Some New Estimators in Linear Mixed Models with Measurement error

Bahareh Yavarizadeh, Master of Statistics

Thesis

Submitted in partial fulfillment
of the requirement for the degree of

Master of Statistics

Faculty of Mathematics and Science, Brock University
St.Catharines, Ontrio

Abstract

Linear mixed models (LMMs) are an important tool for the analysis of a broad range of structures including longitudinal data, repeated measures data (including cross-over studies), growth and dose-response curve data, clustered (or nested) data, multivariate data, and correlated data. In many practical situations, the observation of variables is subject to measurement errors, and ignoring these in data analysis can lead to inconsistent parameter estimation and invalid statistical inference. Therefore, it is necessary to extend LMMs by taking the effect of measurement errors into account. Multicollinearity and fixed-effect variables with measurement errors are two well-known problems in the analysis of linear regression models. Although there exists a large amount of research on these two problems, there is by now no single technique superior to all other techniques for the analysis of regression models when these problems are present. In this thesis, we propose two new estimators using Nakamura's approach in LMM with measurement errors to overcome multicollinearity. We consider that prior information is available on fixed and random effects. The first estimator is the new mixed ridge estimator (NMRE) and the second estimator is the weighted mixed ridge estimator (WMRE). We investigate the asymptotic properties of these proposed estimators and compare the performance of them over the other estimators using the mean square error matrix (MSEM) criterion. Finally, a data example and a Monte Carlo simulation are also provided to show the theoretical results.

Contents

List of Tables

List of Figures

Acknowledgements

I would like to express my gratitude to my supervisor, Prof. Syed Ejaz Ahmed, for the scholarship he provided, his guidance and support over these two years of my studies and research at Brock University.

My sincere thanks also extended to the External examiner Prof. Xuewen Lu, and to the Committee members Prof. Jan Vrbik, and Prof. Pouria Ramazi for their valuable comments.

I appreciate the contributions of faculty members and staff at the Department of Mathematics and Statistics and their support during my studies.

And last, but not least, I want to acknowledge my gratefulness to my husband, Yaser, for his love, understanding, and encouragement throughout my life. I would not have come this far without his support.

Chapter 1

Introduction

1.1 Linear Mixed Models

Statistical models are used for the identification of the data, the determination of the mathematical form that best represents the data structure and the definition of the relationship between the variables. Although the most widely-used statistical model is the linear regression model, it is sometimes possible to encounter with the data structures that accord with the models that include both fixed and random effects. In this case, the linear mixed models (LMMs) are appeared where provide flexibility in fitting models with various combinations of fixed and random effects. Random effects whose levels are randomly sampled from a population of levels being studied while fixed effects are variables the experimenter directly manipulates and have the same levels used when the experiment is repeated. These models are also an important tool for the analysis of a broad range of structures including longitudinal data, repeated measures data (including cross-over studies), growth and dose-response curve data, clustered (or nested) data, multivariate data, and correlated data. Inference on parameters of this model has an extensive history, see Chapter 2 of [54] for a summary. More extensive discussions of LMMs can be found in [9], [27] and [14], among others. Most of the literature on LMMs are related to the estimation of variance components, using either maximum likelihood estimation (ML), (see [25]), or restricted maximum likelihood estimation (REML) which accounts for the loss in degrees of freedom due to fitting fixed effects (see [38], [14] and [23]). However, both methods assume that each fixed and random effects are relevant.

There are different types of linear mixed models, where longitudinal model is one of them. Longitudinal data are very common in practice, either in observational studies or in experimental studies. In a longitudinal study, individuals in the study are followed over a period of time and, for each individual, data are collected at multiple time points. Following [34], a general longitudinal model can be written as:

$$y_i = Z_i\beta + U_i b_i + \varepsilon_i,\ \ i = 1,...m, \tag{1.1}$$

where i represents individuals, $y_i = (y_{i1},...,y_{im_i})'$ is the $n_i \times 1$ vector of observations from the ith individual, n_i is the number of observations on the ith individual. $Z_i = (z_{i1},...,z_{in_i})'$ is the $n_i \times p$ design matrix of fixed effect variables with the $p \times 1$ covariate vector z_{ij}, $U_i = (u_{i1},...,u_{in_i})'$ is the $n_i \times q$ design matrix corresponding to the random effects with the q × 1 covariate vector u_{ij}, $j = 1,...,n_i$, and $b_i = \left(b_{i1},...,b_{iq}\right)'$ is the q × 1 vector of random effects. $\beta = (\beta_1,...,\beta_p)'$ is the $p \times 1$ vector of fixed effects and $\varepsilon_i = (\varepsilon_{i1},...,\varepsilon_{im_i})'$ is the $n_i \times 1$ vector of errors. It is assumed that b_i and ε_i are independent with $b_i \sim N_q(0,\Sigma_i)$, and $\varepsilon_i \sim N_{n_i}\left(0,\sigma^2 W_i\right)$.

Suppose that we have a sample of $n = \sum_{i=1}^{m} n_i$ individuals where m denotes the number of individuals with response from each individual i= 1,...m measured n_i times. By stacking the vectors y_i , b_i , ε_i and matrices Z_i , U_i, the mixed model 1.1 can also be expressed as:

$$y = Z\beta + Ub_1 + \varepsilon_1, \tag{1.2}$$

where y is an $n_1 \times 1$ vector of responses, Z is an $n_1 \times p$ known design matrix for the fixed effects, β is an $p \times 1$ parameter vector of fixed effects, U is an known design matrix for the random effects, b_1 is an $q_1 \times 1$ vector of random effects and ε_1 is an $n_1 \times 1$ vector of random errors. It is assumed that b_1 and ε_1 follow independent and multivariate Gaussian distributions such that $\begin{pmatrix} b_1 \\ \varepsilon_1 \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \sigma^2 \begin{pmatrix} \Sigma_1 & 0 \\ 0 & W \end{pmatrix}\right)$. Then, $Var(y) = \sigma^2 V_1$, in which $V_1 = U\Sigma_1 U' + W$, where Σ_1 and W are known positive definite (p.d) matrices. Σ_1 is a block diagonal matrix with the ith block being $\gamma_{1i} I_{q_{1i}}$ for $\gamma_{1i} = \sigma_{1i}^2/\sigma^2,\ i = 1, \ldots, c_1$, so that $\gamma_1 = (\gamma_{11}, \ldots, \gamma_{1c_1})'$. We use maximum likelihood (ML) or restricted maximum likelihood (REML) estimation methods. To estimate the unknown parameters of the model 1.1 or 1.2, [28] developed a set of equations called mixed model equations that simultaneously yield the best linear unbiased estimator (BLUE) of $\hat{\beta}$ and the best linear unbiased predictor (BLUP) of $\hat{b}_1$ which are given, respectively, as:

$$\hat{\beta} = \left(Z' V_1^{-1} Z\right)' Z' V_1^{-1} y,$$

$$\hat{b}_1 = \Sigma_1 U' V_1^{-1} \left(y - Z\hat{\beta}\right).$$

Here, "best" in BLUE and BLUP refers to minimum variance of estimation and minimum mean squared error of prediction, (see [42]). This approach was considered by [25] to obtain the variance component estimation as:

$$\hat{\sigma}^2 = \frac{1}{n_1}\left(y - Z\hat{\beta}\right)' V_1^{-1} \left(y - Z\hat{\beta}\right),$$

$$\hat{\sigma}_{1i}^2 = \frac{\hat{b}_{1i}' \hat{b}_{1i}}{\left[q_{1i} - tr\left(T_{ii}\right)\right]},\ \ i = 1, \ldots, c_1,$$

where $\hat{b}_{1i} = \hat{D}_{1i}\left(y - Z\hat{\beta}\right)$, $\hat{D}_{1i} = \hat{\gamma}_{1i} U_i' \hat{V}^{-1}$ and T_{ij} is the ij^{th} block of matrix $T = \left(I_{q_1} + U'U\Sigma_1\right)^{-1}$.

1.2 Measurement error models

Measurement error has long been a concern in medical, health and epidemiological studies. It arises commonly in a variety of settings including longitudinal studies, case-control studies, survival data analysis and survey sampling. In nutrition studies, for instance, food frequency questionnaires are commonly used to measure diet, and it is known that this instrument involves a large degree of variation and measurement error, (see [50]). Measurement error is often present with various reasons. Sometimes covariates of interest may be difficult to observe precisely due to physical location or cost. For example, the degree of narrowing of coronary arteries may reflect risk of heart failure, but physicians may measure the degree of narrowing in carotid arteries instead due to the less invasive nature of this method of assessment. Sometimes it is impossible to measure covariates accurately due to the nature of the covariates. For example, the level of exposure to potential risk factors for cancer such as radiation can not be measured accurately, (see [47]). In other situations, a covariate may represent an average of a certain quantity over time, and any practical way of measuring such a quantity necessarily features measurement error. It is known that ignoring measurement error in variables often leads to biased results. For example, in simple linear regression with an error-contaminated covariate that is characterized by a classical additive error model, the estimate of the slope can be attenuated if ignoring error in the covariate.

Measurement error effects could be complex, generally depending on the form of the error model and the relationship between the response and covariates as well as distributions of covariates. There is an enormous literature on this subject. A textbook treatment of measurement error problems is given by [15] for linear regression, and by [5] for nonlinear models.

The classical linear regression model with one independent variable is defined by:

$$Y_i = \beta_0 + \beta_1 x_i + \varepsilon_i,\ i = 1, 2..., n \tag{1.3}$$

where $(x_1, ..., x_n)$ is fixed in repeated sampling and the random variable ε_i are independent from $N(0, \sigma^2)$. When there is measurement error, we will study regression models where one is unable to observe the following sum

$$X_i = x_i + u_i \tag{1.4}$$

where u_i is $(0, \sigma^2)$ random variable. the observed variable X_i is sometimes called the manifest variable or the indicator variable. The unobserved variable x_i is called the latent variable in certain area of application. Model with fixed x_i are called functional models, while models with random x_i are called structural models, (see [15]). It is well-documented in the literature that the maximum likelihood (ML) estimators, without taking into account the measurement error, are generally inconsistent, (see [3], [15] and [55]). Various methods have been proposed to the bias induced by the errors in variables, the unbiased score approach [55], the corrected score approach [45], the estimating equation method [1], and simulation extrapolation (SIMEX), (see [6]). A Common approach is a method in which the score function itself is corrected for measurement errors. This method is based on the corrected log-likelihood of [45]. The idea is that the conditional distribution of the corrected estimate given the true independent variables and the dependent variables are centered around the maximum likelihood estimate, which in turn is centered around the true parameter value. As one of the consequences, the consistent estimates corrected for measurement errors cannot have smaller variance than the maximum likelihood estimates without measurement errors.

The combination of random effects and measurement errors for linear models is worth investigating. As pointed out by [8], it is often the case in practice that covariate values collected on individuals are measured with non-negligible errors. [69] studied the estimation problem when fixed effects have measurement errors. Under normality assumption, they applied the corrected score approach to obtain the estimates of the regression parameters and proved the asymptotic normality, but they did not study the estimation of variance components, so their simulation results were placed in doubts. [68] obtained variance components estimation in LMMs with measurement error on fixed effect, whereas [7] discussed consistency and asymptotic normality in LMMs with measurement error in both fixed and random effects.

However, when the random effects also have errors and the covariates do not have normal distribution assumptions, [7] studied the consistency of the estimates and their asymptotic normality. For a discussion of LMMs with measurement error under the assumption of normality see, for instance, [60] and [4]. Also, [49] extended the functional mixed model proposed by [69] by assuming that each measurement error follows an elliptical distribution.

1.3 Linear mixed models with the measurement error

Consider a LMM with the measurement error in fixed effects as:

$$y = Z\beta + Ub_1 + \varepsilon_1,$$
$$X = Z + L,$$

where y is an $n_1 \times 1$ vector of observations and β is an $p \times 1$ vector of unobservable parameters, which are called fixed effects. Z and $U = [U_1|\, U_2| \ldots |U_{c_1}]$ are $n_1 \times p$ and $n_1 \times q_1$ matrices of regressor, respectively, where U_i is an $n_1 \times q_{1i}$ known design matrix of the random effects factor i, such that $q_1 = \sum_{i=1}^{c_1} q_{1i}$. Also, $b_1 = (b'_{11}, \ldots, b'_{1c_1})'$, where b_{1i} is an $q_{1i} \times 1$ vector of unobservable random effects from $N\left(0, \sigma_i^2 I_{q_{1i}}\right)$, $i = 1, \ldots, c_1$. The error term ε_1 is an $n_1 \times 1$ vector of unobservable random errors from $N\left(0, \sigma^2 I_{n_1}\right)$. X is the observed value of Z with the measurement error L, where L is a $n_1 \times p$ random matrix from $N(0, \Lambda)$ where Λ is a $p \times p$ matrix of known values. We assume that b_{1i}, ε_1 and L are mutually independent. b_1 and y are jointly distributed as $\begin{bmatrix} b_1 \\ y \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ Z\beta \end{bmatrix}, \begin{bmatrix} \sigma^2\Sigma_1 & \sigma^2\Sigma U_1' \\ \sigma^2 U\Sigma_1 & \sigma^2 V_1 \end{bmatrix}\right)$, where Σ_1 is a block diagonal matrix with the ith block being $\gamma_{1i} I_{q_{1i}}$ for $\gamma_{1i} = \sigma_{1i}^2/\sigma^2,\ i = 1, \ldots, c_1$, so that $\gamma_1 = (\gamma_{11}, \ldots, \gamma_{1c_1})'$. Therefore, y has a multivariate normal distribution with $E(y) = Z\beta$ and $Var(y) = \sigma^2 V_1$, in which $V_1 = I_{n_1} + U\Sigma_1 U'$. The conditional distribution of $b_1|y$ is $N\left(\Sigma_1 U' V_1^{-1}(y - Z\beta), \sigma^2\Sigma_1 T\right)$, where $T = \left(I_{q_1} + U'U\Sigma_1\right)^{-1}$. The log-likelihood function of y is given by:

$$\ell\left(\beta, \sigma^2, \gamma_1; Z, y\right) = \frac{-n_1}{2}\log\left(2\pi\sigma^2\right) - \frac{1}{2}\log|V_1| - \frac{1}{2\sigma^2}\left[(y - Z\beta)' V_1^{-1}(y - Z\beta)\right],$$

where $\gamma_1 = (\gamma_{11}, \ldots, \gamma_{1c_1})'$. Also, the conditional distribution of $b_1|y$ is $N\left(\Sigma_1 U' V_1^{-1}(y - Z\beta), \sigma^2\Sigma_1 T\right)$, where $T = \left(I_{q_1} + U'U\Sigma_1\right)^{-1}$. The log-likelihood function of $b_1|y$ is given by:

$$\begin{aligned}\ell_{b_1}\left(\beta, \sigma^2, \gamma_1; Z, y\right) = &\frac{-q_1}{2}\log\left(2\pi\sigma^2\right) - \frac{1}{2}\log\left(|\Sigma_1 T|\right) \\ &- \frac{1}{2\sigma^2}\left[b_1 - \Sigma_1 U' V_1^{-1}(y - Z\beta)\right]'(\Sigma_1 T)^{-1}\left[b_1 - \Sigma_1 U' V_1^{-1}(y - Z\beta)\right],\end{aligned}$$

If we simply replace Z by X without considering the measurement errors, then the estimates obtained from the score functions are not consistent in general. Various ways are proposed in dealing with measurement error models. One useful approach is based upon the corrected score method by [45]. This method proposes to find the corrected score function whose expectation with respect to the measurement error distribution coincides with the usual score function in Z.

