Journal of Statistical Modelling: Theory and Applications
Vol. 4, No. 2, 2023, pp. 1-16
Yazd University Press 2023

JSMTA

Research Paper

Detection of outliers and influential observations in linear mixed measurement error models with Liu estimation

Shadi Borhani[1], Fatemeh Ghapani[2*] , Razieh Jafaraghaie[2]
[1]Department of Statistics, Ahvaz Branch, Islamic Azad University, Ahvaz, Iran
[2]Department of Mathematics and Statistics, Shoushtar Branch, Islamic Azad University, Shoushtar, Iran

Received: September 14, 2023/ Revised: April 16, 2024/ Accepted: May 06, 2024

Abstract: In this paper, the case deletion approach and mean shift outlier model are developed to identify influential and outlier observations using the Liu corrected likelihood estimator in linear mixed measurement error models when multicollinearity is present. We derive a corrected score test statistic for outlier detection based on mean shift outlier models. Furthermore, according to the Liu corrected likelihood estimator, several Cook's distance is constructed for influence diagnostics. A parametric bootstrap procedure is used to obtain empirical distribution of the test statistic and a simulation study is conducted to demonstrate the performance of the diagnostic criteria. Finally, a real example is provided to illustrate the performance of the test statistics.

Keywords: Case deletion model; Cook's distance; Mean shift outlier model; Outlier detection; Score test.
Mathematics Subject Classification (2010): 62J05, 62J20.

1 Introduction

A linear mixed measurement error model is written as

$$\begin{aligned} y_i &= Z_i\beta + U_i b_i + \varepsilon_i, \quad i = 1, \dots, l, \\ X_i &= Z_i + \delta_i, \end{aligned} \tag{1}$$

where y_i shows an $n_i \times 1$ vector of responses, Z_i is an $n_i \times p$ known design matrix of full column rank for the fixed effects and can be observed through the matrix X_i

*Corresponding author: `Fatemeh.Ghapani@iau.ac.ir`

with the measurement error δ_i, which δ_i is $p \times 1$ uncorrelated random vectors with $E(\delta_i) = 0$ and $Var(\delta_i) = \Lambda$. β is a $p \times 1$ vector of unknown regression coefficients, which are called fixed effects; U_i is an $n_i \times q_i$ known design matrix of the ith random effect factor and b_i is a $q_i \times 1$ vector of unobservable random effects from $N(0, \sigma_i^2 I_{q_i})$ and independent of ε_i which is an $n_i \times 1$ vector of unobservable random errors from $N(0, \sigma^2 I_{n_i})$. The variances σ^2 and σ_i^2 are called variance components. We assume that ε_i and δ_i are mutually independent. Let $y = (y_1', y_2', \ldots, y_l')'$, $n = \sum_{i=1}^{l} n_i$, $Z = (Z_1', Z_2', \ldots, Z_l')'$, $U = \oplus_{i=1}^{l} U_i$, $q = \sum_{i=1}^{l} q_i$, $b = (b_1, b_2, \ldots, b_l')'$, $\varepsilon = (\varepsilon_1', \varepsilon_2', \ldots, \varepsilon_l')'$ and $\Delta' = [\delta_1, \delta_2, \ldots, \delta_n]$. In addition it is assumed that b, ε and Δ are mutually independent. As in Zhong et al. (2002), Zare et al. (2012) and Riguelmea et al. (2015) the matrix Λ is considered to be known and if it is unknown, we can estimate it by repeated observations on the independent variables (see Nagelkerke (1992) for details of derivation of unbiased estimation of covariance matrix and Liang et al. (1999)). Let us consider the general linear mixed measurement error model as

$$
\begin{aligned}
y &= Z\beta + Ub + \varepsilon, \\
X &= Z + \Delta,
\end{aligned} \tag{2}
$$

usually assumed that $b \sim N(0, \sigma^2\Sigma)$ and $\varepsilon \sim N(0, \sigma^2 \mathrm{I}_n)$, where Σ is a block diagonal matrix with the ith block being $\gamma_i I_{q_i}$ for $\gamma_i = \frac{\sigma_i^2}{\sigma^2}$, so that y has $MN(Z\beta, \sigma^2 V)$, which $V = I_n + \sum_{i=1}^{l} \gamma_i U_i U_i' = I_n + U\Sigma U'$. The conditional distribution of $b\,|y$ is $N(\Sigma U' V^{-1}(y - Z\beta), \sigma^2 \Sigma T)$, where $T = (I_q - U'V^{-1}U\Sigma) = (I_q + U'U\Sigma)^{-1}$. If the γ_i's (and hence V) are known, the corrected likelihood estimates (CLE) of β, b and σ^2 are given by $\hat{\beta} = \left[X'V^{-1}X - tr(V^{-1})\Lambda\right]^{-1} X'V^{-1}y$, $\tilde{b} = \Sigma U'V^{-1}(y - X\hat{\beta})$ and $\hat{\sigma}^2 = n^{-1}\left[(y - X\hat{\beta})'V^{-1}(y - X\hat{\beta}) - tr(V^{-1})\hat{\beta}'\Lambda\hat{\beta}\right]$ and, respectively, in which $\hat{\sigma}_i^2 = \frac{\left[\tilde{b}_i'\tilde{b}_i - tr(\hat{D}_i'\hat{D}_i)\hat{\beta}'\Lambda\hat{\beta}\right]}{q_i - tr(T_{ii})}$ in which $\hat{D}_i = \hat{\gamma}_i U_i' V^{-1}$ for T_{ij} is ijth block of matrix T (Zhong et al., 2002; Zare et al., 2012). If the γ_i 's are unknown, the corrected score estimates are substituted back into Σ (and/or V) to obtain $\hat{\beta}$, $\hat{\sigma}^2$ and $\tilde{b}$ (Harville, 1977; Jennrich and Sampson, 1976; McCulloch and Searle, 2001; SAS, 1992; Searle et al., 1992). The existence of multicollinearity in the linear regression models leads to higher variance and instable parameter estimates in estimating linear regression models using ordinary least squares estimate. To overcome this problem, many suitable biased estimators have been developed such as the Stein estimator Stein (1956), the ridge regression Hoerl and Kennard (1970) and Liu estimator Liu. (1993). Ghapani (2022) obtained Liu estimation of parameters with additional stochastic linear restrictions for unknown parameter vectors to linear mixed models with measurement errors. The presence of outliers and influential observations in the data is complicated by the existence of multicollinearity. In linear mixed models, diagnostic methods are studied more widely by different authors including Christensen et al. (1992); Banerjee and Frees (1997); Xuping and Bocheng (1999); Haslett and Dillane (2004); Zewotir and Galpin (2005, 2007); Demidenko and Stukel (2005); Li et al. (2009). Fung et al. (2003) investigated estimation and influence diagnostics in linear mixed measurement error (LMME) models. Zare and Rasekh (2011) expressed case deletion and mean shift outlier models for detecting influential observations and outliers in LMME models using the corrected likelihood of Nakamura (1990). Also, Zare and Rasekh (2014) derived residuals and leverage in the LMME models. Recently, Maksaei et al. (2023) concentrate on diagnostic methods in LMME models with Ridge estimation using the corrected score function of Nakamura (1990). To the best of our knowledge, no attention has been paid to the in the literature concerning outlier and influential detection on the outcome of the Liu estimator in LMME models; so in this paper, we assess the diagnostic measures based on the Liu estimator by using the mean shift outlier model (MSOM) and method of approximate case deletion model (CDM) on the estimations of fixed effects, prediction of random effects and predicted values in LMME models. The paper is organized as follows. In Section 2, we derive Liu corrected likelihood estimator (LCLE) of parameters in LMME models using the corrected likelihood of Nakamura (1990). In Section 3, we introduce the diagnostic models: MSOM and CDM. We construct a corrected score test for detecting outliers. In Section 4, we develop influence measures to identify influential observations in LMME models with LCLE. Also, a parametric bootstrap procedure is used to generate the empirical distribution of test statistics. In Section 4, a simulation study has been used to show the performance of the score test statistic and cook distance. A numerical example is given in Section 5. Finally, concluding remarks are given in Section 6.

