $\hat{\beta}_1(\square)$ and $\hat{\beta}_2(\circ)$ for case (II). Left panel:$n = 150$, Right panel: $n = 400$.

Table 1: Mean, standard deviation (SD) and mean squared error (MSE) of each estimator from simulation study using the MPLE based on the true value Z.

Parameter	$q = 50$	$m = 3$		$q = 50$	$m = 8$	
	Mean	SD	MSE	Mean	SD	MSE
$\beta_1 = 1$	0.998	0.051	0.003	1.003	0.046	0.001
$\beta_2 = 2$	1.996	0.053	0.003	2.00	0.047	0.001
$\sigma^2 = 0.36$	0.355	0.066	0.005	0.359	0.050	0.003
$\sigma_1^2 = 0.25$	0.253	0.067	0.005	0.250	0.050	0.003
$f(t)$	–	-	2×10^{-4}	-	-	8×10^{-5}

Table 2: Mean, SD and MSE of each estimator from simulation study of case (I)-(II) using the MPCLE and naive estimator based on X

Parameter	Estimate	$q = 50$	$m = 3$		$q = 50$	$m = 8$	
		Mean	SD	MSE	Mean	SD	MSE
Case(*I*)							
$\beta_1 = 1$	MPCLE	1.002	0.049	0.002	1.000	0.010	0.001
	Naive	0.810	0.061	0.007	0.913	0.045	0.009
$\beta_2 = 2$	MPCLE	2.000	0.043	0.002	2.000	0.012	0.001
	Naive	1.641	0.072	0.007	1.963	0.051	0.008
$\sigma^2 = 0.64$	MPCLE	0.655	0.046	0.005	0.648	0.048	0.002
	Naive	0.536	0.063	0.087	0.765	0.059	0.009
$\sigma_1^2 = 0.25$	MPCLE	0.259	0.045	0.002	0.250	0.034	0.002
	Naive	0.238	0.091	0.095	0.250	0.043	0.007
$f(t)$	MPCLE	-	-	0.003	-	-	0.001
	Naive	-	-	0.015	-	-	0.009
Case(*II*)							
$\beta_1 = 1$	MPCLE	1.011	0.609	0.014	1.005	0.068	0.005
	Naive	0.699	0.093	0.052	0.796	0.053	0.044
$\beta_2 = 2$	MPCLE	2.024	0.012	0.021	2.016	0.081	0.006
	Naive	1.719	0.098	0.167	1.609	0.054	0.059
$\sigma^2 = 0.64$	MPCLE	0.661	0.216	0.0501	0.658	0.085	0.010
	Naive	0.551	0.182	0.960	0.592	0.085	0.081
$\sigma_1^2 = 0.25$	MPCLE	0.251	0.175	0.031	0.250	0.071	0.005
	Naive	0.261	0.103	0.024	0.346	0.071	0.017
$f(t)$	MPCLE	-	-	0.008	-	-	0.002
	Naive	-	-	0.028	-	-	0.019