Let $\ell_1\left(\sigma^2, \gamma_1; Z, y\right) = \ell\left(\tilde{\beta}(\gamma_1), \sigma^2, \gamma_1; Z, y\right)$, in which $\tilde{\beta} = \tilde{\beta}(\gamma_1)$ is ML estimate of β and $\ell_1^*\left(\sigma^2, \gamma_1; X, y\right) = \ell^*\left(\hat{\beta}(\gamma_1), \sigma^2, \gamma_1; X, y\right)$, in which $\hat{\beta} = \hat{\beta}(\gamma_1)$ is the solution of the equation

$\frac{\partial}{\partial\beta}\ell^*\left(\beta,\sigma^2,\gamma_1;X,y\right)=0$. Also, let E^* denotes the conditional mean with respect to X given y. This method proposes a corrected log-likelihood ℓ^* and $\ell^*_{b_1}$, which satisfies:

$$E^*\left[\frac{\partial}{\partial\beta}\ell^*\left(\beta,\sigma^2,\gamma_1;X,y\right)\right]=\frac{\partial}{\partial\beta}\ell\left(\beta,\sigma^2,\gamma_1;Z,y\right),$$

$$E^*\left[\frac{\partial}{\partial\sigma^2}\ell_1^*\left(\sigma^2,\gamma_1;X,y\right)\right]=\frac{\partial}{\partial\sigma^2}\ell_1\left(\sigma^2,\gamma_1;Z,y\right),$$

$$E^*\left[\frac{\partial}{\partial\gamma_{1i}}\ell_1^*\left(\sigma^2,\gamma_1;X,y\right)\right]=\frac{\partial}{\partial\gamma_{1i}}\ell_1\left(\sigma^2,\gamma_1;Z,y\right),\quad i=1,...,c_1,$$

$$E^*\left[\frac{\partial}{\partial b_1}\ell^*_{b_1}\left(\beta,\sigma^2,\gamma_1;X,y\right)\right]=\frac{\partial}{\partial b_1}\ell_{b_1}\left(\beta,\sigma^2,\gamma_1;Z,y\right).$$

The following equation is useful to find such a ℓ^*,

$$E\left(X'V_1^{-1}X\right)=Z'V_1^{-1}Z+tr\left(V_1^{-1}\right)\Lambda,$$

Given Λ, The corrected log-likelihoods of y and $b_1|y$ are obtained, respectively, (see [68]) as:

$$\begin{aligned}\ell^*\left(\beta,\sigma^2,\gamma_1;X,y\right)=&\frac{-n_1}{2}\log\left(2\pi\sigma^2\right)\\&-\frac{1}{2}\log|V_1|-\frac{1}{2\sigma^2}\left[(y-X\beta)'V_1^{-1}(y-X\beta)-tr\left(V_1^{-1}\right)\beta'\Lambda\beta\right],\end{aligned}$$

and

$$\begin{aligned}\ell^*_{b_1}\left(\beta,\sigma^2,\gamma_1;X,y\right)=&\frac{-q}{2}\log\left(2\pi\sigma^2\right)-\frac{1}{2}\log\left(|\Sigma_1T|\right)\\&-\frac{1}{2\sigma^2}\left[b_1-\Sigma_1U'V_1^{-1}(y-X\beta)\right]'(\Sigma_1T)^{-1}\times\left[b_1-\Sigma_1U'V_1^{-1}(y-X\beta)\right]\\&+\frac{1}{2\sigma^2}tr\left(I_{n_1}-V_1^{-1}\right)\beta'\Lambda\beta.\end{aligned}$$

The corrected log-likelihood for the estimation of variance component is

$$\ell^*\left(\beta,\sigma^2,\gamma_1;X,y\right)=\frac{-n_1}{2}\log\left(2\pi\sigma^2\right)-\frac{1}{2}\log|V_1|-\frac{1}{2\sigma^2}y'Ry,$$

where

$$R=V_1^{-1}-V_1^{-1}X\left[X'V_1^{-1}X-tr\left(V_1^{-1}\right)\Lambda\right]X'V_1^{-1},$$

and

$$y'Ry=\left(y-X\hat{\beta}\right)'V_1^{-1}\left(y-X\hat{\beta}\right)-tr\left(V_1^{-1}\right)\hat{\beta}'\Lambda\hat{\beta}.$$

If the elements of γ_1 are known, by solving the equations $\partial\ell^*\left(\beta,\sigma^2,\gamma_1;X,y\right)/\partial\beta=0$ and $\partial\ell_1^*\left(\sigma^2,\gamma_1;X,y\right)/\partial\sigma^2=0$, the corrected score estimates (CSE) of β , σ^2 will be obtained as

$$\hat{\beta}=\left[X'V_1^{-1}X-tr\left(V_1^{-1}\right)\Lambda\right]^{-1}X'V_1^{-1}y,$$

$$\hat{\sigma}^2 = \frac{1}{n_1}\left(y'V_1^{-1}y - \hat{\beta}'X'V_1^{-1}y\right),$$

and by solving the equation $\partial \ell^*_{b_1}\left(\beta, \sigma^2, \gamma_1; X, y\right) \Big/ \partial b_1 = 0$, the corrected score predictor (CSP) of b_1 will be obtained as:

$$\hat{b}_1 = \Sigma_1 U' V_1^{-1}\left(y - X\hat{\beta}\right).$$

If the elements of γ_1 are unknown, their CSE are substituted back into Σ_1 to obtain $\hat{\beta}$, $\hat{\sigma}^2$ and $\hat{b}_1$. For the CSE of $\gamma_{1i}{}'s$, one can use the CSE of $\sigma^2_{11}, \sigma^2_{12}, ..., \sigma^2_{1c_1}$ that are given in [68] as:

$$\hat{\sigma}^2_{1i} = \frac{\left[\hat{b}_{1i}\hat{b}_{1i} - tr\left(\hat{D}_i'\hat{D}_i\right)\hat{\beta}'\Lambda\hat{\beta}\right]}{\left[q_{1i} - tr\left(T_{ii}\right)\right]}, \tag{1.5}$$

where $\hat{b}_{1i} = \hat{D}_i\left(y - X\hat{\beta}\right)$, $\hat{D}_i = \hat{\gamma}_{1i}U_i'\hat{V}_1^{-1}$, $i = 1, ..., c_1$, and T_{ij} is the ij^{th} block of matrix T. An iterative algorithm (presented in Appendix) is used to compute the CSE of parameters.

1.4 Multicollinearity

When there are linear dependencies between the regressor variables, the problem of multicollinearity is said to exist. In the presence of multicollinearity, the least-squares estimates remain unbiased and efficient. The problem is that the estimated Standard error of the coefficient tends to be inflated. This Standard error has a tendency to be larger than it would be in the absence of multicollinearity, because the estimates are very sensitive to any changes in the sample observations or in the model specification. In other words, including or excluding a particular variable or certain observations may greatly change the estimated partial coefficient. If the Standard error is larger than it should be, then the t-value for testing the significance of β_i is smaller than it should be. Thus, it becomes more likely to conclude that a variable x_i is not important in a relationship when, in fact, it is important. In the existence of multicollinearity in LMMs, at least one main diagonal element of $\left(X'V^{-1}X\right)^{-1}$, may be quite large, which in view of $Var\left(\hat{\beta}\right) = \sigma^2\left(X'V^{-1}X\right)^{-1}$, means that at least one element of $\hat{\beta}$ may have a large variance, and $\hat{\beta}$ may be far from its true value. For these reasons, multicollinearity must be examined and removed. From a mathematical standpoint, near multicollinearity makes the $X'X$ matrix ill-conditioned (with X the data matrix), that is the value of its determinant is nearly zero. Thus, attempts to calculate the inverse of the matrix result in numerical snags with uncertain final values. Exact multicollinearity occurs when at least one of the regressor variables predictors is a linear combination of other regressor variables. Therefore, X is not a full rank matrix, the determinant of X is exactly zero, and inverting $X'X$ is not simply difficult, it does not exist.

In literature, there are various multicollinearity diagnostics. Several criteria have been put forth to detect multicollinearity problems. [10] suggested the following: 1: Check if any regression coefficients have the wrong sign, based on prior knowledge. 2: Check if covariates anticipated to be important based on prior knowledge have regression coefficients with small t-statistics. 3: Check if deletion of a row or a column of the X matrix produces a large change in the fitted model. 4: Check the correlations between all pairs of predictor variables to determine if any are unexpectedly high. 5: Examine the variance inflation factor (VIF). The VIF of X_i is given by $VIF = \frac{1}{1-R_{il}^2}$, where R_i^2 is the squared multiple correlation coefficient resulting from the regression of X_i against all other independent variables. If x_i has a strong linear relation with other independent variables, then R_i^2 will be close to one and VIF values tend to be very high. However, in the absence of any linear relation among independent variables, R_i^2 will be zero and the VIF will equal one. It is known that a VIF value greater than one indicates deviation from orthogonality and has tendencies to collinearity. [43] suggest that VIF greater than 10 indicates multicollinearity. 6: The condition number is also used to detect multicollinearity. If X has a full rank, then the condition number can also be wriitten as: $\kappa(x) = \sqrt{\frac{\lambda_{max}}{\lambda_{min}}}$, where λ_{max} and λ_{min} are the largest and smallest eigenvalues of $X'X$. Multicollinearity diagnostics used in linear regression model were extended to linear mixed models by [51]. [21] gave an efficient computational strategy for calculating covariates and their standard errors in LMMs and also considered estimation and estimability for the case when the fixed effects design matrix is not of full rank.

To deal with such instability, different remedial actions have been proposed. One of the most important estimation techniques is to consider biased estimators, such as the Stein estimator [56], the ordinary ridge regression (ORR) estimator [29], and the Liu estimator [39]. [12] obtained ridge predictors in LMMs by using longitudinal data. [40] introduced ridge method to the problem and put forward the ridge predictors in LMMs. The ridge estimator and ridge predictor are derived in the context of Henderson's mixed model equations by [73]. Furthermore, [74] proposed Liu estimator/predictor with the help of the penalized log-likelihood approach to reduce the effects of multicollinearity in LMMs. An alternative technique to combat multicollinearity is to consider parameter estimation with some restrictions on the unknown parameters, which can be an exact or stochastic restriction, (see [63]). However, exact restrictions are often discomforting in many applied work such as economic relations, industrial structures, production planning, and so on. While, as pointed out by [2] using stochastic linear restriction, one can accomplish an examination and analysis of one's own thoughts and feelings (prior information via introspection). In addition, one may also have prior information from a previous sample which usually makes some relations through stochastic subspace restrictions. Therefore, we deal with stochastic restrictions on the parameter vector in this thesis. Some relevant references in the linear model are [40], [26], [72], [70], and so on. Also, there is an ongoing debate in the statistical literature about prior information for LMMs. [37] considered LMMs for longitudinal data under exact linear restrictions and found the estimators for the parameters of interest. [62] studied the prediction problems in LMMs with stochastic linear restrictions on fixed effects. They assumed that the variance parameters of the model were known and their predictors were not examined when the variance parameters of the model were unknown. Recently, [32] derived the estimation of variance parameters in LMMs under stochastic linear restrictions on fixed effects.

As the third solution, considering the combination of two different estimators may inherit the advantages of both estimators, many better estimators are introduced. See [22], [69], and [36], [30], [63], [67], for examples using linear models, also, see [62], and [33] for examples involving LMMs. However, these authors do not consider measurement errors in their studies.

See [20] and [18], for example using linear models with the measurement error, [65] and [17] for example involving LMMs with the measurement error.

[35] proposed a new mixed ridge estimator (NMRE) in the linear model by combining the mixed estimator (ME) and ridge estimator (RE). [18] introduced a NMRE in linear models with the measurement error. Motivated by this, our primary aim in chapter 2 is to obtain a NMRE in LMMs with the measurement error on fixed effects when stochastic linear restrictions are available on fixed and random effects to overcome the multicollinearity problem. [11], [58] and [57] added some stochastic linear restrictions into a sample model and introduced the ordinary mixed estimator (OME) for the regression coefficient vector. Based on the fact that the prior information and the sample information are not equally important, [52] introduced the weighted mixed estimator. [36] proposed the weighted mixed ridge estimator (WMRE) based on ridge estimator. [19] considered the measurement error in linear models and introduced the weighted mixed ridge estimator (WMRE) for regression coefficients. Motivated by this, our primary aim in Chapter 3 in this thesis is to obtain a new ridge-type estimator for LMMs with measurement error on fixed effects, called the weighted mixed ridge estimator (WMRE), when stochastic linear restrictions are available on both fixed and random effects and multicollinearity is present. We consider a case where variance parameters are not known and get the variance parameter estimations.

1.5 The Goal of this thesis

Mixed modeling has become a major area of statistical research including many disciplines where a wide variety of data structures are often encountered by statisticians. The observations in any statistical analysis are assumed to be recorded without any measurement errors. However, this assumption is not satisfied in many applications and the observations are recorded with measurement errors. The presence of measurement errors in the observations disturb the optimal properties of estimators, which are meant for those situations where observations are recorded free of measurement error. Therefore, it is necessary to extend LMMs by taking the effect of measurement errors into account. There are a few methods that can handle measurement errors. We apply the corrected score function of [45] to study LMMs with the measurement error.

To avoid this problem, In Chapter 2, we will propose a new ridge type estimator, called a new mixed ridge estimator/predictor (NMRE/P) in LMMs with the measurement error when the stochastic linear restrictions are available on fixed and random effect and the fixed effect variables are multicollinear. Then, we will derive the asymptotic normality properties of some estimators and compare the estimators in terms of the mean squared error matrix (MSEM) criterion. Finally, the theoretical findings of the proposed estimator are illustrated using a data example and a Monte Carlo simulation.

In chapter 3, we will obtain the weighted mixed estimator/predictor (WME/P) by unifying the sample and prior information when additional stochastic linear restrictions are available on fixed and random effects. Then by adding the ridge method, we will propose a new ridge-type estimator/predictor called the weighted mixed ridge estimator/predictor (WMRE/P) in a LMM with the measurement error on fixed effects. Furthermore, we will study the asymptotic normality properties of some estimators and compare the estimators/predictors based on the mean square matrix (MSEM) criterion. A data example and simulation study are also provided to illustrate the theoretical results.

In chapter 4, some conclusions are also given, the result are summarized and for future work, an outlook is also given.