2 Liu estimation of fixed and random effects

The presence of collinearity has some destructive effects on regression analysis. In order to resolve this problem, we use the Liu estimator in linear mixed measurement error models. For this end, based on Liu. (1993) we consider the linear restriction $d\hat{\beta} = \beta + e, e \sim N(0, \sigma^2 I_p)$ where $0 < d < 1$ is the Liu biasing parameter and $\hat{\beta}$ is the corrected likelihood estimator in linear mixed measurement error model. In order to incorporate the Liu restriction in the estimation of parameters, the corrected log-likelihood function of y and the conditional corrected log-likelihood function of $b\,|y$ which denoted by $l^*(\theta, X, y)$ and $l_b^*(\theta; X, y)$, respectively, are given by

$$\begin{aligned} l^*(\theta, X, y) &= -\frac{n}{2}\log(2\pi\sigma^2) - \frac{1}{2}\log|V| - \frac{1}{2\sigma^2}[(y - X\beta)'V^{-1}(y - X\beta) \\ &\quad -tr(V^{-1})\beta'\Lambda\beta] - \frac{1}{2\sigma^2}\left[(d\hat{\beta} - \beta)'(d\hat{\beta} - \beta)\right], \\ l_b^*(\theta; X, y) &= -\frac{q}{2}\log(2\pi\sigma^2) - \frac{1}{2\sigma^2}\log|\Sigma T_d| - \frac{1}{2\sigma^2}\Big\{\left[b - \Sigma U'{}_d V_d^{-1}(y_d - X_d\beta)\right]' \\ &\quad \times(\Sigma T_d)^{-1}\left[b - \Sigma U'{}_d V_d^{-1}(y_d - X_d\beta)\right] - tr(I - V^{-1})\beta'\Lambda\beta\Big\}, \end{aligned}$$

where $\theta = (\beta, \sigma^2, \gamma, d)$. Let E^* denotes the conditional mean with respect to X given y. Then the $l^*(\theta, X, y)$ and $l_b^*(\theta; X, y)$, should satisfy

$$\begin{aligned} E^*[\frac{\partial}{\partial\beta}l^*(\beta, \sigma^2, \gamma; X, y)] &= \frac{\partial}{\partial\beta}l(\beta, \sigma^2, \gamma; Z, y), \\ E^*[\frac{\partial}{\partial\sigma^2}l_1{}^*(\sigma^2, \gamma; X, y)] &= \frac{\partial}{\partial\sigma^2}l_1(\sigma^2, \gamma; Z, y) \\ E^*[\frac{\partial}{\partial\gamma_i}l_1{}^*(\sigma^2, \gamma; X, y)] &= \frac{\partial}{\partial\gamma_i}l_1(\sigma^2, \gamma; Z, y), \end{aligned}$$

$$E^*[\frac{\partial}{\partial b} l_b{}^*(\beta, \sigma^2, \gamma; X, y)] \quad = \quad \frac{\partial}{\partial b} l_b(\beta, \sigma^2, \gamma; Z, y),$$

where $l_1(\sigma^2, \gamma; Z, y) = l(\hat{\beta}(\gamma), \sigma^2, \gamma; Z, y)$, in which $\hat{\beta} = \hat{\beta}(\gamma)$ is the maximum likelihood estimate of β and $l_1{}^*(\sigma^2, \gamma; X, y) = l^*(\hat{\beta}(\gamma), \sigma^2, \gamma; X, y)$ in which $\hat{\beta} = \hat{\beta}(\gamma)$ is the solution of the equation $\partial l^*(\beta, \sigma^2, \gamma; X, y)/\partial\beta = 0$. By solving the equations $l^*(\theta, X, y)/\partial\beta = 0$ and $l_b^*(\theta; X, y)/\partial b = 0$, the LCLE of β and b are given by

$$\begin{aligned}
\hat{\beta}_d &= (X'V^{-1}X - tr(V^{-1})\Lambda + \mathrm{I}_p)^{-1}(X'V^{-1}y + d\hat{\beta}) = A_p^{-1}(X'V^{-1}y + d\hat{\beta}), \\
\tilde{b}_d &= \Sigma U'V^{-1}(y - X\hat{\beta}_d),
\end{aligned}$$

where $A_p = (X'V^{-1}X - tr(V^{-1})\Lambda + \mathrm{I}_p)$. Also, the LCLE of σ^2 is defined as

$$\hat{\sigma}_d^2 = n^{-1}\left[(y - X\hat{\beta}_d)'V^{-1}(y - X\hat{\beta}_d) - tr(V^{-1})\hat{\beta}_d'\Lambda\hat{\beta}_d\right].$$

If the elements of γ_i are unknown, the LCLE of unknown parameters are substituted back into Σ to obtain $\hat{\beta}_d$, $\hat{\sigma}_d^2$ and $\tilde{b}_d$. According to Zare et al. (2012) we can obtain the LCLE of $\hat{\sigma}_i^2$ as

$$\hat{\sigma}_{id}^2 = \frac{\tilde{b}_{di}'\tilde{b}_{di} - tr(\hat{D}_{di}'\hat{D}_{di})\hat{\beta}_d'\Lambda\hat{\beta}_d}{[q_i - tr(T_{ii})]}, \quad i = 1, \ldots, l,$$

with $\hat{D}_{di} = \hat{\gamma}_{di}U_i'V^{-1}$.

Theorem 2.1. *$\hat{\beta}_d$ has an asymptotically normal distribution with mean $AE(\hat{\beta}_d) = G_d\beta$ and covariance matrix $AVar(\hat{\beta}_d) = G_d(Z'V^{-1}Z)^{-1}(B + \sigma^2 Z'V^{-1}Z)(Z'V^{-1}Z)^{-1}G_d$, which $B = n\sigma^2 + \beta'Z'V^{-1}Z\beta$ and $G_d = (Z'V^{-1}Z + I_p)^{-1}(Z'V^{-1}Z + dI_p)$.*

Proof. See Ghapani (2022). □

Corollary 2.2. *$\hat{\beta}$ has an asymptotically normal distribution with mean $AE(\hat{\beta}) = \beta$ and covariance matrix $AVar(\hat{\beta}) = (Z'V^{-1}Z)^{-1}(\mathrm{B} + \sigma^2 Z'V^{-1}Z)(Z'V^{-1}Z)^{-1}$.*

The advantage of adding Liu condition to the model is find a new estimator whose length is closer to β than $\hat{\beta}$. Ghapani (2022) showed that there exists an estimate of d in interval (0,1) such that $MSEM(\hat{\beta}) - MSEM(\hat{\beta}_d)$ is a positive definite matrix (Note that MSEM is an abbreviation of mean square error matrix).