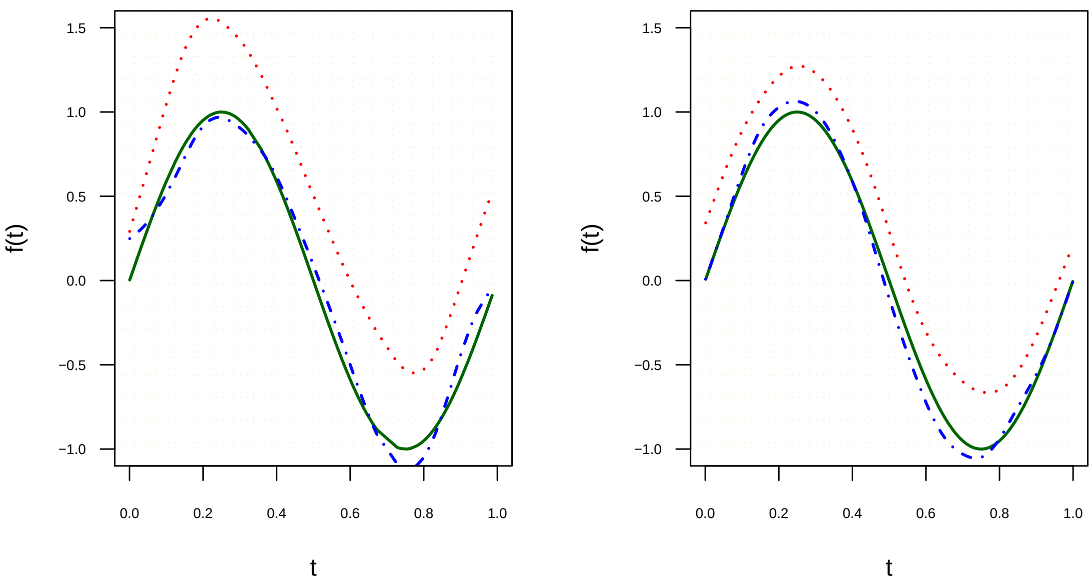

Figure 2: The estimators of the nonparametric component $f(.)$ for case(II), $sin2\pi t$ (solid curve), the proposed estimator (dash-dotted curve) and the naive estimator (dotted curve). Right panel: $n = 150$ ($m = 3$ and $q = 50$). Left panel:$n = 400$ ($m = 8$ and $q = 50$).

4.2 Example 2

Here, for smaller values of n, we explore a model that allows the covariance structure of the random effects to be correlated. Therefore, we modify the simulated model in Example 1 as follows

$$Y_{ij} = \sum_{l=1}^{8} z_{ij}^{(l)}\beta_l + b_{1j} + b_{2j}u_{ij} + f(t_{ij}) + \epsilon_{ij} \qquad i = 1,\ldots,m, j = 1,\ldots,q, \tag{4.2}$$

where β_l is the ith element of arbitrary vector $(2,-2,1,3,5,-4,-1,6)^\tau$, $(b_{1j}, b_{2j})^\tau \sim \mathcal{N}(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix})$ with $\rho = 0.75$, 0.90, 0.99, $u_{ij} = j$ for $i = 1,\ldots,m, j = 1,\ldots q$ and other setting of simulation remains unchanged. We consider $m = 5$, 8 and $q = 10$, therefore we have the sample sizes of $n = 50, 80$. For each combination of ρ and m the experiment is replicated 500 times and the average scalar mean square error (sMSE) of estimated vectors of β and b are calculated, respectively, as follows:

$$sMSE(\hat{\beta}) = \frac{1}{500}\sum_{r=1}^{500}(\hat{\beta}_{(r)} - \beta)^{\tau}(\hat{\beta}_{(r)} - \beta) \qquad sMSE(\hat{b}) = \frac{1}{500}\sum_{r=1}^{500}(\hat{b}_{(r)} - b)^{\tau}(\hat{b}_{(r)} - b), \quad (4.3)$$

where $\hat{\beta}_r$ and $\hat{b}_r$ denote the estimated parameters in the r-th simulation. The simulation results for $m = 5$ and $m = 8$ are summarized in Table 3 and 4, respectively.

Table 3: sMSE values of $\hat{\beta}$ and $\hat{b}$ and MSE value of $\hat{f}(t)$ under Case (I) and (II) with $m = 5$ and $q = 10$ ($n = 50$).

		Case (I)			Case (II)		
		$\hat{\beta}$	$\hat{b}$	$\hat{f}(t)$	$\hat{\beta}$	$\hat{b}$	$\hat{f}(t)$
$\sigma_u^2 = 0.5$	$\rho = 0.75$	0.061	0.004	0.960	0.088	0.007	1.443
	$\rho = 0.90$	0.063	0.006	1.010	0.090	0.007	1.451
	$\rho = 0.99$	0.064	0.005	1.030	0.097	0.007	1.450
$\sigma_u^2 = 0.75$	$\rho = 0.75$	0.103	0.008	1.422	0.110	0.009	1.832
	$\rho = 0.90$	0.106	0.009	1.521	0.115	0.009	1.960
	$\rho = 0.99$	0.102	0.009	1.530	0.111	0.010	1.840

Table 4: sMSE values of $\hat{\beta}$ and $\hat{b}$ and MSE value of $\hat{f}(t)$ under Case (I) and (II) with $m = 8$ and $q = 10$ ($n = 80$).