Chapter 2

The new mixed ridge estimator for linear mixed models with measurement error under stochastic linear mixed restrictions

2.1 Introduction

Consider a linear mixed model with the measurement error in fixed effects as:

$$y = Z\beta + Ub_1 + \varepsilon_1,\ X = Z + L, \tag{2.1}$$

where y is an $n_1 \times 1$ vector of observations, $U = \oplus_{i=1}^{c_1} U_i$ is an $n_1 \times q_{1i}$ regressor matrix corresponding to the random effects with U_i an $n_1 \times q_{1i}$ known design matrix for the ith random effects factor, such that $q_1 = \sum_{i=1}^{c_1} q_{1i}$. Z is an $n_1 \times p$ known design matrix of full column rank for the fixed effects and can be observed through the matrix X with the measurement error L, where L is an $n_1 \times p$ random matrix from $N(0, I_{n_1} \otimes \Lambda)$. It is assumed that Λ is an $p \times p$ matrix of known values. β is an $p \times 1$ vector of unknown fixed effects parameters and $b_1 = (b'_{11}, ..., b'_{1c_1})'$ is an $q_1 \times 1$ vector of unobservable random effects from $N\left(0, \sigma^2 \Sigma_1\right)$, where Σ_1 is a block diagonal matrix with the i^{th} block being $\gamma_{1i} I_{q_{1i}}$ for $\gamma_{1i} = \sigma_{1i}^2 / \sigma^2$, $i = 1, ..., c_1$. The error term ε_1 is an $n_1 \times 1$ vector of unobservable random errors from $N\left(0, \sigma^2 I_{n_1}\right)$. It is assumed that b_1, ε_1 and L are mutually independent. So, y has a multivariate normal distribution with $E(y) = Z\beta$ and $Var(y) = \sigma^2 V_1$, in which $V_1 = I_{n_1} + U\Sigma_1 U'$. The conditional distribution of $b_1|y$ is $N\left(\Sigma_1 U' V_1^{-1}(y - Z\beta), \sigma^2 \Sigma_1 T\right)$, where $T = \left(I_{q_1} + U'U\Sigma_1\right)^{-1}$. The corrected score estimates (CSE) of β and σ^2 are given in [69] and [68], as $\hat{\beta} = \left[X'V_1^{-1}X - tr\left(V_1^{-1}\right)\Lambda\right]^{-1} X'V_1^{-1}y$, $\hat{\sigma}^2 = \frac{1}{n_1}\left(y'V_1^{-1}y - \hat{\beta}'X'V_1^{-1}y\right)$ where $\hat{\sigma}_{1i}^2 = \frac{\left[\hat{b}'_{1i}\hat{b}_{1i} - tr\left(\hat{D}'_i\hat{D}_i\right)\hat{\beta}'\Lambda\hat{\beta}\right]}{\left[q_{1i} - tr(T_{ii})\right]}$, $i = 1, ..., c_1$, in which $\hat{D}_i = \hat{\gamma}_{1i} U'_i \hat{V}_1^{-1}$, and $\hat{b}_1 = \Sigma_1 U' V_1^{-1}\left(y - X\hat{\beta}\right)$ is a random effect prediction and T_{ij} is ij^{th} block of matrix T.

Generally, the variables of the fixed effects design matrix are assumed to be independent. Nonetheless, among the variables of fixed effects design matrix may arise strong or near to strong linear relationships in practice. In this case, the problem of multicollinearity is said to exist and it may cause some serious problems in validation, interpretation, and analysis of the model, such as unstable estimates. unreasonable sign, high-standard errors, and so on. The ridge estimator and the ridge predictor are defined by [40] and [73] to remedy multicollinearity in LMMs.

Another way to deal with multicollinearity is to consider parameter estimation with some additional information on the unknown parameters such as the exact or stochastic restrictions (see [48]). Many authors such as [17], [62] and [30] have studied the estimation of parameters in LMMs with additional restrictions. Also, [66] consider LMMs with the measurement error under stochastic restrictions and found the estimation of parameters.

It is well documented that incorporating a restriction on the parameter of interest and combining the estimator may improve the estimation, Making use of the ridge estimator when some linear restrictions are also present, [65] introduced a stochastic restricted ridge estimator in LMMs with the measurement error. [36] proposed an NMRE in the linear model by combining the ME and RE. [18] introduce an NMRE in a linear model with the measurement error, motivated by this, our primary aim in this paper is to obtain an NMRE in LMMs with the measurement error on fixed effects when stochastic linear restrictions are available on fixed and random effects to overcome the multicollinearity. We consider the case the variance parameters are not known and get the variance parameter estimations. The rest of the paper is organized as follows. We review the definition of the mixed estimator (ME) and mixed ridge estimator (MRE) in section 2.2. We propose the new mixed ridge estimator (NMRE) in section 2.3. We study the asymptotic properties of the new estimator in section 2.4. Some comparisons among the estimators are done to examine their superiority using the mean square error matrix (MSEM) criterion in section 2.5. A real data analysis and a simulation study are provided to illustrate the performance of the new estimator in section 2.6. Some conclusions are presented in section 2.7.

2.2 Mixed estimator and mixed ridge estimator

Consider the LMM with the measurement error under the stochastic restrictions as:

$$y_r = Z_r\beta + U_r b_r + \varepsilon_r,\ X_r = Z_r + L_r, \tag{2.2}$$

where $y_r = \begin{bmatrix} y \\ r \end{bmatrix}$, $Z_r = \begin{bmatrix} Z \\ R \end{bmatrix}$, $U_r = \begin{bmatrix} U & 0 \\ 0 & H \end{bmatrix}$, $b_r = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$, $\varepsilon_r = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \end{bmatrix}$ and $L_r = \begin{bmatrix} L \\ 0 \end{bmatrix}$, where r is an $n_2 \times 1$ observable random vector, R is an $n_2 \times p$ matrix of rank $n_2 < p$, $H = \oplus_{i=1}^{c_2} H_i$ is an $n_2 \times q_2$ matrix of regressors and H_i is an $n_2 \times q_{2i}$ known design matrix for the ith random effects factor, such that $q_2 = \sum\limits_{i=1}^{c_2} q_{2i}$. Also, $b_2 = \left(b_{21}', \ldots, b_{2c_2}'\right)'$ is an $q_2 \times 1$ vector of unobservable random effects from $N\left(0, \sigma^2\Sigma_2\right)$, where Σ_2 is a block diagonal matrix with the i^{th} block being $\gamma_{2i} I_{q_{2i}}$ for $\gamma_{2i} = \sigma_{2i}^2 / \sigma^2$, $i = 1, \ldots, c_2$. The error term ε_2 is an $n_2 \times 1$ vector of unobservable random errors from $N\left(0, \sigma^2 I_{n_2}\right)$ and assumed to be independent of b_{2i}, ε_1 and L. So, r has a multivariate normal distribution with $E(r) = R\beta$ and $Var(r) = \sigma^2 V_2$, in which $V_2 = I_{n_2} + H\Sigma_2 H'$, then b_r and y_r are jointly distributed as $\begin{bmatrix} b_r \\ y_r \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ Z_r\beta \end{bmatrix}, \begin{bmatrix} \sigma^2\Sigma_r & \sigma^2\Sigma_r U_r' \\ \sigma^2 U_r \Sigma_r & \sigma^2 V_r \end{bmatrix}\right)$ where $V_r = I_n + U_r\Sigma_r U_r' = \begin{bmatrix} V_1 & 0 \\ 0 & V_2 \end{bmatrix}$ and $\Sigma_r = \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix}$ in which $\gamma_r = (\gamma_{11}, \ldots, \gamma_{1c_1}, \gamma_{21}, \ldots, \gamma_{2c_2})'$. The mixed estimates (ME) of β, σ^2 and the mixed predictor (MP) of b_r for model 2.2, will be obtained as (see [66]):

$$\hat{\beta}_{me} = \left[X'V_1^{-1}X + R'V_2^{-1}R - tr\left(V_1^{-1}\right)\Lambda\right]^{-1}\left(X'V_1^{-1}y + R'V_2^{-1}r\right),$$

$$\hat{\sigma}^2_{me} = \frac{1}{n}\left(y'V_1^{-1}y - \hat{\beta}'_{me}X'V_1^{-1}y - \hat{\beta}'_{me}R'V_2^{-1}r + r'V_2^{-1}r\right),$$

$$\hat{b}_{mp} = \begin{bmatrix} \Sigma_1 U'V_1^{-1}\left(y - X\hat{\beta}_{me}\right) \\ \Sigma_2 H'V_2^{-1}\left(r - R\hat{\beta}_{me}\right) \end{bmatrix}.$$

Using lemma 1 (presented in Appendix), we can rewrite $\hat{\beta}_{me}$ as:

$$\hat{\beta}_{me} = \hat{\beta} + S^{-1}R'\left(V_2 + RS^{-1}R'\right)^{-1}\left(r - R\hat{\beta}\right), \tag{2.3}$$

where $\hat{\beta} = S^{-1}X'V_1^{-1}y$ and $S = \left[X'V_1^{-1}X - tr\left(V_1^{-1}\right)\Lambda\right]$.

If the elements of γ_r are unknown, the ME of σ^2_{1i}'s and σ^2_{2i}'s are given as:

$$\hat{\sigma}^2_{me_{1i}} = \frac{\left[\hat{b}'_{mp_{1i}}\hat{b}_{mp_{1i}} - tr\left(\hat{D}'_{me_i}\hat{D}_{me_i}\right)\hat{\beta}'_{me}\Lambda\hat{\beta}_{me}\right]}{\left[q_{1i} - tr\left(T_{ii}\right)\right]},$$

with $\hat{b}_{mp_{1i}} = \hat{D}_{me_i}\left(y - X\hat{\beta}_{me}\right)$ and $\hat{D}_{me_i} = \hat{\gamma}_{me_{1i}}U'_i\hat{V}_1^{-1}$, $i = 1, ..., c_1$,

$$\hat{\sigma}^2_{me_{2i}} = \frac{\hat{b}'_{mp_{2i}}\hat{b}_{mp_{2i}}}{\left[q_{2i} - tr\left(F_{ii}\right)\right]},$$

with $\hat{b}_{mrp_{2i}} = \hat{\gamma}_{mre_{2i}}H'_i\hat{V}_2^{-1}\left(r - R\hat{\beta}_{mre}\right)$ and $F = \left(I_{q_2} + H'H\Sigma_2\right)^{-1}$, $i = 1, ..., c_2$, (see [66]). An iterative algorithm is used to obtain the ME of parameters.

Now, consider the augmented model

$$y_{rk} = Z_{rk}\beta + U_{rk}b_r + \varepsilon_{rk},\; X_{rk} = Z_{rk} + L_{rk}, \tag{2.4}$$

and define $y_{rk} = \begin{bmatrix} y \\ r \\ 0 \end{bmatrix}$, $Z_{rk} = \begin{bmatrix} Z \\ R \\ \sqrt{k}I_p \end{bmatrix}$, $U_{rk} = \begin{bmatrix} U & 0 \\ 0 & H \\ 0 & 0 \end{bmatrix}$, $\varepsilon_{rk} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \end{bmatrix}$ and $L_{rk} = \begin{bmatrix} L \\ 0 \\ 0 \end{bmatrix}$. The error term ε_{rk} is an error vector with $E(\varepsilon_{rk}) = 0$ and $var\left(\varepsilon_{rk}\right) = \sigma^2 I_N$, where $N = n_1 + n_2 + p$, and $I_N = \begin{bmatrix} I_{n_1} & 0 & 0 \\ 0 & I_{n_2} & 0 \\ 0 & 0 & I_p \end{bmatrix}$. So, y_{rk} has a multivariate normal distribution with $E\left(y_{rk}\right) = Z_{rk}\beta$ and $Var\left(y_{rk}\right) = \sigma^2 V_{rk}$, in which $V_{rk} = I_N + U_{rk}\Sigma_r U'_{rk}$. Also, the conditional distribution of b_r given y_{rk} is $N\left(\Sigma_r U'_{rk}V_{rk}^{-1}\left(y_{rk} - Z_{rk}\beta\right), \sigma^2\Sigma_r T_{rk}\right)$, where $T_{rk} = \left(I_q + U'_{rk}U_{rk}\Sigma_r\right)^{-1}$. The mixed ridge estimates (MRE) of β, σ^2 and the mixed ridge predictor (MRP) of b_r for the model 2.2, will be obtained as (see [66]):

$$\hat{\beta}_{mre} = \left[X'V_1^{-1}X + R'V_2^{-1}R + kI_p - tr\left(V_1^{-1}\right)\Lambda\right]^{-1}\left(X'V_1^{-1}y + R'V_2^{-1}r\right),$$

$$\hat{\sigma}^2_{mre} = \frac{1}{N}\left(y'V_1^{-1}y - \hat{\beta}'_{mre}X'V_1^{-1}y - \hat{\beta}'_{mre}R'V_2^{-1}r + r'V_2^{-1}r\right),$$

$$\hat{b}_{mrp} = \begin{bmatrix} \Sigma_1 U' V_1^{-1} \left(y - X\hat{\beta}_{mre} \right) \\ \Sigma_2 H' V_2^{-1} \left(r - R\hat{\beta}_{mre} \right) \end{bmatrix}.$$

Using lemma 1 (presented in Appendix), we can rewrite $\hat{\beta}_{mre}$ as:

$$\hat{\beta}_{mre} = \hat{\beta}_{re} + S_k^{-1} R' \left(V_2 + R S_k^{-1} R' \right)^{-1} \left(r - R\hat{\beta}_{re} \right),$$

where $\hat{\beta}_{re} = S_k^{-1} X' V_1^{-1} y$ and $S_k = \left[X' V_1^{-1} X - tr\left(V_1^{-1} \right) \Lambda + kI_p \right]$.

If the elements of γ_r are unknown, the MRE of σ_{1i}^2's and σ_{2i}^2's are given as:

$$\hat{\sigma}^2_{mre_{1i}} = \frac{\left[\hat{b}'_{mrp_{1i}} \hat{b}_{mrp_{1i}} - tr\left(\hat{D}'_{mre_i} \hat{D}_{mre_i} \right) \hat{\beta}'_{mre} \Lambda \hat{\beta}_{mre} \right]}{\left[q_{1i} - tr\left(T_{ii} \right) \right]},$$

with $\hat{b}_{mrp_{1i}} = \hat{D}_{mre_i} \left(y - X\hat{\beta}_{mre} \right)$ and $\hat{D}_{mre_i} = \hat{\gamma}_{mre_{1i}} U_i' \hat{V}_1^{-1}$, $i = 1, ..., c_1$,

$$\hat{\sigma}^2_{mre_{2i}} = \frac{\hat{b}'_{mre_{2i}} \hat{b}_{mre_{2i}}}{\left[q_{2i} - tr\left(F_{ii} \right) \right]},$$

with $\hat{b}_{mrp_{2i}} = \hat{\gamma}_{mre_{2i}} H_i' \hat{V}_2^{-1} \left(r - R\hat{\beta}_{mre} \right)$, $i = 1, ..., c_2$, (see [65]). An iterative algorithm is used to obtain the MRE of parameters.