3 Diagnostic methods using Liu estimator

In order to identify outlier observations in LMME models, we use MSOM. In the following subsections, we study some of these measures in the context of the LMME models with Liu estimation. We rearrange the matrices so that the ith deleted case to be in the first row. Then

$$y \quad = \quad \begin{bmatrix} y_i \\ y_{(i)} \end{bmatrix}, X = \begin{bmatrix} x'_i \\ X_{(i)} \end{bmatrix}, Z = \begin{bmatrix} z'_i \\ Z_{(i)} \end{bmatrix},$$

$$C \;=\; V^{-1} = \begin{bmatrix} c_{ii} & c'_{i(i)} \\ c_{i(i)} & V_{[i]}^{-1} + c_{i(i)}c'_{i(i)}/c_{ii} \end{bmatrix}.$$

For notational simplicity x_i' denotes the ith row of X, $X_{(i)}$ denotes matrix X with the ith row removed, c_{ii} the ith elements of the diagonal matrix V^{-1}, $c_{(i)}$ denotes the ith row of V^{-1}, $c_{(i)}$ denotes vector c with the ith element removed and c_i the ith element of c, $V_{[i]}$ denotes matrix V without the ith row and the ith column.

3.1 Mean shift outlier model

Suppose that the ith case is suspected as being an outlier, so the MSOM for the model can be represented as

$$\begin{aligned} y_i &= x'_i\beta + u'_i b + \tau + \varepsilon_i, \\ y_j &= x'_j\beta + u'_j b + \varepsilon_j, \quad j = 1, \ldots, n, j \neq i, \text{ subject to} \\ d\hat{\beta} &= \beta + e, \end{aligned} \tag{3}$$

where τ is an extra parameter to indicate the presence of an outlier in the ith.

Theorem 3.1. *For model MSOM in* (3)*, the parameter estimation is as*

$$\hat{\beta}_{md} = \hat{\beta}_d - A_p^{-1}X'c_i\hat{\tau}_{md}, \qquad \hat{\sigma}^2_{md} = \frac{n\hat{\sigma}^2_d - (\hat{v}^2_{di}/r_{pii})}{n}, \qquad \hat{\tau}_{md} = \frac{\hat{v}_{di}}{r_{pii}},$$

and $\tilde{b}_{md} \simeq \tilde{b}_d - \Sigma U' r_{pi}\frac{\hat{v}_{di}}{r_{pii}}$. *Where* $c_i{}'$ *and* r'_{pi} *are ith rows of* V^{-1} *and* $R_p = V^{-1} - V^{-1}XA_p^{-1}X'V^{-1}$*, respectively,* r_{pii} *denotes the ith diagonal elements of* R_p *and* $\hat{v}_{di} = y_i - x_i'\hat{\beta}_d - u_i'\tilde{b}_d$ *is the ith elements of residual of the model.*

Proof. The log-likelihood and the conditional log-likelihood for MSOM, respectively are given by

$$\begin{aligned} l^*_m(\theta, \tau; X, y) = &-\frac{n}{2}\log(2\pi\sigma^2) - \frac{1}{2}\log|V| \\ &-\frac{1}{2\sigma^2}\Big[(y_{(i)} - X_{(i)}\beta)'(V_{[i]}^{-1} + c_{i(i)}c'_{i(i)}\big/{}_{i(i)}c'_{i(i)}c_{ii}c_{ii})(y_{(i)} - X_{(i)}\beta) \\ &+c_{ii}(y_i - x'_i\beta - \tau)^2 + 2(y_i - x'_i\beta - \tau)c'_{i(i)}(y_{(i)} - X_{(i)}\beta) \\ &-tr(V^{-1})\beta'\Lambda\beta\Big] - \frac{1}{2\sigma^2}\left[(d\hat{\beta} - \beta)'(d\hat{\beta} - \beta)\right], \\ l^*_{bm}(\theta, \tau; X, y) = &-\frac{q}{2}\log(2\pi\sigma^2) - \frac{1}{2\sigma^2}\log|\Sigma \mathrm{T}| - \frac{1}{2\sigma^2}\Big\{b'(\Sigma T)^{-1}b \\ &-2b'(\Sigma T)^{-1}\Sigma\Big[c_{ii}(y_i - x'_i\beta - \tau)b_i + (y_i - x'_i\beta - \tau)U_i'c_{i(i)} \\ &+(y_i - X_{(i)}\beta)'\left[I - V_{[i]}^{-1} - c_{i(i)}c'_{i(i)}/c_{ii}\right](y_i - X_{(i)}\beta) \\ &+(1 - c_{ii})(y_i - x'_i\beta - \tau)^2 - 2(y_i - x'_i\beta - \tau)c'_{i(i)}(y_i - X_{(i)}\beta) \\ &-tr(I_n - V^{-1})\beta'\Lambda\beta + u_ic'_{i(i)}(y_i - X_{(i)}\beta) \end{aligned}$$

$$+U_i'(V_{[i]}^{-1} + c_{i(i)}c_{i(i)}' \Big/ {}_{i(i)}c_{i(i)}' c_{ii}c_{ii})(y_i - X_{(i)}\beta)\Big]\Big\}.$$

Equating the partial derivatives of $l_m^*(\theta, \tau; X, y)$ with respect to the elements of β and τ to zero and using $\hat{\beta}_{md}$ and $\hat{\tau}_{md}$, to denote the Liu corrected score estimator solutions gives

$$\begin{aligned}\hat{\beta}_{md} &= (X'V^{-1}X - tr(V^{-1})\Lambda + I_p)^{-1}(X'V^{-1}y - x_i\hat{\tau}_{md}) \\ &= \hat{\beta}_d - (X'V^{-1}X - tr(V^{-1})\Lambda + I_p)^{-1}X'c_i\hat{\tau}_{md} \\ &= \hat{\beta}_d - A_p^{-1}X'c_i\hat{\tau}_{md},\end{aligned}$$

and $\hat{\tau}_{md}(c_{ii} - c_i'XA_p^{-1}X'c_i) = \hat{v}_{di}$, therefore $\hat{\tau}_{md} = \frac{\hat{v}_{di}}{r_{pii}}$. Replacing $\hat{\tau}_{md}$ into $\hat{\beta}_{md}$ we obtain $\hat{\beta}_{md} = \hat{\beta}_d - A_p^{-1}X'c_i\frac{\hat{v}_{di}}{r_{pii}}$. Also, by solving the equation $l_{bm}^*(\theta, \tau; X, y)/\partial b = 0$, the Liu corrected score predictor of$b$ is given by $\tilde{b}_{md} = \Sigma U'V^{-1}(y - X\hat{\beta}_{md})$. Substituting $\hat{\beta}_{md}$ in to $\tilde{b}_{md}$ we have $\tilde{b}_{md} \simeq \tilde{b}_d - \Sigma U' r_{pi}\frac{\hat{v}_{di}}{r_{pii}}$. In addition, the LCLE of σ^2 is defined as

$$n\hat{\sigma}_{md}^2 = (y - X\hat{\beta}_{md})'V^{-1}(y - X\hat{\beta}_{md}) - (y_i - x_i\hat{\beta}_{md})^2 - tr(V^{-1})\hat{\beta}_{md}'\Lambda\hat{\beta}_{md},$$

by put $\hat{\beta}_{md}$ in above Equation, we obtain $\hat{\sigma}_{md}^2 = \frac{n\hat{\sigma}_d^2 - (\hat{v}_{di}^2 / r_{pii})}{n}$. □