		Case (I)			Case (II)		
		$\hat{\beta}$	$\hat{b}$	$\hat{f}(t)$	$\hat{\beta}$	$\hat{b}$	$\hat{f}(t)$
$\sigma_u^2 = 0.5$	$\rho = 0.75$	0.039	0.001	0.725	0.045	0.003	0.088
	$\rho = 0.90$	0.042	0.001	0.730	0.049	0.002	0.902
	$\rho = 0.99$	0.041	0.001	0.741	0.046	0.002	1.001
$\sigma_u^2 = 0.75$	$\rho = 0.75$	0.061	0.002	1.017	0.079	0.005	1.207
	$\rho = 0.90$	0.072	0.003	1.182	0.083	0.005	1.280
	$\rho = 0.99$	0.077	0.003	1.193	0.083	0.006	1.294

From these tables, it can be seen that the performance of the discussed estimation procedure for the small sample size and almost large variance of measurement error is good. Furthermore, we see that at any level of dependency among the random effects (ρ), increasing the measurement error's variance σ_u^2 increases the sMSE ($\hat{\beta}$), sMSE ($\hat{b}$) and MSE ($\hat{f}(t)$). Besides, at any level of σ_u^2 decreasing the degree of dependency results in a decrease in MSE's.

5 The Iranian Households and Expenditure Data

In this section, to study the performance of the proposed model, we analyze Iranian urban household income and expenditure data (IUHIE)[1]. We study household members' incomes from paying jobs (on log scale) as response variable (Y), the number of employees in the household (NEH), the number of literates in the household (NLH) and the number of days they work in a week (NDW) as covarites, which are selected by stepwise regression. This data set was analyzed by Zarei et al. (2007) and Arima and Zarei (2023) in the small area estimation via linear mixed models and when some covariates are measured with error. People usually tend not to tell the truth about their income. Also, try to make sure that their lies remain hidden. Therefore, they do not tell the whole truth that might lead to finding out the real income, so we expect that there is an error in the covariates. Therefore, these data are in line with the research objectives. Similar to Zarei et al. (2007) and Arima and Zarei (2023), we deal with 1700 households sampled in 4 provinces of Iran including Lorestan, Hamedan, Tehran and Khorasan Razavi. Here, fitting a semiparametric linear mixed model with measurement error obliges us to identify the nonparametric part of the model, so we seek the added variable plots to get help. The added-variable plots can be used to detect the nonparametric component of the model and enable us to visually assess the effect of each predictor, having adjusted for the effects of the other predictors. From the added-variable plot in Fig. 3, it is reasonable to assume that the relationship between Y and NLH is nonlinear. Therefore, we can say that NEH and NDW are expressed as fixed effects; NLH's are explained as nonparametric term and, since these provinces are selected randomly from the 31 provinces, the provinces factor effect on the response is expressed as random effect. Then the model is formed as:

$$Y_{ij} = \beta_0 + \beta_1 \mathsf{NEH}_{ij} + \beta_2 \mathsf{NDW}_{ij} + \mathsf{P}_j + f(\mathsf{NLH}_{ij}) + \epsilon_{ij} \quad , \tag{5.1}$$

where $j = 1, 2, 3, 4$, $i = 1, \ldots, n_j$ and $n_j \in \{335, 236, 166, 936\}$. Since in each province we have repeated data, the estimation of non-diagonal covariance matrix of measurement

[1] Available at www.amar.org.ir/english/Statistics-by-Topic/Household-Expenditure-and-Income.

error for the fixed effects is obtained as $\hat{\mathbf{\Lambda}} = \begin{bmatrix} 0.657 & 1.647 \\ 1.647 & 12.922 \end{bmatrix}$. The smoothing parameter $\lambda = 0.461$ is selected by cross validation. Table 5 shows the MPCLEs (t-ratios) for the parametric term. As can be found in Table 5, the t-ratios of the variables NEH show that after correcting for the effect of the measurement errors, this variable is statistically significant for MPCLE; while it is not significant in MPLE . Fig. 6 shows the nonparametric estimate of the nonparametric function which is computed based the MPCLEs and MPLEs.