2.3 The new estimator

Based on [36], a new ridge type estimator is obtained by combining the ME and RE. Substituting $\hat{\beta}$ with $\hat{\beta}_{re}$ in 2.3, we can get the new mixed ridge estimator (NMRE) as:

$$\hat{\beta}_{nmre} = \hat{\beta}_{re} + S^{-1} R' \left(V_2 + R S^{-1} R' \right)^{-1} \left(r - R\hat{\beta}_{re} \right),$$

Let $T_k = \left(I_p + kS^{-1} \right)^{-1} S_k^{-1} S = S S_k^{-1}$. So, we can write the NMRE of β as:

$$\begin{aligned}
\hat{\beta}_{nmre} &= \hat{\beta}_{re} + S^{-1} R' \left[V_2 + R S^{-1} R' \right]^{-1} \left(r - R\hat{\beta}_{re} + R S^{-1} R' V_2^{-1} r - R S^{-1} R' V_2^{-1} r \right) \\
&= \hat{\beta}_{re} + S^{-1} R' \left[V_2 + R S^{-1} R' \right]^{-1} \left[\left(I_p + R S^{-1} R' V_2^{-1} \right) r - R\hat{\beta}_{re} - R S^{-1} R' V_2^{-1} r \right] \\
&= S^{-1} T_k X' V_1^{-1} y + S^{-1} R' V_2^{-1} r - S^{-1} R' \left[V_2 + R S^{-1} R' \right]^{-1} R S^{-1} T_k X' V_1^{-1} y \\
&\quad - S^{-1} R' \left[V_2 + R S^{-1} R' \right]^{-1} R S^{-1} R' V_2^{-1} r \\
&= \left[S^{-1} - S^{-1} R' \left(V_2 + R S^{-1} R' \right)^{-1} R S^{-1} \right] \left(T_k X' V_1^{-1} y + R' V_2^{-1} r \right) \\
&= \left(S + R' V_2^{-1} R \right)^{-1} \left(T_k X' V_1^{-1} y + R' V_2^{-1} r \right),
\end{aligned}$$

We obtain the NMRE of σ^2 and the new mixed ridge predictor (NMRP) of b_r as:

$$\hat{\sigma}^2_{nmre} = \frac{1}{N} \left(y' V_1^{-1} y - \hat{\beta}'_{nmre} X' V_1^{-1} y - \hat{\beta}'_{nmre} R' V_2^{-1} r + r' V_2^{-1} r \right),$$

$$\hat{b}_{nmrp} = \begin{bmatrix} \Sigma_1 U' V_1^{-1} \left(y - X\hat{\beta}_{nmre} \right) \\ \Sigma_2 H' V_2^{-1} \left(r - R\hat{\beta}_{nmre} \right) \end{bmatrix}.$$

If the elements of γ_r are unknown, their NMRE will be substituted back into Σ_r to obtain $\hat{\beta}_{nmre}$, $\hat{\sigma}^2_{nmre}$ and $\hat{b}_{nmrp}$. For the NMRE of $\gamma_{ji}{}'s$, we obtain the NMRE of σ^2_{1i}'s and σ^2_{2i}'s as:

$$\hat{\sigma}^2_{nmre_{1i}} = \frac{\left[\hat{b}'_{nmrp_{1i}} \hat{b}_{nmrp_{1i}} - tr\left(\hat{D}'_{nmre_i} \hat{D}_{nmre_i}\right) \hat{\beta}'_{nmre} \Lambda \hat{\beta}_{nmre}\right]}{\left[q_{1i} - tr\left(T_{ii}\right)\right]},$$

with $\hat{b}_{nmrp_{2i}} = \hat{\gamma}_{nmre_{2i}} H'_i \hat{V}_2^{-1} \left(r - R\hat{\beta}_{nmre} \right)$, $\hat{D}_{nmre_i} = \hat{\gamma}_{nmre_{1i}} U'_i \hat{V}_1^{-1},\ i = 1, ..., c_1,$

$$\hat{\sigma}^2_{nmre_{2i}} = \frac{\hat{b}'_{nmrp_{2i}} \hat{b}_{nmrp_{2i}}}{\left[q_{2i} - tr\left(F_{ii}\right)\right]},$$

with $\hat{b}_{nmrp_{2i}} = \hat{\gamma}_{nmre_{2i}} H'_i \hat{V}_2^{-1} \left(r - R\hat{\beta}_{nmre} \right)$, $F = \left(I_{q_2} + H'H\Sigma_2 \right)^{-1}$, $i = 1, ..., c_2$. We must use an iterative algorithm to derive the NMRE of parameters.

2.4 Asymptotic properties of the new estimator

In this section, we study the asymptotic behaviour of the fixed effects estimators, we assume the parameter β is identifiable and assume that as n tends to infinity, the limits of $n^{-1} tr\left(V_1^{-1}\right)$, $n^{-1}\left(Z'V_1^{-1}Z + R'V_2^{-1}R + kI_p\right)$, $n^{-1}\left(Z'V_1^{-2}Z\right)$, $n^{-1}\left(Z'V_1^{-1}Z + R'V_2^{-1}R\right)$, $n^{-1}\left(Z'V_1^{-1}Z + kI_p\right)$, $n^{-1}\left(Z'V_1^{-1}Z\right)$, and $n^{-1}\left(G_k X'V_1^{-1}y + R'V_2^{-1}r\right)$ exist, where $G_k = \left[I_p + k\left(Z'V_1^{-1}Z\right)^{-1}\right]^{-1}$, (see [69]).

Theorem 2.4.1. *The asymptotic distribution of $\hat{\beta}_{nmre}$ is normal with mean vector $M^{-1}M_k\beta$ and covariance matrix*

$$AVar\left(\hat{\beta}_{nmre}\right) = M^{-1}\left[G_k B G_k + \sigma^2\left(G_k Z'V_1^{-1}ZG_k + R'V_2^{-1}R\right)\right] M^{-1},$$

where $M_k = G_k Z'V_1^{-1}Z + R'V_2^{-1}R$, $M = M_{k=0} = Z'V_1^{-1}Z + R'V_2^{-1}R$, $G_k = \left[I_p + k\left(Z'V_1^{-1}Z\right)^{-1}\right]^{-1}$, and $B = \left[\beta'Z'V_1^{-2}Z\beta + \sigma^2 tr\left(V_1^{-1}\right)\right]\Lambda$.

Proof. Since $E\left(X'V_1^{-1}X\right) = Z'V_1^{-1}Z + tr\left(V_1^{-1}\right)\Lambda$ (see [16]), we have

$$X'V_1^{-1}X = Z'V_1^{-1}Z + tr\left(V_1^{-1}\right)\Lambda + O_p\left(n^{1/2}\right). \tag{2.5}$$

Moreover, we can obtain $T_k = G_k + O_p\left(n^{-1/2}\right)$, so, It follows from 2.3 and 2.5 that

$$\begin{aligned} \sqrt{n}\hat{\beta}_{nmre} &= \left[n^{-1}\left(Z'V_1^{-1}Z + R'V_2^{-1}R\right)\right]^{-1}\left[I_p + O_p\left(n^{-1/2}\right)\right]^{-1} n^{-1/2}\left(T_k X'V_1^{-1}y + R'V_2^{-1}r\right) \\ &= \left[n^{-1}\left(Z'V_1^{-1}Z + R'V_2^{-1}R\right)\right]^{-1}\left[I_p + O_p\left(n^{-1/2}\right)\right] \qquad (2.6) \\ &\quad \times n^{-1/2}\left[G_k X'V_1^{-1}y + O_p\left(n^{-1/2}\right) X'V_1^{-1}y + R'V_2^{-1}r\right], \end{aligned}$$

where $\left[I_p+O_p\left(n^{-1/2}\right)\right]^{-1}=\left[I_p+O_p\left(n^{-1/2}\right)\right]$ is obtained from Taylor series expansion. Moreover, since the limit of $C=n^{-1}\left(Z'V_1^{-1}Z+R'V_2^{-1}R\right)$ exists, then 2.6 can be written as $\sqrt{n}\hat{\beta}_{nmre}=C^{-1}h+O_p\left(n^{-1/2}\right)$, where $h=n^{-1/2}\left(G_kX'V_1^{-1}y+R'V_2^{-1}r\right)$ is asymptotically normal (see [16]). So, it follows from $E\left(G_kX'V_1^{-1}y+R'V_2^{-1}r\right)=\left(G_kZ'V_1^{-1}Z+R'V_2^{-1}R\right)\beta$, that $E(h)=n^{-1/2}M_k\beta$. Consequently, we have $\sqrt{n}\left(\hat{\beta}_{nmre}-M^{-1}M_k\beta\right)=C^{-1}\left[h-E(h)\right]+O_p\left(n^{-1/2}\right)$, which indicates that $\sqrt{n}\left(\hat{\beta}_{nmre}-M^{-1}M_k\beta\right)$ is asymptotically normal with zero mean. So, we can asymptotically write

$$\begin{aligned}E\left(\hat{\beta}_{nmre}\right)&=M^{-1}G_kZ'V_1^{-1}Z\beta+M^{-1}R'V_2^{-1}R\beta\\&=\beta+M^{-1}\left(G_k-I_p\right)Z'V_1^{-1}Z\beta.\end{aligned}$$

Furthermore, we have $AVar\left(\sqrt{n}\hat{\beta}_{nmre}\right)=C^{-1}Var(h)C^{-1}$. The variance of h can be obtained as:

$$\begin{aligned}Var(h)&=E_{y_r}\left[Var(h|y_r)\right]+Var_{y_r}\left[E(h|y_r)\right]\\&=n^{-1}E_{y_r}\left(y'V_1^{-2}y\right)\Lambda+n^{-1}Var_{y_r}\left(G_kZV_1^{-1}y+R'V_2^{-1}r\right),\end{aligned}$$

where E_{y_r} and Var_{y_r} denote the expectation and variance with respect to the random vector y_r. Since $Var\left(G_kZ'V_1^{-1}y+R'V_2^{-1}r\right)=\sigma^2\left(G_kZ'V_1^{-1}ZG_k+R'V_2^{-1}R\right)$, and $E\left(y'V_1^{-2}y\right)=\beta'Z'V_1^{-2}Z\beta+\sigma^2tr\left(V_1^{-1}\right)$, therefore $Var(h)=n^{-1}\left[G_kBG_k+\sigma^2\left(G_kZ'V_1^{-1}ZG_k+R'V_2^{-1}R\right)\right]$, whose limit exists as n tends to infinity. Thus

$$AVar\left(\hat{\beta}_{nmre}\right)=M^{-1}\left[G_kBG_k+\sigma^2\left(G_kZ'V_1^{-1}ZG_k+R'V_2^{-1}R\right)\right]M^{-1},$$

this complete the proof. □

Corollary 2.4.1. *$\hat{\beta}_{me}$ has an asymptotic normal distribution with mean vector β and covariance matrix $AVar\left(\hat{\beta}_{me}\right)=M^{-1}\left(B+\sigma^2M\right)M^{-1}$, where $M=\left(Z'V_1^{-1}Z+R'V_2^{-1}R\right)$.*

Corollary 2.4.2. *$\hat{\beta}_{mre}$ has an asymptotic normal distribution with mean vector $A_k^{-1}M\beta$ and covariance matrix $AVar\left(\hat{\beta}_{mre}\right)=A_k^{-1}\left(B+\sigma^2A_k\right)A_k^{-1}$, where $A_k=\left(Z'V_1^{-1}Z+R'V_2^{-1}R+kI_p\right)$.*

2.5 Mean square error comparison of the estimators

In this section, we consider the comparison of the NMRE to the ME and MRE. One of the criteria proposed for measuring the goodness of an estimator is taken to be the MSE criterion. The MSE matrix concept to a $p\times 1$ vector of estimators, say $\hat{\beta}$, is the $p\times p$ matrix defined as $MSEM\left(\hat{\beta}\right)=E\left(\hat{\beta}-\beta\right)\left(\hat{\beta}-\beta\right)'=Var\left(\hat{\beta}\right)+Bias\left(\hat{\beta}\right)Bias\left(\hat{\beta}\right)'$, where $Var(\hat{\beta})$ is the variance-covariance matrix for $\hat{\beta}$ and $Bias\left(\hat{\beta}\right)$ is the $p\times 1$ vector of the bias term for each $\hat{\beta}_i$. We obtain the asymptotic mean square error matrices of $\hat{\beta}_{nmre}$, $\hat{\beta}_{mre}$ and $\hat{\beta}_{me}$ as:

$$AMSEM\left(\hat{\beta}_{nmre}\right)=M^{-1}\left[G_kBG_k+\sigma^2\left(G_kZ'V_1^{-1}ZG_k+R'V_2^{-1}R\right)\right]M^{-1}+d_1'd_1,$$

$$AMSEM\left(\hat{\beta}_{mre}\right) = A_k^{-1}\left(B+\sigma^2 A_k\right)A_k^{-1} + d_2' d_2,$$

$$AMSEM\left(\hat{\beta}_{me}\right) = M^{-1}\left(B+\sigma^2 M\right)M^{-1},$$

respectively, where $d_1 = M^{-1}\left(G_k - I_p\right)Z'V_1^{-1}Z\beta$ and $d_2 = -kA_k^{-1}\beta$. In order to compare $\hat{\beta}_{nmre}$ with $\hat{\beta}_{me}$, in the MSEM sense, we consider the following difference

$$\Delta_1 = AMSEM\left(\hat{\beta}_{me}\right) - AMSEM\left(\hat{\beta}_{nmre}\right) = D_1 - d_1 d_1', \tag{2.7}$$

in which $D_1 = M^{-1}\left[B - G_k B G_k + \sigma^2\left(Z'V_1^{-1}Z - G_k Z'V_1^{-1}ZG_k\right)\right]M^{-1}$. We express the following theorem for the superiority of $\hat{\beta}_{nmre}$ over $\hat{\beta}_{me}$.

Theorem 2.5.1. *$\hat{\beta}_{nmre}$ is superior to the estimator $\hat{\beta}_{me}$ in the MSEM criterion, if $d_1' D_1^{-1} d_1 \leq 1$.*

Proof. The MSEM difference between NMRE and ME given in 2.7 is $\Delta_1 = D_1 - d_1 d_1'$. To show that $\Delta_1 > 0$ apply Lemma 2 (presented in Appendix), we need to prove that D_1 is a p.d. matrix. For $A = Z'V_1^{-1}Z > 0$ there exists some orthogonal matrix Q , such that $A = Q\Lambda Q'$, where $\Lambda = diag\left(\lambda_1,...,\lambda_p\right)$, $\lambda_i > 0$. Therefore, we can compute that $Z'V_1^{-1}Z - G_k Z'V_1^{-1}ZG_k = Q\Gamma Q' = Q\left(\gamma_1,...,\gamma_p\right)Q'$, where $\Gamma = \Lambda - \left(I_p + k\Lambda^{-1}\right)^{-1}\Lambda\left(I_p + k\Lambda^{-1}\right)^{-1}$. For $k > 0$, $\lambda_i > 0$, we have $\gamma_i > 0$, which means $Z'V_1^{-1}Z - G_k Z'V_1^{-1}ZG_k > 0$. Furthermore, $B - G_k B G_k > 0$ and $M^{-1} = \left(Z'V_1^{-1}Z + R'V_2^{-1}R\right)^{-1} > 0$, this implies that $D_1 > 0$. Hence according to Lemma 2, we can get $\Delta_1 > 0$ if $d_1' D_1^{-1} d_1 \leq 1$. □

We consider matrix difference Δ_1 and note that $B - G_k B G_k$ is a p.d. matrix, we have that Δ_1 is a p.d. matrix if $\sigma^2\left(Z'V_1^{-1}Z - G_k Z'V_1^{-1}ZG_k\right) - \left(G_k - I_p\right)Z'V_1^{-1}Z\beta\beta' Z'V_1^{-1}Z\left(G_k - I_p\right)M$ is a p.d. matrix. Denoting $\alpha = Q'\beta$, we have $\beta = Q\alpha$. In this case, we can write

$$\sigma^2\left(Z'V_1^{-1}Z - G_k Z'V_1^{-1}ZG_k\right) - \left(G_k - I_p\right)Z'Z\beta\beta' Z'Z\left(G_k - I_p\right) = QTQ',$$

where $T = diag\left(\tau_1,...,\tau_p\right)$ and $\tau_i = \frac{\sigma^2 k\lambda_i(k+2\lambda_i) - \alpha_i^2\lambda_i^2 k^2}{(\lambda_i + k)^2}$. For $k > 0$ and $\lambda_i > 0$, we have $\tau_i > 0$ if and only if $k < \frac{2\sigma^2}{\alpha_i^2 - \lambda_i^{-1}\sigma^2}$. We suggest the replacement of σ^2, α_i and λ_i by their corresponding appropriate estimators. Therefore, we have $\hat{k}_{nmre} = \frac{2\hat{\sigma}^2}{\hat{\alpha}_{\max}^2 - \lambda_{\max}^{-1}\hat{\sigma}^2}$, where $\hat{\alpha}_{\max}^2$ and $\lambda_{\max}$ are denote the maximum element of $\alpha = Q\hat{\beta}$ and Λ . Moreover, based on [65], a sufficient condition for $\hat{\beta}_{mre}$ to be superior over $\hat{\beta}_{me}$ is $\hat{k}_{mre} \leq 2\hat{\sigma}^2 \big/ \hat{\beta}'\hat{\beta}$. Note that $\hat{k}_{mre} < \hat{k}_{nmre}$, therefore, in this paper, we suggest to choose $\hat{k} = \hat{k}_{mre}$.