Theorem 3.2. *The score test statistic for the ith observation of the MSOM for testing $H_0 : \tau = 0$ versus $H_1 : \tau \neq 0$ is defined as $SC_i(\beta) = s_i^2(1 + \frac{\hat{\beta}_{md}\Lambda\hat{\beta}_{md}}{\hat{\sigma}_{md}^2})$, where $s_i = \frac{\hat{v}_{di}}{\hat{\sigma}_v\sqrt{r_{pii}}}$ is the ith studentized residual of the model with $\hat{\sigma}_v^2 = \hat{\sigma}_d^2 + \hat{\beta}_d'\Sigma\hat{\beta}_d$.*

Proof. The score test statistic for the ith observation based on the corrected observed information matrix $J(\beta, \gamma)$ of the MSOM, is given by

$$SC_i(\beta) = \left[\frac{\partial l_m^*(\theta, \tau; X, y, r)}{\partial\tau}\right]^2 J^{\tau\tau},$$

in which $J^{\tau\tau}$ is the lower right corner of the inverse of Fisher information matrix $J(\beta, \tau)$. The corrected observed information matrix $J(\beta, \tau)$ for model (3) is defined as

$$J(\beta, \tau) = \frac{1}{\sigma^2}\begin{bmatrix} X'V^{-1}X - tr(V^{-1})\Lambda + I_p & X'c_i \\ c'_iX & c_{ii}\end{bmatrix} = \frac{1}{\sigma^2}\begin{bmatrix} A_p & X'c_i \\ c'_iX & c_{ii}\end{bmatrix}.$$

Also, we have $\frac{\partial l_m^*(\theta,\tau;X,y)}{\partial\tau} = \frac{1}{\sigma^2}c_i'(y - X\beta)$ and $J^{\tau\tau} = \frac{\sigma^2}{r_{pii}}$. Substituting $\tau = 0$, $\beta = \hat{\beta}_d$ and $\sigma^2 = \hat{\sigma}_d^2$, we have $SC_i(\beta) = \frac{(y - X\hat{\beta}_d)'c_ic'_i(y - X\hat{\beta}_d)}{\hat{\sigma}_d^2 r_{pii}} = \frac{\hat{v}_{di}^2}{\hat{\sigma}_d^2 r_{pii}} = s_i^2(1 + \frac{\hat{\beta}_d\Lambda\hat{\beta}_d}{\hat{\sigma}_d^2})$. If H_0 is rejected, then the ith case may not come from the original model and so is an outlier. □

3.2 Case deletion diagnostic

The aim of the analysis of influential observations is to evaluate the impact of the ith observations and the most common approach is case deletion diagnostics with the ith case deleted. There are a number of different statistics used by statisticians to detect influential observations in the data set. We study some of these measures in the context of LMME models for Liu estimator. The case deletion model based on Liu restriction with the ith observation deleted is defined as

$$\begin{aligned} y_{(i)} &= Z_{(i)}\beta + U_{(i)}b + \varepsilon_{(i)}, X_{(i)} = Z_{(i)} + \Delta_{(i)}, i = 1, 2, \ldots, n, \\ d\hat{\beta} &= \beta + e, \end{aligned} \tag{4}$$

The corrected log-likelihoods of and the conditional corrected log-likelihood of $b\,|y$ for model (4) are defined as

$$\begin{aligned} l_{ci}{}^{*}(\theta, X, y) &= -\frac{n-1}{2}\log(2\pi\sigma^2) - \frac{1}{2}\log\left|V_{[i]}\right| \\ &\quad -\frac{1}{2\sigma^2}[(y_{(i)} - X_{(i)}\beta)'V^{-1}(y_{(i)} - X_{(i)}\beta) - tr({V_{[i]}}^{-1})\beta'\Lambda\beta] \\ &\quad -\frac{1}{2\sigma^2}\left[(d\hat{\beta} - \beta)'(d\hat{\beta} - \beta)\right], \\ l^{*}_{bci}(\theta; X, y) &= -\frac{q}{2}\log(2\pi\sigma^2) - \frac{1}{2}[\log(\left|U'_{(i)}U_{(i)} + \Sigma^{-1}\right|^{-1})] \\ &\quad -\frac{1}{2\sigma^2}\Big\{\Big[b'[U'_{(i)}U_{(i)} + \Sigma^{-1}]b \\ &\quad -2b'[U'_{(i)}U_{(i)} + \Sigma^{-1}]^{-1}\Sigma[U'_{(i)}V^{-1}_{[i]}(y_{(i)} - X_{(i)}\beta)\Big] \\ &\quad +(y_{(i)} - X_{(i)}\beta)'(I - V^{-1}_{[i]})(y_{(i)} - X_{(i)}\beta) - tr(I - V^{-1}_{[i]})\beta'\Lambda\beta\Big\}. \end{aligned}$$

Theorem 3.3. *For model* (4), *the parameter estimation is as*

$$\hat{\beta}_{d(i)} \simeq \hat{\beta}_d - A_p^{-1}\frac{X'c_i\hat{v}_{di}}{r_{pii}}, \tilde{b}_{d(i)} \simeq \tilde{b}_d - \Sigma U' r_{pi}\frac{\hat{v}_{di}}{r_{pii}},$$

and $\hat{\sigma}^2_{d(i)} \simeq \frac{n\hat{\sigma}^2_d - (\hat{v}^2_{di}/r_{pii})}{n-1}$ *where* $\hat{\beta}_{d(i)}$, $\hat{\sigma}^2_{d(i)}$ *and* $\tilde{b}_{d(i)}$ *denote the estimates of* β, σ^2 *and* b *when the ith case is deleted, respectively.*

Proof. With differentiating $l^{*}_{ci}(\theta; X, y)$ with respect to β , we have

$$\hat{\beta}_{d(i)} = (X_{(i)}{}'V^{-1}_{[i]}X_{(i)} - tr(V^{-1}_{[i]})\Lambda + \mathrm{I}_p)^{-1}(X_{(i)}{}'V^{-1}_{[i]}y_{(i)} + d\hat{\beta}),$$

where $V^{-1}_{[i]}$ denotes inverse matrix V with the ith row and column removed. Using

$$(A + BCD)^{-1} = A^{-1} - A^{-1}B(C^{-1} + DA^{-1}B)^{-1}DA^{-1},$$

(Rao et al. (2008), Theorem A. 18), can be written as

$$\begin{aligned} \hat{\beta}_{d(i)} &= (X'V^{-1}X - \frac{X'c_i c'_i X}{c_{ii}} - tr(V^{-1})\Lambda + c_{ii}\Lambda + \mathrm{I}_p)^{-1}(X'V^{-1}y - \frac{X'c_i c'_i y}{c_{ii}} + d\hat{\beta}) \\ &= (A_p^{-1} + A_p^{-1}\frac{X'c_i}{c_{ii}}(I_n - \frac{c'_i X A_p^{-1} X' c_i}{c_{ii}})^{-1}c'_i X A_p^{-1}) \end{aligned}$$

$$
\begin{aligned}
&\times(X'V^{-1}y + d\hat{\beta} - \frac{X'c_i c'_i y}{c_{ii}}) + O_p(n^{-1}) \\
&\simeq \hat{\beta}_d - A_p^{-1}\frac{X'c_i\hat{v}_{di}}{r_{pii}}.
\end{aligned}
$$