6 Conclusion

In this paper, we have studied the estimation of semiparametric mixed linear models when some of the covariates are measured with errors. Since the corrected score method of Nakamura (1990) provide a valuable tool for estimating in measurement errors, we constructed corrected estimators for for the parametric and nonparametric component by taking the measurement errors into account, and showed that they were consistent and asymptotically normal. The simulation study and analyzing of IUHIE data showed that the proposed estimators are also performing well in finite sample cases and with moderate sample sizes.

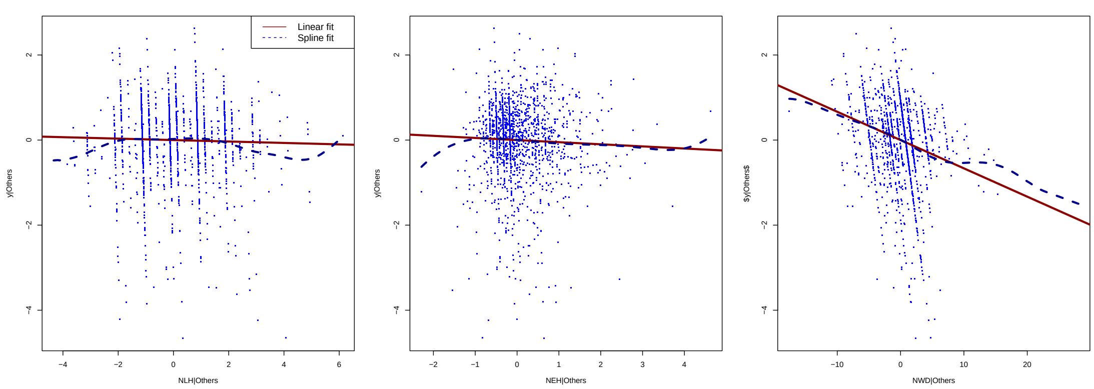

Figure 3: Added variable plot for NLH, NEH and NWD from left to right.

Table 5: MPCLEs and MPLEs for the Iranian households and expenditure data. The t-ratios are in parentheses. * Significant at 0.05 level.

Parameter	MPCLE	MPLE
β_0	16.132*(14.253)	10.918*(12.603)
β_1	0.515*(3.011)	0.259(1.072)
β_2	0.0928*(7.419)	0.106*(8.308)
σ_1^2	0.318	0.610
σ^2	0.898	1.417

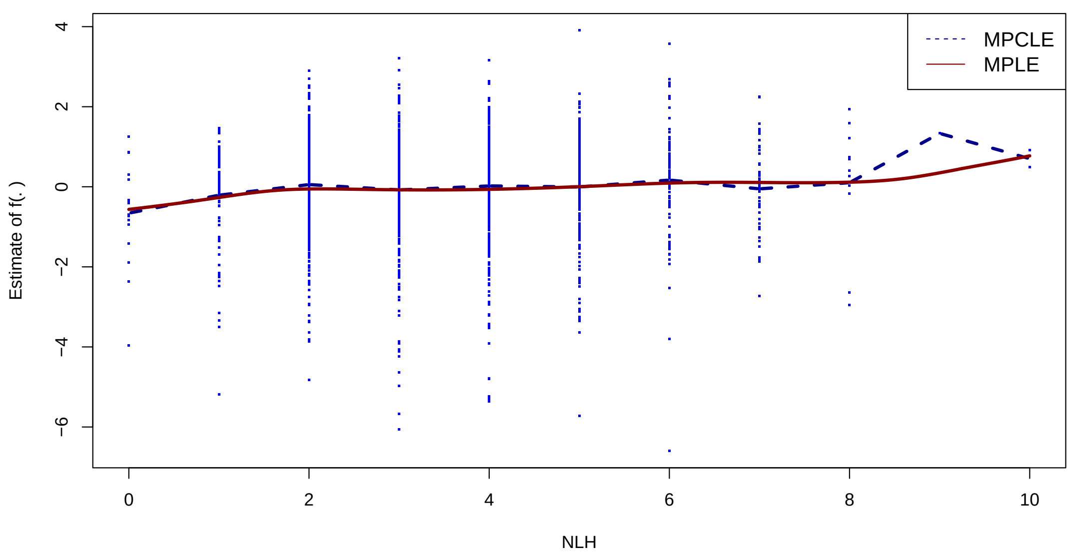

Figure 4: Estimated curve of $f(.)$ with a consideration of measurement error for the the Iranian households and expenditure data. The dots are the partial residuals ($r = \mathbf{W}_x Y$). The solid curve is the estimated $f(.)$ based MPLEs and dashed curve is the estimated $f(.)$ based MPCLEs.