In order to compare $\hat{\beta}_{nmre}$ with $\hat{\beta}_{mre}$, in the MSEM sense, we consider the following difference

$$\Delta_2 = AMSEM\left(\hat{\beta}_{mre}\right) - AMSEM\left(\hat{\beta}_{nmre}\right) = D_2 + d_2 d_2' - d_1 d_1', \tag{2.8}$$

where $D_2 = A_k^{-1}\left(B+\sigma^2 M\right)A_k^{-1} - M^{-1}\left[G_k B G_k + \sigma^2\left(G_k Z'V_1^{-1}ZG_k + R'V_2^{-1}R\right)\right]M^{-1}$. We express the following theorem for the superiority of $\hat{\beta}_{nmre}$ over $\hat{\beta}_{mre}$.

Theorem 2.5.2. *when the maximum eigenvalue of the matrix* $M^{-1}\left[G_kBG_k+\sigma^2\left(G_kZ'V_1^{-1}ZG_k+R'V_2^{-1}R\right)\right]M^{-1}\left[A_k^{-1}\left(B+\sigma^2M\right)A_k^{-1}\right]^{-1}$ *is less than 1,* $\hat{\beta}_{nmre}$ *will be superior to* $\hat{\beta}_{mre}$*, in the MSEM sense if* $d_1'(D_2+d_2d_2')^{-1}d_1 \leq 1$.

Proof. The MSEM difference between NMRE and MRE given in 2.8 is $\Delta_2 = D_2 + d_2d_2' - d_1d_1'$. To show that $\Delta_2 > 0$, Lemma 3 (presented in Appendix) can be used. A requirement to apply Lemma 3, is that D_2 to be a p.d. matrix. It is clear that $A_k^{-1}\left(B+\sigma^2A_k\right)A_k^{-1} > 0$ and $M^{-1}\left[G_kBG_k+\sigma^2\left(G_kZ'V_1^{-1}ZG_k+R'V_2^{-1}R\right)\right]M^{-1} \geq 0$. According to Lemma 4 (presented in Appendix), $A_k^{-1}\left(B+\sigma^2A_k\right)A_k^{-1} > M^{-1}\left[G_kBG_k+\sigma^2\left(G_kZ'V_1^{-1}ZG_k+R'V_2^{-1}R\right)\right]M^{-1}$, if and only if $\lambda < 1$, where λ is the maximum eigenvalue of

$$M^{-1}\left[G_kBG_k+\sigma^2\left(G_kZ'V_1^{-1}ZG_k+R'V_2^{-1}R\right)\right]M^{-1}\left[A_k^{-1}\left(B+\sigma^2A_k\right)A_k^{-1}\right]^{-1}.$$

Therefore, D_2 is a p.d. matrix, if and only if $\lambda < 1$. Then according to lemma 3, Δ_2 is a p.d. matrix if $d_1'(D_2+d_2d_2')^{-1}d_1 \leq 1$, this complete the proof. □

2.6 Mean square error comparison of the predictors

Prediction of linear combination of β and b_1 can be expressed as $\mu = L\beta + Sb_1$, for specific matrix $L \in {}^{p\times s}$ and $S \in {}^{q\times s}$. This type of prediction problem was investigated by [46] and [62] for the situation $s = 1$. We assume $b_2 = 0$ and obtain the predictor of μ under the NMRP and NMP. Following [62], The mean square error matrix (MSEM) of any estimator or predictor is defined as:

$$\begin{aligned} MSEM(\hat{\mu}) &= E\left[(\hat{\mu}-\mu)(\hat{\mu}-\mu)'\right] = Var(\hat{\mu}) + Var(\mu) \\ &+ bias(\hat{\mu})\, bias(\hat{\mu})' - \operatorname{cov}(\hat{\mu},\mu) - \operatorname{cov}(\mu,\hat{\mu}), \end{aligned} \tag{2.9}$$

where $bias\left(\hat{\beta}\right) = E\left(\hat{\beta}\right) - \beta$. To get asymptotic $MSEM(\hat{\mu}_{wmrp})$ from equation (2.9), we obtain $Var(\hat{\mu}_{nmrp})$, $Var(\mu)$, $bias(\hat{\mu}_{nmrp})$ and $\operatorname{cov}(\hat{\mu}_{nmrp},\mu)$ as:

$$\begin{aligned} Var(\hat{\mu}_{nmrp}) &= QVar\left(\hat{\beta}_{nmre}\right)Q' + \sigma^2S'\Sigma_1U'V_1^{-1}U\Sigma_1S + Q\operatorname{cov}\left(\hat{\beta}_{nmre},y\right)V_1^{-1}U\Sigma_1S \qquad (2.10)\\ &+ S'\Sigma_1U'V_1^{-1}\operatorname{cov}\left(y,\hat{\beta}_{nmre}\right)Q' \\ &= QVar\left(\hat{\beta}_{nmre}\right)Q' \\ &+ \sigma^2\left[S'\Sigma_1U'V_1^{-1}U\Sigma_1S + QA_rT_kX'V_1^{-1}U\Sigma_1S + S'\Sigma_1U'V_1^{-1}XT_k'A'_rQ'\right] \end{aligned}$$

where, $\operatorname{cov}\left(\hat{\beta}_{nmre},y\right) = \operatorname{cov}\left(A_r\left(T_kX'V_1^{-1}y+R'V_2^{-1}r\right),y\right) = \operatorname{cov}\left(A_r\left(T_kX'V_1^{-1}y\right),y\right) = \sigma^2A_rT_kX'$, and $A_r = \left[X'V_1^{-1}X+R'V_2^{-1}R-tr\left(V_1^{-1}\right)\Lambda\right]^{-1}$.

$$Var(\mu) = Var\left(L'\beta+S'b\right) = \sigma^2S'\Sigma_1S, \tag{2.11}$$

$$\begin{aligned} bias(\hat{\mu}_{nmrp}) &= E(\hat{\mu}_{nmrp}-\mu) = E\left(Q\hat{\beta}_{nmre}+S'\Sigma_1U'V_1^{-1}y-L'\beta-S'b\right) \\ &= QE\left(\hat{\beta}_{nmre}-\beta\right) = Qbias\left(\hat{\beta}_{nmre}\right), \end{aligned} \tag{2.12}$$

$$\begin{aligned}\operatorname{cov}\left(\hat{\mu}_{nmre},\mu\right)&=\operatorname{cov}\left(Q\hat{\beta}_{nmre}+S'\Sigma_1U'V_1^{-1}y,L'\beta+S'b\right)=\\&=\sigma^2QA_rT_kX'V_1^{-1}U\Sigma_1S+\sigma^2S'\Sigma_1U'V_1^{-1}U\Sigma_1S,\end{aligned}\tag{2.13}$$

in which, $\operatorname{cov}\left(\hat{\beta}_{nmre},b\right)=\operatorname{cov}\left(A_r\left(T_kX'V_1^{-1}y+R'V_2^{-1}r\right),b\right)=\sigma^2A_rT_kX'V_1^{-1}U\Sigma_1$. Then equations (2.10), (2.11), (2.12) and (2.13) are put into equation (2.9) to get

$$AMSEM\left(\hat{\mu}_{nmrp}\right)=QAMSEM\left(\hat{\beta}_{nmre}\right)Q'+\sigma^2S'\left[\Sigma_1-\Sigma_1U'V_1^{-1}U\Sigma_1\right]S.\tag{2.14}$$

Similarly, we can obtain the predictor of μ under the WMP as:

$$AMSEM\left(\hat{\mu}_{mrp}\right)=QAMSEM\left(\hat{\beta}_{mre}\right)Q'+\sigma^2S'\left[\Sigma_1-\Sigma_1U'V_1^{-1}U\Sigma_1\right]S.\tag{2.15}$$

We are now interested in knowing under which conditions $\hat{\mu}_{nmrp}$ is better than $\hat{\mu}_{mp}$. For this, we investigate the difference

$$\Delta=AMSEM\left(\hat{\mu}_{mrp}\right)-AMSEM\left(\hat{\mu}_{nmrp}\right),\tag{2.16}$$

and from equations (2.14) and (2.15), Δ in equation (2.16) can equivalently be written as $\Delta=Q\Delta_1Q'$, where $\Delta_1=AMSEM\left(\hat{\beta}_{mre}\right)-AMSEM\left(\hat{\beta}_{nmre}\right)$. Similarly, we derive that superiority of $\hat{\mu}_{nmrp}$ over $\hat{\mu}_{mp}$ is equivalent to the superiority of $\hat{\beta}_{nmre}$ over $\hat{\beta}_{me}$. This means that if we want to compare two different predictors, we just need to compare the corresponding estimators under the MSEM criterion, which is discussed in previous section.

2.7 Data example and simulation study

To evaluate the performance of the new estimator, we consider a sample of real data, Boston housing data taken from [24]. [69] used this data set and considered the data of 132 census tracts in the 15 districts of Boston (as a part of 506 observations on census tracts in the Boston Standard Metropolitan Statistical Area [SMSA] in 1970). The census tracts within each district are taken as repeated measurements. The pollution variable NOX is taken to have measurement errors. Therefore, they employed a LMM with the measurement error. We use the same data in this example (as presented in [16]).

For this data set, the condition number of $\hat{Z}'\hat{V}_1^{-1}\hat{Z}$ is equal to 159, which indicates that there is collinearity among the fixed effects variables. We can improve the CSE by adding stochastic linear restrictions $r=R\beta+Hb_2+\varepsilon_2$, $\varepsilon_2\sim N\left(0,I_7\right)$ to model 2.1, where the rows are the 133th-139th observations of the historical data (see [66]).

Then, we can write our model as (2.4) and find the mixed ridge estimated MSE of β. A MRE of Z is derived as $\hat{Z}_{mre}=X+\hat{v}_{mre}\hat{\beta}'_{mre}\Lambda\hat{\sigma}^{-2}_{\hat{v}_{mre}}$, where $\hat{v}_{mre}=y-X\hat{\beta}_{mre}-U\hat{b}_{1_{mre}}$ and $\hat{\sigma}^2_{\hat{v}_{mre}}=\hat{\sigma}^2_{mre}+\hat{\beta}'_{mre}\Lambda\hat{\beta}_{mre}$.

The estimated MSE values of ME, MRE and NMRE and the values of $d_1'\left(D_2+d_2d_2'\right)^{-1}d_1$ and $d_1'D_1^{-1}d_1$ are presented in Table 3.1. We can find that estimated MSE values of the NMRE and MRE are indeed smaller than of the ME, and the NMRE has smaller estimated MSE values than the MRE. Moreover, $d_1'D_1^{-1}d_1\leq 1$ and $d_1'\left(D_2+d_2d_2'\right)^{-1}d_1\leq 1$, which agree with

ME	MRE	$NMRE$	$d_1'D_1^{-1}d_1$	$d_1'(D_2+d_2d_2')d_1 \leq 1$
0.1547	0.1525	0.1482	0.0.0519	0.255

Table 2.1: The Estimated MSE values of ME, MRE, NMRE

our theoretical findings in Theorems 2.5.1 and 2.5.2. This can imply that the new estimator $\hat{\beta}_{nmre}$ can perform better than the $\hat{\beta}_{mre}$ and $\hat{\beta}_{me}$.

For more convenience, a plot of the estimated MSE values of the estimators versus k in the interval [0,0.005] is depicted in Figure 1. Since the $\hat{\beta}_{me}$ does not depend on k, its estimated MSE value is the same for all k values. As seen in Figure 1, $\hat{\beta}_{nmre}$ dominates $\hat{\beta}_{mre}$ for k greater than about 0.0005 and it looks better than $\hat{\beta}_{me}$ for $0.0006 < k < 0.0034$. Moreover, $\hat{\beta}_{mre}$ looks better than $\hat{\beta}_{me}$ for $0.0008 < k < 0.0024$. In order to further illustrate the behaviour of the new