As the same way by solving the equation $l^*_{bci}(\theta; X, y)/\partial b = 0$, the corrected score predictor of$b$is given by

$$
\begin{aligned}
\tilde{b}_{d(i)} &= \Sigma U'_{(i)} V_{[i]}^{-1}(y_{(i)} - X_{(i)}\hat{\beta}_{d(i)}) \simeq \tilde{b}_d - \Sigma U'(c_i - V^{-1}XA_p^{-1}X'c_i)\frac{\hat{v}_{di}}{r_{pii}} \\
&= \tilde{b}_d - \Sigma U' r_{pi}\frac{\hat{v}_{di}}{r_{pii}},
\end{aligned}
$$

and the LCLE of σ^2 is defined as

$$
(n-1)\hat{\sigma}^2_{d(i)} = (y_{(i)} - X_{(i)}\hat{\beta}_{d(i)})'V_{[i]}^{-1}(y_{(i)} - X_{(i)}\hat{\beta}_{d(i)}) - tr(V_{[i]}^{-1})\hat{\beta}_{d(i)}\Lambda\hat{\beta}_{d(i)}.
$$

Then, $\hat{\sigma}^2_{d(i)} \simeq \frac{n\hat{\sigma}^2_d - (\hat{v}^2_{di}/r_{pii})}{n-1}$. □

A commonly used method to detect influential observations for the fixed and random parameters is Cook's distance, which will be discussed in the following sub sections.

3.3 Generalized Cook's distance for fixed effects

To assessment changes in estimated vector of parameters, Cook's distance based on Cook (2000) for the deletion of the ith observation in LMME models with Liu estimation is defined as

$$
CD_i(\beta) = (\hat{\beta}_d - \hat{\beta}_{d(i)})'\Phi(\hat{\beta}_d - \hat{\beta}_{d(i)})',
$$

where $\Phi = \hat{\sigma}_d^{-2}\left[X'V^{-1}X - tr(V^{-1})\Lambda + \mathrm{I}_p\right]$. Then, we have

$$
CD_i(\beta) = \frac{(\hat{\beta}_d - \hat{\beta}_{d(i)})'\left[X'V^{-1}X - tr(V^{-1})\Lambda + \mathrm{I}_p\right](\hat{\beta}_d - \hat{\beta}_{d(i)})}{\hat{\sigma}^2_d},
$$

we have approximately $CD_i(\beta) = \frac{(c_{ii} - r_{pii})\hat{v}^2_{di}}{\hat{\sigma}^2_d r^2_{pii}} + O_p(n^{-1})$. Large values of $CD_i(\beta)$ indicate that observation i has a substantial effect on the full sample estimate.

3.4 Generalized Cook's distance for random effects

A convenient measure of influence for random effects in LMME models with Liu estimation based on the difference between two estimators, one includes the ith observation in the data set; the other excludes the ith observation. The Cook's distance for random effects is define as

$$
CD_i(b) = \frac{(\tilde{b}_d - \tilde{b}_{d(i)})'(U'U + \Sigma^{-1})(\tilde{b}_d - \tilde{b}_{d(i)})}{\hat{\sigma}^2_d}.
$$

Since $\tilde{b}_{d(i)} \simeq \tilde{b}_d - \Sigma U' r_{pi}\frac{\hat{v}_{di}}{r_{pii}}$, we can get $CD_i(b) = r'_{pi}(V - I_n)Vr_{pi}\frac{\hat{v}^2_{di}}{\hat{\sigma}^2_d r^2_{pii}}$. A large $CD_i(b)$ indicates that the k observations indexed by i are jointly influential on the predictions of the random effects.

3.5 Conditional Cook's distance

To investigate the influence of observations on the predicted values, we define the conditional Cook's distance similar to Tan et al. (2001) in LMME models with Liu estimation as follows

$$CD_{cond_i} = \hat{\sigma}_d^{-2}(\hat{y}_d - \hat{y}_{d(i)})'(\hat{y}_d - \hat{y}_{d(i)}),$$

which can be decomposed in to the three components

$$CD_{cond_i} = CD^1_{cond_i} + CD^2_{cond_i} + CD^3_{cond_i},$$

with

$$\begin{aligned}
CD^1_{cond_i} &= \hat{\sigma}_d^{-2}(\hat{\beta}_d - \hat{\beta}_{d(i)})'X'X(\hat{\beta}_d - \hat{\beta}_{d(i)}),\\
CD^2_{cond_i} &= \hat{\sigma}_d^{-2}(\tilde{b}_d - \tilde{b}_{d(i)})'U'U(\tilde{b}_d - \tilde{b}_{d(i)}),\\
CD^3_{cond_i} &= 2\hat{\sigma}_d^{-2}(\hat{\beta}_d - \hat{\beta}_{d(i)})'X'U(\tilde{b}_d - \tilde{b}_{d(i)}).
\end{aligned}$$

Large values of CD_{cond_i} indicate that the subset of the observations indexed by i are jointly influential on fixed effects and random effects.

3.6 Empirical distribution

To generate an empirical distribution of the test statistics under the null hypothesis that no outlier (influential) observations are exist in the data, we perform the following Algorithm (Lin et al., 1993; Rebai et al., 1994).

Algorithm 3.4. *Generate an empirical distribution of the test statistics. The algorithm is carried out in four steps:*
Step 1: We fit the model (1) *to the data and calculate LCLE of parameters. An estimate of can be derived as,* $\hat{Z}_d = X + \hat{\sigma}_v^{-2}\hat{v}_d\hat{\beta}'_d\Lambda$ *(Zare et al., 2012).*
Step 2a: Generate a new data vector as

$$\begin{aligned}
y^* &= \hat{Z}_d\hat{\beta}_d + Ub^* + \varepsilon^*,\\
X^* &= \hat{Z}_d + \Delta,
\end{aligned}$$

which Δ *is randomly generated as*$MN(0, I_n \otimes \Lambda)$*,* b^* *and* ε^* *are randomly generated as* $N(0, \hat{\sigma}^2_{1d}I_q)$ *and* $N(0,\hat{\sigma}^2_d I_n)$*, respectively.*
Step 2b: Compute the test statistics of SC_i*,* $CD_i(\beta)$*,* $CD_i(b)$ *and* CD_{cond_i} *for* $i = 1, 2, \ldots, n$ *and save the order statistics of the set.*
Step 3. Repeat steps 2a and 2b, N times, for N reasonably large. This generates an empirical distribution for each order statistic.
Step 4: Calculate the $100(1-\alpha)$ *percentile for each order statistic for the level* α *which used as a threshold for the test statistic from the original analysis. If the ith largest values of the test statistic from the original data all exceed their respective thresholds, then it is concluded that these data are all outliers (or influential) observations.*