Acknowledgements

We would like to thank Editor and four anonymous reviewers for their valuable comments and suggestions on the earlier version of this article which significantly improved the presentation.

References

Anderson, T. W. (1984), Estimating linear statistical relationships. *The Annals of Statistics*. **12**(1), 1-45.

Arima, S. and Zarei, S. (2023), Robust bayesian small area estimation using sub-Gaussian α-stable distribution when covariates are measured with error, *Computational Statistics*.

Belsley, DA., Kuh, E., and Welsch, RE. (1989), Regression diagnostics: identifying influential data and sources of collinearity. Wiley, New York.

Carroll, R. J., Ruppert, D. and Stefanski, L. A. (1995). *Measurement Error In Nonlinear Models*. Chapman & Hall, London.

Chen K., and Jin Z. (2006), Partial linear regression models for clustered data. *Journal of the American Statistical Association*. **101**,195-204.

Diggle PJ., Heagerty P., Liang KY., and Zeger SL. (2002), Analysis of longitudinal data. Oxford University Press; Oxford, U. K.

Emami, H., and Mansoori, M. (2018), Influence diagnostics in semiparametric linear mixed measurement error models. *Journal of Statistical Sciences*, **11**, 219-240.

Fan J, Li R. (2001), Variable selection via nonconcave penalized likelihood and its oracle properties. *Journal of the American Statistical Association*. **96**, 1348-1360.

Fung, W. K., Zhu, Z. Y., Wei, B. C., and He, X. M. (2002), Influence Diagnostics and Outlier Tests for Semiparametric Mixed Models, *Journal of Royal Statitical Society*, **64**, 565- 579.

Fuller, W. A. (1987), *Measurement error models*. John Wiley, New York.

Green, P. J., Silverman, B. W. (1994). *Nonparmetric Regression and Generalized Linear Models.* London: Chapman & Hall.

Harvill, D. A. (1977), Maximum likelihood approaches to variance component estimation and related problems (with discussion). *Journal of the American Statistical Association*, **72**, 320-340.

Kuran, O., and Yalaz, S. (2022), Kernel ridge prediction method in partially linear mixed measurement error model. *Communication in Statistics- Simulation and Computation* DOI: 10.1080/03610918.2022.2075389.

Nakamura, T. (1990), Corrected score functions for error in variables models: methodology and application to generalized linear models. *Biometrika*, **77**, 127-137.

Riquelmea,M., Bolfarine, H., and Galea, M. (2015), Robust linear functional mixed models. *Journal of Multivariate Analysis*, **134**, 82-98.

Ruppert D., Wand MP., and Carroll RJ. (2003), *Semiparametric Regression*. Cambridge University Press; Cambridge, New York.

Wu, L. (2010). *Mixed Effects Models for Complex Data*. Chapman and Hall, New York.

Yalaz, S., and Kuran, O. (2020), Kernel estimator and predictor of partially linear mixed-effect errors-in-variables model. *Journal of Statistical Computation and Simulation*, 934-951. DOI:10.1080/00949655.2020.1836642.

Yatchew, A. (2003). *Semiparametric Regression For the Applied Econometrican*. Cambridge University press, Cambridge.

Zare, K., Rasekh, A., and Rasekhi, A. (2011), Estimation of variance components in linear mixed measurement error models. *Statistical Papers*, **53**, 849-863.

Zarei, S., Gerami, A., and Jafari Khaledi, M. (2007), Comparison of Small Area Estimation of the Mean of Household's Income in Certain Provinces of Iran with Hierarchical Bayes Approach. *Journal of Statistical Research of Iran*, **4**(1), 71-90.

Zhang, D., Lin, X., Raz, J., and Sower, M. F. (1998), Semiparametric stochastic mixed models for longitudinal data. *Journal of the American Statistical Association*, **93**, 710-719.

Zhong, X. P., Fung, W. K., and Wei, B. C. (2002), Estimation in linear models with random effects and errors-in-variables. *Annals of the Institute of Statistical Mathematics*, **54**, 595-606.