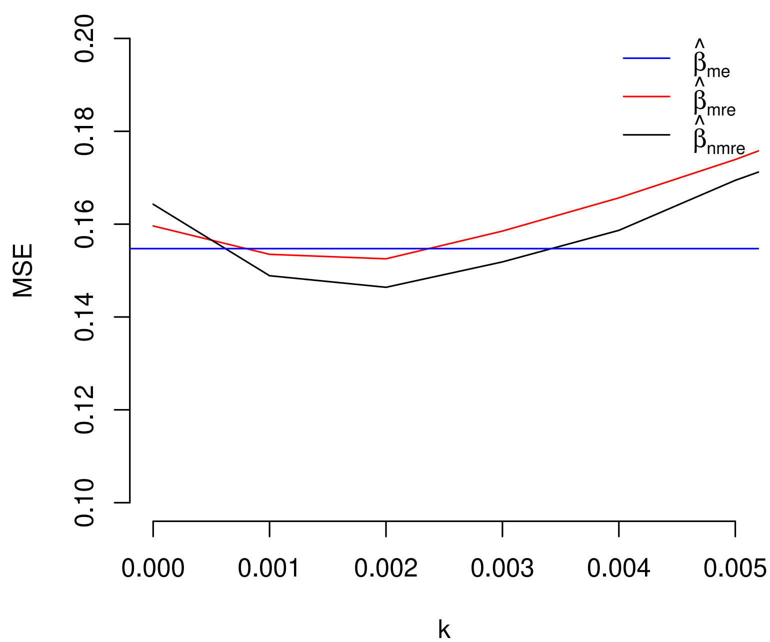

Figure 2.1: Estimated MSE values of the ME, MRE and NMRE

estimator, we will perform a Monte Carlo simulation study by considering the different degrees of multicollinearity. To achieve different degree of multicollinearity, following [31], the fixed effect variables are generated as:

$$z_{kt} = \left(1-\rho^2\right)^{\frac{1}{2}} u_{kt} + \rho u_{kp+1},\ k = 1, \ldots, n_1,\ t = 1, \ldots, p,$$

where u_{kt}'s are independent standard normal pseudo-random numbers and ρ is specified so that the correlation between any two fixed effect variables is given by ρ^2. Four different sets of correlation were considered, corresponding to $\rho = 0.85, 0.9, 0.95$ and 0.99 . In this study, q_1 can be interpreted as the number of independent clusters and m_1 is the cluster size in a longitudinal study and $n_1 = m_1 \times q_1$ is the total size. We generate the sth set of simulated data as:

$$y_s = Z\beta + Ub_{1s} + \varepsilon_{1s}, X_s = Z + L_s,$$
$$r_s = R\beta + Hb_{2s} + \varepsilon_{2s},$$

where $y_s = \left(y_{11s}, ..., y_{1m_1s}, y_{21s}, ..., y_{2m_1s}, ..., y_{q_11s}, ..., y_{q_1m_1s}\right)'$ and $b_{1s} = \left(b_{11s}, ..., b_{1q_1s}\right)'$. We have $Z = \left(Z^{(1)}, Z^{(2)}, Z^{(3)}\right)$, where $Z^{(t)} = \left(Z^{(t)}_{11}, ..., Z^{(t)}_{1m_1}, Z^{(t)}_{21}, ..., Z^{(t)}_{2m_1}, ..., Z^{(t)}_{q_11}, ..., Z^{(t)}_{q_1m_1}\right)$; $t = 1, 2, 3.$, we assume that $c_1 = 1$. The vector β is chosen as the eigenvector corresponding to the largest eigenvalues of the matrix $Z'V_1^{-1}Z$. We consider $U = I_{q_1} \otimes 1_{m_1}$, in which 1_{m_1} is a $m_1 \times 1$ vector where all elements are 1 and $\varepsilon_{1s} \sim N(0, \sigma^2)$ is rewritten in accordance with y_s. Moreover, $R^{(t)} = \left(R^{(t)}_{11}, ..., R^{(t)}_{1m_2}, R^{(t)}_{21}, ..., R^{(t)}_{2m_2}, ..., R^{(t)}_{q_11}, ..., R^{(t)}_{q_1m_2}\right)$; $t = 1, 2, 3, R_{lt} \sim N(0, 1)$, $l = 1, ..., n_2$, also $R = \left(R^{(1)}, R^{(2)}, R^{(3)}\right)$, and $r_s = (r_{1s}, r_{2s})'$. We consider $L = I_{q_2} \otimes 1_{m_2}$, in which 1_{m_2} is a $m_2 \times 1$ vector with all elements 1 and $b_{2s} = \left(b_{21s}, ..., b_{2q_2s}\right)'$. Also, $\varepsilon_{2s} \sim N(0, \sigma^2)$ is rewritten in accordance with r_s. We assume the following combinations for simulation:
$p = 3$, $m_1 = 5$, $q_1 = 5$, or $q_1 = 15$, and so $n_1 = 25$ or 75, $m_2 = 1$, $q_2 = 2$, $\left(\sigma^2, \sigma_1^2, \sigma_2^2\right) = (2, 0.2, 0.2)$ or $(8, 0.8, 0.8)$. $b_{1s} \sim N\left(0, \sigma_1^2 I_{q_1}\right)$, $b_{2s} \sim N\left(0, \sigma_2^2 I_{q_2}\right)$, $L_s \sim N(0, I_{n_1} \otimes \Lambda)$, $\Lambda = diag(0.002, 0.002, 0.002)$, $\Lambda = diag(0.008, 0.008, 0.008)$, $\Lambda = diag(0.01, 0.01, 0.01)$ and $\Lambda = diag(0.05, 0.05, 0.05)$.

The experiment is replicated 1000 times and then we assess the performance of the estimators via its scaler mean squared error (MSE) values which is defined as $MSE\left(\hat{\beta}\right) = \frac{1}{1000}\sum\limits_{s=1}^{1000}\left(\hat{\beta}_s - \beta\right)'\left(\hat{\beta}_s - \beta\right)$, where $\hat{\beta}_s$ is any estimator considered in this study in the sth repetition.

For relative comparison, we consider $\hat{\beta}$ as the baseline estimator. For any estimator $\hat{\beta}^*$, we compute the relative mean square (RMSE) as $RMSE(\hat{\beta} : \hat{\beta}^*) = \frac{MSE(\hat{\beta})}{MSE(\hat{\beta}^*)}$. It is clear that if RMSE is greater than one, it indicates the superiority of the estimator $\hat{\beta}^*$ over $\hat{\beta}$.

The simulation study was conducted using R software. For four different levels of multicollinearity, the results are presented in Tables 2.2 to 2.9. As can be seen from the results, for all cases, $\hat{\beta}_{nmre}$ has the smaller MSE values than the $\hat{\beta}_{mre}$ and $\hat{\beta}_{me}$. Moreover, the performance of the estimators depends on the degree of correlation ρ, the variance of random variables $\left(\sigma^2, \sigma_1^2, \sigma_2^2\right)$, the variance of measurement error Λ and sample size n_1. When $\left(\sigma^2, \sigma_1^2, \sigma_2^2\right)$, Λ and n_1 are fixed and ρ increases, the simulated MSE values of all estimators increase. With increasing the Λ, the simulated MSE values of all estimators increase for fixed ρ, n_1 and $\left(\sigma^2, \sigma_1^2, \sigma_2^2\right)$ values. Also, as $\left(\sigma^2, \sigma_1^2, \sigma_2^2\right)$ increases from $(2, 0.2, 0.2)$ to $(8, 0.8, 0.8)$, the simulated MSE values of all estimators increase for fixed ρ , Λ and n_1. Finally, as n_1 increases from 25 to 75 for fixed ρ, Λ and $\left(\sigma^2, \sigma_1^2, \sigma_2^2\right)$ values, the simulated MSE values of all estimators decrease.

2.8 Conclusion

In this chapter, we proposed the new mixed ridge estimator (NMRE) for the vector of parameters using Nakamura's corrected score function in a linear mixed model with measurement error on fixed effects to overcome multicollinearity. The new estimator we proposed is a combination of the mixed estimator (ME) and ridge estimator (RE). We investigated the asymptotic property of the new estimator, and proved that the NMRE is superior to the other estimators in terms of MSEM criterion under certain conditions. Finally, we illustrate our findings with a data example and simulation study.

Acknowledgments. ”The research of this chapter has been published in the journal of Communications in Statistics-Simulation and Computation, refer to [64]”

	$\Lambda = diag(0.002, 0.002, 0.002)$			
$(\sigma^2, \sigma_1^2, \sigma_2^2)$	$(2, 0.2, 0.2)$		$(8, 0.8, 0.8)$	
	n_1=25	n_1=75	n_1=25	n_1=75
$\hat{\beta}$	0.9547	0.2468	3.8190	0.9873
$\hat{\beta}_{re}$	0.5910	0.1853	1.9893	0.5610
$\hat{\beta}_{me}$	0.7869	0.2324	3.1477	0.9297
$\hat{\beta}_{mre}$	0.5244	0.1774	1.8224	0.5485
$\hat{\beta}_{nmre}$	0.5127	0.1766	1.7194	0.5378
$RMSE(\hat{\beta}_{me}; \hat{\beta}_{nmre})$	1.53	1.31	1.83	1.72
$RMSE(\hat{\beta}_{mre}; \hat{\beta}_{nmre})$	1.02	1.01	1.05	1.02
$d_1' D_1^{-1} d_1$	0.0173	0.0177	0.067	0.0431
$d_1'(D_2 + d_2 d_2')^{-1} d_1$	0.0171	0.0156	0.086	0.0428
	$\Lambda = diag(0.008, 0.008, 0.008)$			
$\hat{\beta}$	1.0191	0.2598	4.0763	1,0394
$\hat{\beta}$	0.6330	0.1947	2.1453	0.5917
$\hat{\beta}_{me}$	0.8341	0.2440	3.3364	0.9761
$\hat{\beta}_{mre}$	0.5566	0.1859	1.9444	0.5768
$\hat{\beta}_{nmre}$	0.5459	0.1853	1.8483	0.5659
$RMSE(\hat{\beta}_{me}; \hat{\beta}_{nmre})$	1.52	1.31	1.80	1.72
$RMSE(\hat{\beta}_{mre}; \hat{\beta}_{nmre})$	1.02	1.003	1.05	1.02
$d_1' D_1^{-1} d_1$	0.0156	0.0578	0.0843	0.0421
$d_1'(D_2 + d_2 d_2')^{-1} d_1$	0.0044	0.0175	0.0078	0.0043

Table 2.2: The estimated MSE and corresponding RMSE values of $\hat{\beta}$, $\hat{\beta}_{re}$, $\hat{\beta}_{me}$, $\hat{\beta}_{mre}$, $\hat{\beta}_{nmre}$ at $\rho = 0.85$ with $\Lambda = diag(0.002, 0.002, 0.002)$ and $diag(0.008, 0.008, 0.008)$.

	$\Lambda = diag(0.01, 0.01, 0.01)$			
$(\sigma^2, \sigma_1^2, \sigma_2^2)$	$(2, 0.2, 0.2)$		$(8, 0.8, 0.8)$	
	n_1=25	n_1=75	n_1=25	n_1=75
$\hat{\beta}$	1.0422	0.2644	4.1684	1.0577
$\hat{\beta}_{re}$	0.6481	0.19808	2.2018	0.6026
$\hat{\beta}_{me}$	0.8513	0.2481	3.4043	0.9923
$\hat{\beta}_{mre}$	0.5690	0.1889	1.9884	0.5868
$\hat{\beta}_{nmre}$	0.5579	0.1883	1.8884	0.5758
$RMSE(\hat{\beta}_{me}; \hat{\beta}_{nmre})$	1.52	1.31	1.80	1.72
$RMSE(\hat{\beta}_{mre}; \hat{\beta}_{nmre})$	1.02	1.003	1.05	1.02
$d_1' D_1^{-1} d_1$	0.0224	0.0174	0.0976	0.0419
$d_1'(D_2 + d_2 d_2')^{-1} d_1$	0.0165	0.0169	0.0063	0.0408
	$\Lambda = diag(0.05, 0.05, 0.05)$			
$\hat{\beta}$	1.7839	0.3911	7.1226	1.5627
$\hat{\beta}_{re}$	1.1637	0.2901	4.1528	0.9124
$\hat{\beta}_{me}$	1.3879	0.3593	5.5247	1.4344
$\hat{\beta}_{mre}$	0.9701	0.2710	3.5210	0.8671
$\hat{\beta}_{nmre}$	0.9582	0.2608	3.4158	0.8551
$RMSE(\hat{\beta}_{me}; \hat{\beta}_{nmre})$	1.45	1.32	1.61	1.67
$RMSE(\hat{\beta}_{mre}; \hat{\beta}_{nmre})$	1.01	1.04	1.03	1.01
$d_1' D_1^{-1} d_1$	0.0142	0.0139	0.0218	0.0317
$d_1'(D_2 + d_2 d_2')^{-1} d_1$	0.0048	0.0166	0.0024	0.0305

Table 2.3: The estimated MSE and corresponding RMSE values of $\hat{\beta}$, $\hat{\beta}_{re}$, $\hat{\beta}_{me}$, $\hat{\beta}_{mre}$, $\hat{\beta}_{nmre}$ at $\rho = 0.85$ with $\Lambda = diag(0.01, 0.01, 0.01)$ and $\Lambda = diag(0.05, 0.05, 0.05)$

	$\Lambda = diag(0.002, 0.002, 0.002)$			
$(\sigma^2, \sigma_1^2, \sigma_2^2)$	$(2, 0.2, 0.2)$		$(8, 0.8, 0.8)$	
	n_1=25	n_1=75	n_1=25	n_1=75
$\hat{\beta}$	1.3940	0.3584	5.5764	1.4339
$\hat{\beta}_{re}$	0.8015	0.2482	2.7809	0.7506
$\hat{\beta}_{me}$	1.0812	0.3280	4.3256	1.3122
$\hat{\beta}_{mre}$	0.6895	0.2340	2.4684	0.7300
$\hat{\beta}_{nmre}$	0.6620	0.2323	2.2920	0.7049
$RMSE(\hat{\beta}_{me}; \hat{\beta}_{nmre})$	1.63	1.41	1.88	1.03
$RMSE(\hat{\beta}_{mre}; \hat{\beta}_{nmre})$	1.04	1.01	1.07	1.86
$d_1' D_1^{-1} d_1$	0.0356	0.0434	0.0159	0.0186
$d_1'(D_2 + d_2 d_2')^{-1} d_1$	0.0081	0.0182	0.0016	0.0244
	$\Lambda = diag(0.008, 0.008, 0.008)$			
$\hat{\beta}$	1.5391	0.3870	6.1562	1.5483
$\hat{\beta}_{re}$	0.8934	0.2674	3.1315	0.8159
$\hat{\beta}_{me}$	1.1796	0.3522	4.7186	1.4093
$\hat{\beta}_{mre}$	0.7565	0.2507	2.7288	0.7878
$\hat{\beta}_{nmre}$	0.7318	0.2490	2.5603	0.7624
$RMSE(\hat{\beta}_{me}; \hat{\beta}_{nmre})$	1.61	1.41	1.84	1.84
$RMSE(\hat{\beta}_{mre}; \hat{\beta}_{nmre})$	1.03	1.01	1.06	1.03
$d_1' D_1^{-1} d_1$	0.0697	0.0184	0.0531	0.0223
$d_1'(D_2 + d_2 d_2')^{-1} d_1$	0.0071	0.0181	0.0018	0.0136

Table 2.4: The estimated MSE and corresponding RMSE values of $\hat{\beta}$, $\hat{\beta}_{re}$, $\hat{\beta}_{me}$, $\hat{\beta}_{mre}$, $\hat{\beta}_{nmre}$ at $\rho = 0.9$ with $\Lambda = diag(0.002, 0.002, 0.002)$ and $\Lambda = diag(0.008, 0.008, 0.008)$.

$(\sigma^2,\sigma_1^2,\sigma_2^2)$	$(2,0.2,0.2)$		$(8,0.8,0.8)$	
	n_1=25	n_1=75	n_1=25	n_1=75
	$\Lambda = diag(0.01,0.01,0.01)$			
$\hat{\beta}$	1.5931	0.3973	6.3720	1.5896
$\hat{\beta}_{re}$	0.9280	0.2744	3.2641	0.8396
$\hat{\beta}_{me}$	1.2174	0.3611	4.8668	1.4442
$\hat{\beta}_{mre}$	0.7826	0.2657	2.8304	0.8088
$\hat{\beta}_{nmre}$	0.7571	0.2567	2.6582	0.7833
$RMSE(\hat{\beta}_{me};\hat{\beta}_{nmre})$	1.59	1.40	1.83	1.84
$RMSE(\hat{\beta}_{mre};\hat{\beta}_{nmre})$	1.03	1.03	1.06	1.03
$d_1'D_1^{-1}d_1$	0.056	0.0183	0.0194	0.0767
$d_1'(D_2+d_2d_2')^{-1}d_1$	0.065	0.0178	0.273	0.0217
	$\Lambda = diag(0.05,0.05,0.05)$			
$\hat{\beta}$	4.0329	0.7414	16.0832	2.9641
$\hat{\beta}_{re}$	2.6861	0.5118	10.0939	1.6855
$\hat{\beta}_{me}$	2.9735	0.6460	11.412	2.5784
$\hat{\beta}_{mre}$	2.1976	0.4572	7.9638	1.5823
$\hat{\beta}_{nmre}$	2.0109	0.4516	7.7560	1.5067
$RMSE(\hat{\beta}_{me};\hat{\beta}_{nmre})$	1.47	1.43	1.47	1.71
$RMSE(\hat{\beta}_{mre};\hat{\beta}_{nmre})$	1.09	1.01	1.02	1.05
$d_1'D_1^{-1}d_1$	0.0996	0.0123	0.0873	0.0724
$d_1'(D_2+d_2d_2')^{-1}d_1$	0.0033	0.0158	0.0031	0.0160

Table 2.5: The estimated MSE and corresponding RMSE values of $\hat{\beta}$, $\hat{\beta}_{re}$, $\hat{\beta}_{me}$, $\hat{\beta}_{mre}$, $\hat{\beta}_{nmre}$ at $\rho = 0.9$ with $\Lambda = diag(0.01,0.01,0.01)$ and $\Lambda = diag(0.05,0.05,0.05)$.