4 Simulation study

In this section, we perform a simulation study to evaluate the performance of different diagnostic measures in terms of type I error and power of test in LMME models with LCLE. For this purpose, the jth set of simulated data is generated as

$$\begin{aligned} y_j &= Z\beta + Ub_j + \varepsilon_j, j = 1, \ldots, 1000, \\ X &= Z + \Delta, \end{aligned} \tag{5}$$

where $y_j = (y_{11j}, \ldots, y_{1mj}, y_{21j}, \ldots, y_{2mj}, \ldots, y_{q1j}, \ldots, y_{qmj})$, $b_j = (b_{1j}, b_{2j}, \ldots, b_{qj})'$, $Z = (z^{(1)}, \ldots, z^{(p)})$ and $z^{(t)} = (z_{11}^{(t)}, \ldots, z_{1m}^{(t)}, z_{21}^{(t)}, \ldots, z_{2m}^{(t)}, \ldots, z_{q1}^{(t)}, \ldots, z_{qm}^{(t)})', t = 1, \ldots, p$, ε_j is rewritten in accordance with y_j. In addition q can be interpreted as the number of independent groups; m is the group size, so $n = m \times q$ is the total size in data set. $U = I_q \otimes 1_m$ is an $n \times q$ matrix, where 1_m is $m \times 1$ vector all of whose elements of 1's. To achieve different degrees of collinearity, following McDonald and Galarneau (1975) the fixed effects variables are computed as $z_{it} = (1 - \rho^2)^{\frac{1}{2}} w_{it} + \rho w_{i,p+1}, i = 1, \ldots, n, t = 1, \ldots, p$ where w_{it} are independent standard normal pseudo-random numbers and ρ^2 represents the correlation between any two fixed effects. We consider four different values of ρ corresponding to 0.75, 0.85, 90 and 0.95. The following combinations were taken for simulation: $n = 40$ or $n = 90$, $p = 3$, $\varepsilon_{ij} \sim N(0, \sigma^2)$, $b_{kj} \sim N(0, \sigma_1^2), k = 1, \ldots, q$, $(\sigma_1^2, \sigma^2) = (0.16, 0.36)$ or $(\sigma_1^2, \sigma^2) = (0.09, 0.49)$ and $\Lambda = diag(0.05, 0.05, 0.05)$. For each set of explanatory variables the parameter vector is chosen as the eigenvector corresponding to the largest eigenvalues of $Z'V^{-1}Z$. The simulation study was conducted using the R software. For each simulated data set, diagnostic measures were calculated for the first observation. The choice of the first observation was arbitrary. Now, to generate an empirical distribution of the test statistics under the null hypothesis, the datasets were simulated as

$$\begin{aligned} y_{jh}^* &= \hat{Z}_{dj}\hat{\beta}_{dj} + Ub_{jh}^* + \varepsilon_{jh}^*, h = 1, \ldots, 1000, \\ X_{jh}^* &= \hat{Z}_{dj} + \Delta, \end{aligned}$$

where ε_{jh}^*, b_{jh}^* are randomly generated as $N(0, \hat{\sigma}_{dj}^2 I_n)$, $N(0, \hat{\sigma}_{1dj}^2 I_q)$ and Δ is randomly generated as $N(0, I_n \otimes \Lambda)$, respectively. Additionally, $\hat{\beta}_{dj}$, $\hat{Z}_{dj}$, $\hat{\sigma}_{dj}^2$ and $\hat{\sigma}_{1dj}^2$ are the LCLE of parameters from model (5). The diagnostic measures mentioned were performed for the first observation of each simulated data and $100(1 - \alpha)$ percentile from the empirical distribution of test statistics were used as threshold value of the test statistics of the model (5). The probability of a type I error estimate for the different test statistic and $\alpha = 0.05$ were calculated as the number of data sets for which the test statistic exceeded the $100(1 - \alpha)$ percentile of the empirical distribution, divided by the number of replicates.

The results are listed in Tables 1-2 . Also, in order to evaluate the relative sensitivity of the score test statistic, we introduce the shift values 1 and 3 for the first observation and again for each combination of parameters, 1000 data sets are generated from the following model:

$$\begin{aligned} y_j &= Z\beta + Ub_j + \varphi\tau + \varepsilon_j, j = 1, \ldots, 1000, \\ X &= Z + \Delta \end{aligned}$$

for $\varphi = 1$ or $\varphi = 1$ where τ is an $n \times 1$ vector with value 1 in the first element and zero elsewhere. It appears that in general the type I error of score test statistic for different

Table 1: Type of I error and power of score test statistics with different combination of parameters.

n	ρ	σ^2	σ_1^2	Sig.level	Power	
					$\varphi = 1$	$\varphi = 3$
40	0.75	0.36	0.16	0.051	0.247	0.976
		0.49	0.09	0.051	0.238	0.938
	0.85	0.36	0.16	0.049	0.217	0.952
		0.49	0.09	0.044	0.209	0.915
	0.90	0.36	0.16	0.045	0.190	0.888
		0.49	0.09	0.033	0.190	0.860
	0.95	0.36	0.16	0.043	0.154	0.627
		0.49	0.09	0.051	0.144	0.618
90	0.75	0.36	0.16	0.036	0.294	0.989
		0.49	0.09	0.058	0.242	0.967
	0.85	0.36	0.16	0.045	0.294	0.991
		0.49	0.09	0.054	0.229	0.966
	0.90	0.36	0.16	0.042	0.312	0.987
		0.49	0.09	0.057	0.264	0.964
	0.95	0.36	0.16	0.045	0.304	0.988
		0.49	0.09	0.039	0.225	0.957

combinations of parameters are close to the nominal value of . Also, with the increase of the displacement, φ , the power of the score test statistic increases in general.
In the next step, in order to assess power of the Cook's distances, we consider two cases:
(i) $z_{11} \sim U(0,5)$, $z_{12} \sim U(0,5)$ and $z_{13} = z_{12} + \upsilon$ with $\varepsilon_1 = 1.5$,
(ii) $z_{11} \sim U(5,10)$, $z_{12} \sim U(5,10)$ and $z_{13} = z_{12} + \upsilon$ with $\varepsilon_1 = 2.5$,
which are generated as points with high influence for the first observation and again for each combination of parameters, 1000 data sets are generated from the following model

$$\begin{aligned} y_j &= Z\beta + Ub_j + \varepsilon_j, j = 1, \ldots, 1000, \\ X_j &= Z + \Delta \end{aligned}$$

The parameters remained as in the evaluation of type I error. Again, for each simulated data set, we derive the Liu estimate of parameters and the different cook's distances for the first observation. The power of the cook's distance was calculated as the number of data sets for which the cook's distances exceeded the percentile $100(1-\alpha)$ of the empirical distribution, divided by the number of replicates. According to Table 3, we can see that power of the different cook's distances increase as the sample size increases. Moreover, we can see that in case (ii) power of the cook's distances increased in general.

5 Real data analysis

In this section, to evaluate the performance of the proposed diagnostic measures, we consider a sample of real data which is known as the Boston Housing data set. This data set was the basis for a 1978 paper by Harrison and Rubinfeld (1978) which discussed approaches for using housing market data to estimate the willingness to pay for clean air. Zhong et al. (2002) considered this data and used the data of 132 census tracts

Table 2: Type of I error of Cook's distance with different combination of parameters.