$(\sigma^2,\sigma_1^2,\sigma_2^2)$	$(2,0.2,0.2)$		$(8,0.8,0.8)$	
	n_1=25	n_1=75	n_1=25	n_1=75
	$\Lambda=diag(0.002,0.002,0.002)$			
$\hat{\beta}$	2.7791	0.7070	11.117	2.8283
$\hat{\beta}_{re}$	1.4497	0.4208	5.3061	1.3339
$\hat{\beta}_{me}$	1.8811	0.6018	7.5271	2.4076
$\hat{\beta}_{mre}$	1.1487	0.3832	4.3150	1.2689
$\hat{\beta}_{nmre}$	1.0743	0.3732	3.9077	1.1840
$RMSE(\hat{\beta}_{me};\hat{\beta}_{nmre})$	1.75	1.61	1.92	2.03
$RMSE(\hat{\beta}_{mre};\hat{\beta}_{nmre})$	1.06	1.02	1.10	1.07
$d_1'D_1^{-1}d_1$	0.3151	.0081	0.0779	0.0949
$d_1'(D_2+d_2d_2')^{-1}d_1$	0.0385	0.0028	0.0986	0.0543
	$\Lambda=diag(0.008,0.008,0.008)$			
$\hat{\beta}$	3.4234	0.8267	13.6930	3.3070
$\hat{\beta}_{re}$	1.8516	0.4942	6.8825	1.6026
$\hat{\beta}_{me}$	2.2505	0.6929	8.9977	2.7715
$\hat{\beta}_{mre}$	1.4061	0.4417	5.3332	1.4849
$\hat{\beta}_{nmre}$	1.3355	0.4315	4.9388	1.3994
$RMSE(\hat{\beta}_{me};\hat{\beta}_{nmre})$	1.68	1.60	1.82	1.98
$RMSE(\hat{\beta}_{mre};\hat{\beta}_{nmre})$	1.05	1.02	1.07	1.06
$d_1'D_1^{-1}d_1$	0.0165	0.0070	0.0504	0.0148
$d_1'(D_2+d_2d_2')^{-1}d_1$	0.0424	0.0066	0.0147	0.0159

Table 2.6: The estimated MSE and corresponding RMSE of $\hat{\beta}$, $\hat{\beta}_{re}$, $\hat{\beta}_{me}$, $\hat{\beta}_{mre}$, $\hat{\beta}_{nmre}$ at $\rho=0.95$ with $\Lambda=diag(0.002,0.002,0.002)$ and $\Lambda=diag(0.008,0.008,0.008)$.

$(\sigma^2,\sigma_1^2,\sigma_2^2)$	$(2,0.2,0.2)$		$(8,0.8,0.8)$	
	n_1=25	n_1=75	n_1=25	n_1=75
	$\Lambda = diag(0.01,0.01,0.01)$			
$\hat{\beta}$	3.6939	0.8736	14.7739	3.4944
$\hat{\beta}_{re}$	2.0260	0.5232	7.5682	1.7102
$\hat{\beta}_{me}$	2.4057	0.7283	9.6197	2.9127
$\hat{\beta}_{mre}$	1.5151	0.4646	5.7631	1.5742
$\hat{\beta}_{nmre}$	1.4438	0.4543	5.3713	1.4878
$RMSE(\hat{\beta}_{me};\hat{\beta}_{nmre})$	1.66	1.60	1.79	1.95
$RMSE(\hat{\beta}_{mre};\hat{\beta}_{nmre})$	1.04	1.02	1.07	1.05
$d_1'D_1^{-1}d_1$	0.0245	0.0062	0.0541	0.0629
$d_1'(D_2+d_2d_2')^{-1}d_1$	0.0182	0.0527	0.0012	0.0055
	$\Lambda = diag(0.05,0.05,0.05)$			
$\hat{\beta}$	5.1986	0.9341	19.0513	0.0992
$\hat{\beta}_{re}$	4.8501	0.6592	16.7439	0.0673
$\hat{\beta}_{me}$	3.9699	1.7972	16.8328	3.7944
$\hat{\beta}_{mre}$	3.2950	1.6855	10.9651	2.3699
$\hat{\beta}_{nmre}$	3.1532	1.6418	10.0283	2.2497
$RMSE(\hat{\beta}_{me};\hat{\beta}_{nmre})$	1.25	1.09	1.67	1.68
$RMSE(\hat{\beta}_{mre};\hat{\beta}_{nmre})$	1.04	1.02	1.09	1.05
$d_1'D_1^{-1}d_1$	0.0132	0.0436	0.0423	0.0387
$d_1'(D_2+d_2d_2')^{-1}d_1$	0.0017	0.0064	0.0013	0.0068

Table 2.7: The estimated MSE and corresponding RMSE values of $\hat{\beta}$, $\hat{\beta}_{re}$, $\hat{\beta}_{me}$, $\hat{\beta}_{mre}$, $\hat{\beta}_{nmre}$ at $\rho = 0.95$ with $\Lambda = diag(0.01,0.01,0.01)$ and $\Lambda = diag(0.05,0.05,0.05)$.

$(\sigma^2, \sigma_1^2, \sigma_2^2)$	(2, 0.2, 0.2)		(8, 0.8, 0.8)	
	n_1=25	n_1=75	n_1=25	n_1=75
	$\Lambda = diag(0.002, 0.002, 0.002)$			
$\hat{\beta}$	17.0931	14.8230	30.7632	19.0817
$\hat{\beta}_{re}$	8.9553	6.3213	10.9437	9.0612
$\hat{\beta}_{me}$	7.2746	2.5910	29.0916	10.3640
$\hat{\beta}_{mre}$	4.6830	1.5220	18.5636	5.8516
$\hat{\beta}_{nmre}$	4.2632	1.3452	16.6849	4.9875
$RMSE(\hat{\beta}_{me}; \hat{\beta}_{nmre})$	1.70	1.92	1.74	2.07
$RMSE(\hat{\beta}_{mre}; \hat{\beta}_{nmre})$	1.09	1.13	1.11	1.17
$d_1' D_1^{-1} d_1$	0.0343	0.0321	0.0341	0.0416
$d_1'(D_2 + d_2 d_2')^{-1} d_1$	0.0162	0.0656	0.0099	0.0078
	$\Lambda = diag(0.008, 0.008, 0.008)$			
$\hat{\beta}$	70.0436	20.0675	200.0102	40.0023
$\hat{\beta}_{re}$	50.1390	9.6541	180.9023	28.9056
$\hat{\beta}_{me}$	47.7931	6.8282	190.6856	27.3153
$\hat{\beta}_{mre}$	42.9928	4.3891	173.7659	17.2719
$\hat{\beta}_{nmre}$	40.0814	4.3616	160.2468	17.0271
$RMSE(\hat{\beta}_{me}; \hat{\beta}_{nmre})$	1.19	1.56	1.18	1.59
$RMSE(\hat{\beta}_{mre}; \hat{\beta}_{nmre})$	1.07	1.01	1.08	1.01
$d_1' D_1^{-1} d_1$	0.0467	0.0456	0.0984	0.0124
$d_1'(D_2 + d_2 d_2')^{-1} d_1$	0.0162	0.0086	0.0556	0.0054

Table 2.8: The estimated MSE and corresponding RMSE of $\hat{\beta}$, $\hat{\beta}_{re}$, $\hat{\beta}_{me}$, $\hat{\beta}_{mre}$, $\hat{\beta}_{nmre}$ at $\rho = 0.99$ with $\Lambda = diag(0.002, 0.002, 0.002)$ and $\Lambda = diag(0.008, 0.008, 0.008)$.

$(\sigma^2,\sigma_1^2,\sigma_2^2)$	$(2,0.2,0.2)$		$(8,0.8,0.8)$	
	n_1=25	n_1=75	n_1=25	n_1=75
	$\Lambda = diag(0.01,0.01,0.01)$			
$\hat{\beta}$	21.5078	17.9899	1302.707	86.0312
$\hat{\beta}_{re}$	14.0706	12.6776	1273.428	55.5881
$\hat{\beta}_{me}$	96.7799	11.7140	301.1803	46.8935
$\hat{\beta}_{mre}$	82.1682	8.4606	236.2710	33.6338
$\hat{\beta}_{nmre}$	81.0281	8.1138	213.2983	32.9706
$RMSE(\hat{\beta}_{me};\hat{\beta}_{nmre})$	1.19	1.44	1.41	1.42
$RMSE(\hat{\beta}_{mre};\hat{\beta}_{nmre})$	1.01	1.04	1.10	1.02
$d_1'D_1^{-1}d_1$	0.0661	0.0921	0.0422	0.0504
$d_1'(D_2+d_2d_2')^{-1}d_1$	0.0697	0.0038	0.0889	0.0809
	$\Lambda = diag(0.05,0.05,0.05)$			
$\hat{\beta}$	325.8573	101.0034	388.0907	180.3223
$\hat{\beta}_{re}$	318.4693	50.7300	350.6043	98.3400
$\hat{\beta}_{me}$	122.7581	46.3770	356.1285	96.9524
$\hat{\beta}_{mre}$	106.1720	26.8124	262.1792	85.6824
$\hat{\beta}_{nmre}$	97.6989	25.2771	249.8838	82.6816
$RMSE(\hat{\beta}_{me};\hat{\beta}_{nmre})$	1.25	1.83	1.42	1.17
$RMSE(\hat{\beta}_{mre};\hat{\beta}_{nmre})$	1.08	1.06	1.04	1.03
$d_1'D_1^{-1}d_1$	0.0356	0.0878	0.0365	0.0329
$d_1'(D_2+d_2d_2')^{-1}d_1$	0.0162	0.0034	0.0044	0.0063

Table 2.9: The estimated MSE and corresponding RMSE values of $\hat{\beta}$, $\hat{\beta}_{re}$, $\hat{\beta}_{me}$, $\hat{\beta}_{mre}$, $\hat{\beta}_{nmre}$ at $\rho = 0.99$ with $\Lambda = diag(0.01,0.01,0.01)$ and $\Lambda = diag(0.05,0.05,0.05)$.

Chapter 3

The weighted ridge estimation for linear mixed models with measurement error under stochastic linear mixed restrictions

3.1 Introduction

A common statistical model is the linear mixed model (LMM), which provides flexibility in fitting models with various combinations of fixed and random effects. These models are commonly used for the analysis of a broad range of structures including longitudinal data, repeated measures data, clustered data, multivariate data, and correlated data.

In many practical situations, the observation of variables is subject to measurement errors, and ignoring these in data analysis can lead to inconsistent parameter estimation and invalid statistical inference. Therefore, it is necessary to extend LMMs by taking the effect of measurement errors into account. There are a few methods which can handle measurement errors. One method is based on the corrected score function proposed by [45]. Many authors, including [69] and [68], have considered a situation in which only the fixed effects variables are subject to measurement error and obtained a corrected score estimate (CSE) of the parameters by using Nakamura's approach.

Generally, fixed effects variables are assumed to be linearly independent. However, in practice, there may be strong linear relationships among the fixed effect variables. The assumption of independence is no longer valid in such a case, which causes the problem of multicollinearity. To overcome this problem, there have been many attempts to provide better estimators, an example of which is the incorporation of prior information available in the form of exact or stochastic restrictions [48]. To reduce the effect of multicollinearity when prior information comes to stochastic linear restrictions, [66] derived the stochastic restricted estimation of parameters in LMMs with the measurement errors, to reduce the effects of multicollinearity.

The ridge regression estimator is one of the most popular techniques to deal with the multicollinearity problem. [12] obtained ridge predictors in LMMs by using longitudinal data. [40] introduced the ridge estimator method to the multicollinearity problem and put forward the ridge predictors in LMMs. Both the ridge estimator and ridge predictor are derived in the context of Henderson's mixed model equations by [73].

The combination of two different estimators may inherit the advantages of both. See [35], [71] and [35] for examples using linear models, and [62] and [32] for examples involving LMMs. However, these authors do not consider measurement errors in their studies. As a remedy for multicollinearity, [66] developed the ridge estimation with additional stochastic linear restrictions for unknown parameter vector using Nakamura's approach to LMMs with measurement errors. When the prior information and the sample information are not equally important,

[36] propose the weighted mixed ridge estimator (WMRE) by unifying the sample and prior information in a linear model with additional stochastic linear restrictions. When there is a known measurement error, [19] introduce the weighted mixed ridge estimator (WMRE) for regression coefficients in linear models. Motivated by this, our primary aim in this paper is to obtain a new ridge-type estimator for LMMs with measurement error on fixed effects, called the weighted mixed ridge estimator (WMRE), when stochastic linear restrictions are available on both fixed and random effects and multicollinearity is present. We consider a case where variance parameters are not known and get the variance parameters estimations.

The rest of the paper is organized as follows: We define the model and propose a weighted ridge estimator in section 3.2. The proposed estimator (WMRE) is given in section 3.3. Asymptotic normality property of the proposed estimator is discussed in section 3.4. The WMRE is compared to the WRE and RE based on the mean square error matrix (MSEM) criterion in section 3.5 and some comparisons among linear combinations of the weighted mixed ridge estimator and the predictor are done in section 3.6. A data example and a simulation study are performed in section 3.7 to illustrate the theoretical findings of the proposed estimator. Finally, we draw some conclusions in section 3.8.