				Sig. level		
n	ρ	σ^2	σ_1^2	$CD_i(\beta)$	$CD_i(b)$	CD_{condi}
40	0.75	0.36	0.16	0.040	0.052	0.044
		0.49	0.09	0.044	0.056	0.047
	0.85	0.36	0.16	0.042	0.058	0.056
		0.49	0.09	0.038	0.050	0.047
	0.90	0.36	0.16	0.032	0.050	0.041
		0.49	0.09	0.026	0.035	0.029
	0.95	0.36	0.16	0.033	0.036	0.034
		0.49	0.09	0.035	0.037	0.029
90	0.75	0.36	0.16	0.039	0.048	0.048
		0.49	0.09	0.034	0.043	0.044
	0.85	0.36	0.16	0.031	0.038	0.040
		0.49	0.09	0.029	0.050	0.047
	0.90	0.36	0.16	0.025	0.039	0.039
		0.49	0.09	0.043	0.058	0.055
	0.95	0.36	0.16	0.031	0.040	0.048
		0.49	0.09	0.046	0.044	0.058

Table 3: Power of of Cook's distance with different combination of parameters.

				Power					
n	ρ	σ^2	σ_1^2	$CD_i(\beta)$		$CD_i(b)$		CD_{condi}	
				case(i)	case(ii)	case(i)	case(ii)	case(i)	case(ii)
40	0.75	0.36	0.16	0.892	0.969	0.325	0.325	0.803	0.953
		0.49	0.09	0.885	0.962	0.247	0.250	0.792	0.945
	0.85	0.36	0.16	0.866	0.955	0.299	0.329	0.769	0.934
		0.49	0.09	0.854	0.942	0.230	0.259	0.776	0.920
	0.90	0.36	0.16	0.826	0.942	0.238	0.322	0.720	0.903
		0.49	0.09	0.761	0.919	0.167	0.243	0.668	0.874
	0.95	0.36	0.16	0.672	0.854	0.159	0.196	0.525	0.732
		0.49	0.09	0.679	0.828	0.128	0.146	0.502	0.680
90	0.75	0.36	0.16	0.987	0.996	0.521	0.661	0.922	0.988
		0.49	0.09	0.958	0.990	0.334	0.478	0.890	0.982
	0.85	0.36	0.16	0.990	0.994	0.514	0.666	0.941	0.993
		0.49	0.09	0.954	0.989	0.347	0.474	0.882	0.981
	0.90	0.36	0.16	0.982	0.993	0.501	0.652	0.934	0.987
		0.49	0.09	0.939	0.980	0.350	0.454	0.881	0.967
	0.95	0.36	0.16	0.945	0.982	0.447	0.564	0.882	0.960
		0.49	0.09	0.903	0.958	0.324	0.422	0.840	0.931

within the 15 districts of the Boston city [as a part of 506 observations on census tracts in the Boston Standard Metropolitan Statistical Area (SMSA) in 1970]. The authors examined criteria to check the level of satisfaction with housing prices, with structural characteristics (such as size, age, and condition) as well as neighborhood Characteristics (such as crime rate, accessibility, and environmental factors). They used the followings fixed effects variables: Average number of rooms per dwelling (RM), proportions of owner-occupied units built prior to 1940 (AGE), weighted distances to the employment centers (DIS), blacks population proportion (B), lower status population proportion (LSTAT), crime rate (CRIM), location contiguous to the Charles River (CHAS), and level of nitrogen oxide (NOX). All fixed effects variables can be measured precisely except the pollution variable (NOX), which is taken to have measurement error. This method is often used to quantify the effects of environmental factors that affect the value of a property. Since the range of census tracts among the districts is measured with repetition, therefore, the mixed linear model should be used. More description of this data set can be found in Harrison and Rubinfeld (1978); Belsley et al. (1980). Also analysis of this data set can be found in, for example, Zare and Rasekh (2011); Zare et al. (2012); Zare and Rasekh (2014); Ghapani (2022). We apply the same data set and illustrate the use of the proposed diagnostic measures for a linear mixed model with the measurement error under Liu restriction. To investigate whether the data set is ill-conditioned, we obtain the condition number. We first estimated component variance by considering $\sigma^2 = 0.36$ and $\sigma_1^2 = 0.16$. The eigenvalues of $\hat{Z}'V^{-1}\hat{Z}$ are calculated as $\lambda_1 = 9.962542\text{e}+04$, $\lambda_2 = 1.706426\text{e}+04, ...,$ and $\lambda_9 = 2.662981\text{e}-01$. Then the condition number $\sqrt{\lambda_{\max}/\lambda_{\min}} = 611.647$ which indicate severe collinearity. The CLE and LCLE for linear mixed measurement error model are displayed in Table 4.

Table 4: The CLE and LCLE of LMM for Boston data.

Variable	CLE	LCLE	Variable	CLE	LCLE
Intercept	9.06673	9.06351	CRIM	-0.00733	-0.00733
RM	-0.00143	-0.00143	CHAS	-0.02970	-0.02974
AGE	0.00076	0.00077	NOXSQ	-0.01032	-0.01029
DIS	0.08789	0.08847	σ^2	0.02844	0.02844
B	0.45788	0.45773	σ_1^2	0.04783	0.04782
LSTAT	-0.53433	-0.53457			

Now, we derive different diagnostic measures for linear mixed measurement error model with Liu restriction given in previous section. Figure 1 give the plot of the score test and Cook's distance for fixed effects. Figure 2 shows plot of Cook's distance for random effects and plot of condition Cook's distance versus observation numbers. According to Figure 1, it is obvious that observation 36 stands out as a possible outlier with relatively large value of score test statistic. A glance at Figures 1-2 shows that observation 9 has more influence on fixed effect and observations 9, 16 and 36 have more influence on random effects. Finally, from Figure 2 we conclude that observations 9, 15 and 16 have more influence on the predicted values.

6 Discussion and conclusions

In this study, we extended mean shift outlier and case deletion models to detect outliers and influential observations under using Liu estimates. Based on the corrected likelihood, we derived a score test statistic to test whether an observation stands out as a possible outlier. In addition, we obtained analogous Cook's distance for detecting influential observations of the proposed model. Using parametric bootstrap simulation different diagnostic measures were studied in terms of type I error and test power. It was found that the type I error of the test statistics for different combinations of parameters is close to the nominal value of and with an increase in sample size the power of the test statistics increases. Furthermore, real data analysis showed that the proposed diagnostic measures successfully identified outliers and influential observations in LMME models with Liu estimates.

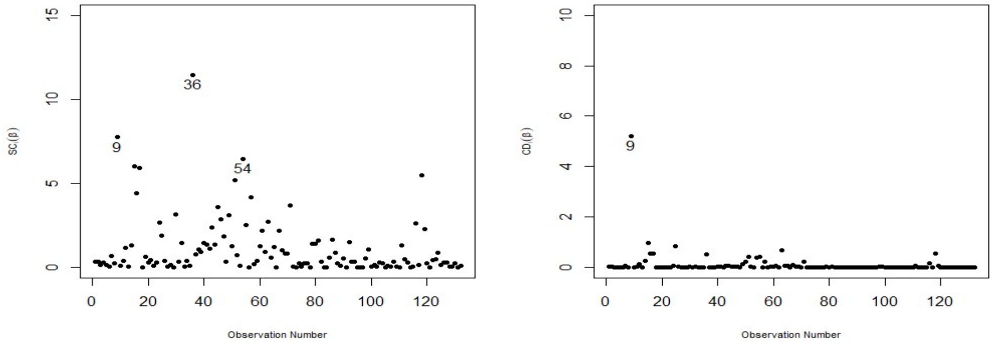

Figure 1: Plot of score test statistic and Cook's distance for fixed effects.

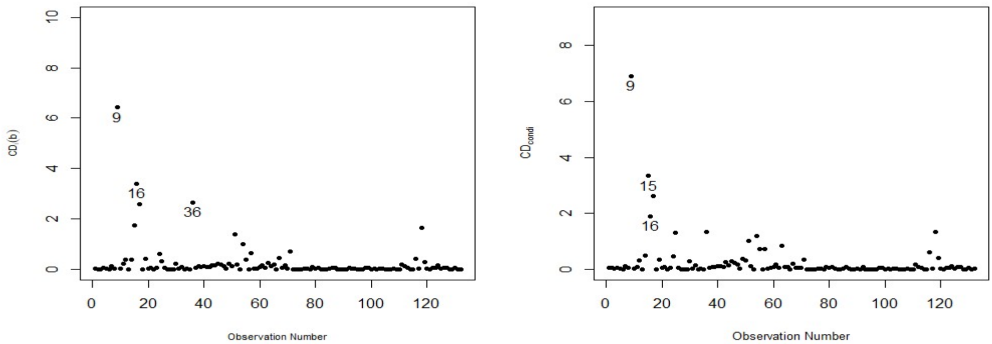

Figure 2: Plot of Cook's distance for random effects and Conditional Cook's distance.

Acknowledgement

The authors would like to thank anonymous referees for several helpful comments and suggestions, which resulted in a significant improvement in the presentation of this paper.

References

Banerjee, M. and Frees, E.W. (1997). Influence diagnostics for linear longitudinal models. *Journal of the American Statistical Association*, **92**(439):999–1005.

Belsley, D.A., Kuh, E. and Welsch, R.E. (1980). *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity.* New York: John Wiley & Sons.

Christensen, R., Pearson, L.M. and Johnson, W. (1992). Case-deletion diagnostics for mixed-models. *Technometrics* **34**:38–45.

Cook, R.D. (2000). Detection of influential observation in linear regression. *Technometrics,* **42**(1):65–68.

Demidenko, E. and Stukel, T.A. (2005). Influence analysis for linear mixed-effects models. *Statatistics in Medicine*, **24**(6):893–909.

Fung, W.K., Zhong, X.P. and Wei, B.C. (2003). On estimation and influence diagnostics in linear mixed measurement error models. *American Journal of Mathematical and Management Sciences*, **23**(1-2):37–59.

Ghapani, F. (2022). Stochastic restricted Liu estimator in linear mixed measurement error models. *Communications in Statistics - Simulation and Computation*, **51**(3):1220–1233.

Harrison, D. and Rubinfeld, D.L. (1978). Hedonic housing prices and the demand for clean air. *Journal of Environmental Economics and Management*, **5**(1):81–102.

Harville, D.A. (1977). Maximum likelihood approaches to variance component estimation and to related problems (with discussion). *Journal of the American Statistical Association*, **72**(358):320–340.

Haslett, J. and Dillane, D. (2004). Application of ‘delete= replace’to deletion diagnostics for variance component estimation in the linear mixed model. *Journal of the Royal Statistical Society, Series B*, **66**(1):131–143.

Hoerl, A.E. and Kennard, R.W. (1970). Ridge regression: biased estimation for nonorthogonal problems. *Technometrics*, **12**(1):55–67.

Jennrich, R.I. and Sampson, P.F. (1976). Newton-Raphson and related algorithms for maximum likelihood estimation of variance components. *Technometrics,* **18**(1):11–18.

Li, Z., Xu, W. and Zhu, L. (2009). Influence diagnostics and outlier tests for varying coefficient mixed models. *Journal of Multivariate Analysis*, **100**(9):2002–2017.

Liang, H., Härdle, W. and Carroll, R.J. (1999). Estimation in a semiparametric partially linear errors-in variables model. *The Annals of Statistics*, **27**(5):1519–1535.

Lin, D.Y., Wei, L.J. and Ying, Z. (1993). Checking the Cox model using cumulative sums of martingale- based residuals. *Biometrika*, **80**(3):557–572.

Liu, K. (1993). A new class of biased estimate in linear regression. *Communications in Statistics-Theory and Methods*, **22**(2):393–402.

Maksaei, N., Rasekh, A. and Babadi, B. (2023). Influence measures and outliers detection in linear mixed measurement error models with Ridge estimation. *Communications in Statistics-Simulation and Computation*, **52**(7):3062–3078.

McCulloch, C.E. and Searle, S.R. (2001). *Generalized, Linear and Mixed Models.* New York: John Wiley & Sons.

McDonald, G.C., and Galarneau, D.I. (1975). A Monte Carlo evaluation of some ridge-type estimators. *Journal of the American Statistical Association*, **70**(350):407–416.

Nagelkerke, N.J. (1992). *Maximum Likelihood Estimation of Functional Relationships.* New York: Springer.

Nakamura, T. (1990). Corrected score function for errors-in-variables models: Methodology and application to generalized linear models. *Biometrika*, **77**(1):127–137.

Rao, C.R., Toutenburg, H. Shalabh, and Heumann, C. (2008). *Linear Models and Generslizations.* Berlin: Springer.

Rebai, A., Goffinet, B. and Mangin, B. (1994). Approximate thresholds of interval mapping tests for QTL detection. *Genetics*, **138**(1):235–240.

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

SAS Institute Inc. (1992). *SAS Technical Report P-229 SAS/STAT Software: Changes and Enhancement.* Release 6.07. Cary, NC: SAS Institute.

Searle, S. R., Casella, G., McCulloch, C. E. (1992). *Variance Components.* New York: John Wiley.

Stein, C. (1956). Inadmissibility of the usual estimator for mean of multivariate normal distribution. In *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability*, Volume 1: Contributions to the Theory of Statistics (Vol. 3, pp. 197-207), University of California Press.

Tan, F.E., Ouwens, M.J. and Berger, M.P. (2001). Detection of influential observations in longitudinal mixed effects regression models. *Journal of the Royal Statistical Society, Series D*, **50**(3):271–284.

Zare, K. and Rasekh, A. (2011). Diagnostic measures for linear mixed measurement error models. *SORT-Statistics and Operations Research Transactions*, **35**(2):125–144.

Zare, K. and Rasekh, A. (2014). Residuals and leverages in the linear mixed measurement error models. *Journal of Statistical Computation and Simulation,* **84**(7):1427–1443.

Zare, K., Rasekh, A. and Rasekhi, A.A. (2012). Estimation of variance components in linear mixed measurement error models. *Statistical Papers,* **53**(4):849–863.

Zewotir, T. and Galpin, J.S. (2005). Influence diagnostics for linear mixed models. *Journal of Data Science*, **3**(2):153–177.

Zewotir, T. and Galpin, J.S. (2007). A unified approach on residuals, leverages and outliers in the linear mixed model. Test, **16**:58–75.

Xuping, Z. and Bocheng, W. (1999). Influence analysis on linear models with random effects. *Applied Mathematics-A Journal of Chinese Universities*, **14**:169–176.

Zhong, X.P., Fung, W.K. and Wei, B.C. (2002). Estimation in linear models with random effects and errors-invariables. *Annals of Applied Statistics*, **54**(3):595–606.