3.2 Model and estimation

We consider a linear mixed model with a measurement error in fixed effects as

$$y = Z\beta + Ub_1 + \varepsilon_1, \;\; X = Z + L, \tag{3.1}$$

where y is a $n_1 \times 1$ vector of observations and β is a $p \times 1$ vector of unobservable parameters, which are called fixed effects. Z and $U = [U_1|U_2|...|U_{c_1}]$ are $n_1 \times p$ and $n_1 \times q_1$ matrices of the regressor, respectively, where U_i is a $n_1 \times q_{1i}$ known design matrix of the random effects factor i, such that $q_1 = \sum_{i=1}^{c_1} q_{1i}$. Also, $b_1 = (b'_{11}, ..., b'_{1c_1})'$, where b_{1i} is a $q_{1i} \times 1$ vector of unobservable random effects from $N\left(0, \sigma_i^2 I_{q_{1i}}\right)$, $i = 1, ..., c_1$. The error term ε_1 is a $n_1 \times 1$ vector of unobservable random errors from $N\left(0, \sigma^2 I_{n_1}\right)$. X is the observed value of Z with the measurement error L, where L is an $n_1 \times p$ random matrix from $N\left(0, I_{n_1} \otimes \Lambda\right)$ and Λ is a $p \times p$ matrix of known values. We assume that b_{1i}, ε_1 and Λ are mutually independent. b_1 and y are jointly distributed as $\begin{bmatrix} b_1 \\ y \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ Z\beta \end{bmatrix}, \begin{bmatrix} \sigma^2\Sigma_1 & \sigma^2\Sigma U'_1 \\ \sigma^2 U\Sigma_1 & \sigma^2 V_1 \end{bmatrix}\right)$, where Σ_1 is a block diagonal matrix with the i^{th} block being $\gamma_{1i} I_{q_{1i}}$ for $\gamma_{1i} = \sigma_{1i}^2/\sigma^2$, $i = 1, ..., c_1$, so that $\gamma_1 = (\gamma_{11}, ..., \gamma_{1c_1})'$. Therefore, y has a multivariate normal distribution with $E(y) = Z\beta$ and $Var(y) = \sigma^2 V_1$ in which $V_1 = I_{n_1} + U\Sigma_1 U'$. The conditional distribution of $b_1|y$ is $N\left(\Sigma_1 U' V_1^{-1}(y - Z\beta), \sigma^2\Sigma_1 T\right)$, where $T = \left(I_{n_1} + U'U\Sigma_1\right)^{-1}$.

If the elements of γ_1 are known, the corrected score estimator (CSE) of β, σ^2 are given as $\hat{\beta} = \left[X'V_1{}^{-1}X - \mathrm{tr}\left(V_1^{-1}\right)\Lambda\right]^{-1} X'V_1^{-1}y$, $\hat{\sigma}^2 = \frac{1}{n_1}\left(y'V_1^{-1}y - \hat{\beta}'X'V_1^{-1}y\right)$, and the corrected score predictor of (CSP) of b_1 is given as $\hat{b_1} = \Sigma_1 U' V_1^{-1}\left(y - X\hat{\beta}\right)$, (see [68] and [69]). If the elements of γ_1 are unknown, the CSE of σ_{1i}^2's are given as $\hat{\sigma}_{1i}^2 = \frac{\left[\hat{b}'_{1i}\hat{b}_{1i} - tr\left(\hat{D}'_i\hat{D}_i\right)\hat{\beta}'\Lambda\hat{\beta}\right]}{\left[q_{1i} - tr(T_{ii})\right]}$, where $\hat{b_{1i}} = \hat{D}_i \left(y - X\hat{\beta} \right)$, $\hat{D}_i = \hat{\gamma}_{1i} U_i' \hat{V}_1^{-1}$, $i = 1, ..., c_1$, and T_{ij} is the ij^{th} block of matrix $T = \left(I_{q_1} + U'U\Sigma_1 \right)^{-1}$. An iterative algorithm is used to compute the CSE of parameters (see [68]).

If there exists multicollinearity among the columns of matrix Z in model (3.1), we can assume that the vectors of β and b_2 are subject to the following stochastic linear restrictions

$$r = R\beta + Hb_2 + \varepsilon_2, \tag{3.2}$$

where r is an $n_2 \times 1$ observable random vector, R is an $n_2 \times p$ $(n_2 < p)$ known full row rank matrix, $H = [H_1 | H_2 ... | H_{c_2}]$ is an $n_2 \times q_2$ matrix of regressors, and H_i is an $n_2 \times q_{2i}$ known design matrix of the random effects factor i, such that $q_2 = \sum_{i=1}^{c_2} q_{2i}$. Also, $b_2 = \left(b_{21}', ..., b_{2c_2}' \right)'$ is a $q_2 \times 1$ vector of unobservable random effects from $N\left(0, \sigma^2 \Sigma_2\right)$, where Σ_2 is a block diagonal matrix with the i^{th} block being $\gamma_{2i} I_{q_{2i}}$ for $\gamma_{2i} = \sigma_{2i}^2 / \sigma^2$, $i = 1, ..., c_2$. The error term ε_2 is an $n_2 \times 1$ vector of unobservable random errors from $N\left(0, \sigma^2 I_{n_2}\right)$ and assumed to be independent of b_{2i}, ε_1 and L . Therefore, r has a multivariate normal distribution with $E(r) = R\beta$ and $Var(r) = \sigma^2 V_2$, in which $V_2 = I_{n_2} + H\Sigma_2 H'$. When the sample information given by (3.1) and the prior information given by (3.2) are not equally important, we extend the method of weighted mixed regression estimation using a linear model, as suggested by [52], to a linear mixed model with the measurement error. We consider the augmented model

$$y_r = Z_r\beta + U_r b_r + \varepsilon_r, \quad X_r = Z_r + L_r, \tag{3.3}$$

where $y_r = \begin{bmatrix} y \\ \sqrt{\omega} r \end{bmatrix}$, $Z_r = \begin{bmatrix} Z \\ \sqrt{\omega} R \end{bmatrix}$, $U_r = \begin{bmatrix} U & 0 \\ 0 & H \end{bmatrix}$, $b_r = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$, $\varepsilon_r = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \end{bmatrix}$ and $L_r = \begin{bmatrix} L \\ 0 \end{bmatrix}$. The weight $\omega \, (0 \leqslant \omega \leqslant 1)$ reflects the degree of importance of the prior information in relation to the sample information. The error term ε_r is a random vector with $E(\varepsilon_r) = 0$ and $Var(\varepsilon_r) = \sigma^2 I_n$, where $I_n = \begin{bmatrix} I_{n_1} & 0 \\ 0 & I_{n_2} \end{bmatrix}$ and $n = n_1 + n_2$. Also, the conditional distribution of $b_r | y_r$ is $N\left(\Sigma_r U_r' V_r^{-1} (y_r - Z_r\beta), \sigma^2 \Sigma_r T_r \right)$, where $V_r = I_n + U_r \Sigma_r U_r'$, with $\Sigma_r = \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix}$ a block diagonal matrix with the i^{th} block being $\gamma_{ji} I_{q_{ji}}$, $\left(i = 1, ..., c_j, \; j = 1, 2 \right)$, $\gamma_r = (\gamma_{11}, ..., \gamma_{1c_1}, \gamma_{21}, ..., \gamma_{2c_2})'$ and $T_r = \left(I_q + U_r' U_r \Sigma_r \right)^{-1}$ with $I_q = \begin{bmatrix} I_{q_1} & 0 \\ 0 & I_{q_2} \end{bmatrix}$ and $q = q_1 + q_2$.

Now, we obtain the corrected log-likelihood of y_r and the conditional corrected log-likelihood of $b_r | y_r$ for model (3.3) as:

$$\begin{aligned} \ell^*\left(\beta, \sigma^2, \gamma_r, \omega; X, y, r\right) &= \frac{-n}{2} \log\left(2\pi\sigma^2\right) - \frac{1}{2} \log|V_1| - \frac{1}{2} \log|V_2| \\ &- \frac{1}{2\sigma^2} \left[(y - X\beta)' V_1^{-1} (y - X\beta) + \omega (r - R\beta)' V_2^{-1} (r - R\beta) \right] \\ &+ \frac{1}{2\sigma^2} \left[tr\left(V_1^{-1}\right) \beta' \Lambda \beta \right], \end{aligned}$$

and

$$\ell^*_{b_r}\left(\beta,\sigma^2,\gamma_r,\omega;X,y,r\right)=\frac{-q}{2}\log\left(2\pi\sigma^2\right)-\frac{1}{2}\log\left(|\Sigma_rT_r|\right)$$
$$-\frac{1}{2\sigma^2}\left[b_r-\Sigma_rU_r'V_r^{-1}\left(y_r-X_r\beta\right)\right]'(\Sigma_rT_r)^{-1}\times\left[b_r-\Sigma_rU_r'V_r^{-1}\left(y_r-X_r\beta\right)\right]$$
$$+\frac{1}{2\sigma^2}tr\left(I_{n_1}-V_1^{-1}\right)\beta'\Lambda\beta.$$

Let $\ell_1\left(\sigma^2,\gamma_1;Z,y\right)=\ell\left(\tilde{\beta}\left(\gamma_1\right),\sigma^2,\gamma_1;Z,y\right)$, in which $\tilde{\beta}=\tilde{\beta}(\gamma_1)$ is an ML estimate of β, and $\ell_1^*\left(\sigma^2,\gamma_1;X,y\right)=\ell^*\left(\hat{\beta}\left(\gamma_1\right),\sigma^2,\gamma_1;X,y\right)$, in which $\hat{\beta}=\hat{\beta}\left(\gamma_1\right)$ is the solution of $\frac{\partial}{\partial\beta}\ell^*\left(\beta,\sigma^2,\gamma_1;X,y\right)=0$. Also, let E^* denote the conditional mean with respect to X given y, then ℓ^*, ℓ_1^* and $\ell^*_{b_r}$ have the following properties:

$$E^*\left[\frac{\partial}{\partial\beta}\ell^*\left(\beta,\sigma^2,\gamma_r,\omega;X,y,r\right)\right]=\frac{\partial}{\partial\beta}\ell\left(\beta,\sigma^2,\gamma_r,\omega;Z,y,r\right),$$
$$E^*\left[\frac{\partial}{\partial\sigma^2}\ell_1^*\left(\sigma^2,\gamma_r,\omega;X,y,r\right)\right]=\frac{\partial}{\partial\sigma^2}\ell_1\left(\sigma^2,\gamma_r,\omega;Z,y,r\right),$$
$$E^*\left[\frac{\partial}{\partial\gamma_{ji}}\ell_1^*\left(\sigma^2,\gamma_r,\omega;X,y,r\right)\right]=\frac{\partial}{\partial\gamma_{ji}}\ell_1\left(\sigma^2,\gamma_r,\omega;Z,y,r\right),$$
$$i=1,...,c_j,\quad j=1,2,$$
$$E^*\left[\frac{\partial}{\partial b_r}\ell^*_{b_r}\left(\beta,\sigma^2,\gamma_r,\omega;X,y,r\right)\right]=\frac{\partial}{\partial b_r}\ell_{b_r}\left(\beta,\sigma^2,\gamma_r,\omega;Z,y,r\right),$$

If the elements of γ_r are known, we can obtain the weighted mixed estimator/predictor (WME/P) of the parameters by differentiating $\ell^*\left(\beta,\sigma^2,\gamma_r,\omega;X,y,r\right)$, $\ell_1^*\left(\sigma^2,\gamma_r,\omega;X,y,r\right)$, and $\ell^*_{b_r}\left(\beta,\sigma^2,\gamma_r,\omega;X,y,r\right)$ with respect to β, σ^2 and b_r, respectively, as:

$$\hat{\beta}_{wme}=\left[X'V_1^{-1}X+\omega R'V_2^{-1}R-tr\left(V_1^{-1}\right)\Lambda\right]^{-1}\left(X'V_1^{-1}y+\omega R'V_2^{-1}r\right),$$

$$\hat{\sigma}^2_{wme}=\frac{1}{n}\left(y'V_1^{-1}y-\hat{\beta}'_{wme}X'V_1^{-1}y-\omega\hat{\beta}'_{wme}R'V_2^{-1}r+\omega r'V_2^{-1}r\right),$$

$$\hat{b}_{wmp}=\begin{bmatrix}\Sigma_1UV_1^{-1}\left(y-X\hat{\beta}_{wme}\right)\\ \Sigma_2H'V_2^{-1}\sqrt{\omega}\left(r-R\hat{\beta}_{wme}\right)\end{bmatrix}=\begin{bmatrix}\hat{b}_{1_{wmp}}\\ \hat{b}_{2_{wmp}}\end{bmatrix}.$$

Using lemma 1(presented in Appendix), we can rewrite $\hat{\beta}_{wme}$ as:

$$\hat{\beta}_{wme}=\hat{\beta}$$
$$+\omega\left[X'V_1^{-1}X-tr\left(V_1^{-1}\right)\Lambda\right]^{-1}R'\left[V_2+\omega R\left(X'V_1^{-1}X-tr\left(V_1^{-1}\right)\Lambda\right)R'\right]^{-1}$$
$$\times\left(r-R\hat{\beta}\right).$$

If the elements of γ_r are unknown, their WME, given by $\hat{\gamma}_{1i_{wme}}$ and $\hat{\gamma}_{2i_{wme}}$, are substituted back into Σ_r to obtain $\hat{\beta}_{wme}$, $\hat{\sigma}^2_{wme}$ and $\hat{b}_{wmp}$. For the WME of γ_{ji}'s, we obtain the WME of σ^2_{1i}'s and σ^2_{2i}'s as:

$$\hat{\sigma}^2_{1i_{wme}} = \frac{\left[\hat{b}'_{1i_{wmp}}\hat{b}_{1i_{wmp}} - tr\left(\hat{D}'_{i_{wme}}\hat{D}_{i_{wme}}\right)\hat{\beta}'_{wme}\Lambda\hat{\beta}_{wme}\right]}{\left[q_{1i} - tr\left(T_{ii}\right)\right]},\ i = 1,...,c_1,$$
$$\hat{b}_{1i_{wmp}} = \hat{D}_{i_{wme}}\left(y - X\hat{\beta}_{wme}\right),\ \hat{D}_{i_{wme}} = \hat{\gamma}_{1i_{wme}}U'_i\hat{V}_1^{-1},$$

and

$$\hat{\sigma}^2_{2i_{wme}} = \frac{\hat{b}'_{2i_{wmp}}\hat{b}_{2i_{wmp}}}{\left[q_{2i} - tr\left(F_{ii}\right)\right]},\ i = 1,...,c_2,$$
$$\hat{b}_{2i_{wmp}} = \hat{\gamma}_{2i_{wme}}H'_i\hat{V}_2^{-1}\sqrt{\omega}\left(r - R\hat{\beta}_{wme}\right).$$

We use the iterative algorithm to obtain the WME/P of parameters.

3.3 The weighted mixed ridge estimation

If multicollinearity still remains a problem with the restricted model, we need to apply innovative methods as a remedy for the multicollinearity problem. In this section, we propose the ridge estimate of β under weighted stochastic linear restrictions. We consider the augmented model

$$y_* = Z_*\beta + U_*b_r + \varepsilon_*,\ \ X_* = Z_* + L_*, \tag{3.4}$$

where, $y_* = \begin{bmatrix} y \\ \sqrt{\omega}r \\ 0 \end{bmatrix}, Z_* = \begin{bmatrix} Z \\ \sqrt{\omega}R \\ \sqrt{k}I_p \end{bmatrix}, U_* = \begin{bmatrix} U & 0 \\ 0 & H \\ 0 & 0 \end{bmatrix}, \varepsilon_* = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \end{bmatrix}, L_* = \begin{bmatrix} L \\ 0 \\ 0 \end{bmatrix}$. The error term ε_* is a vector with $E(\varepsilon_*) = 0$ and $Var(\varepsilon_*) = \sigma^2 I_N$, where $I_N = \begin{bmatrix} I_{n_1} & 0 & 0 \\ 0 & I_{n_2} & 0 \\ 0 & 0 & I_p \end{bmatrix}$ and $N = n_1 + n_2 + p$. Moreover, y_* has a multivariate normal distribution with $E(y_*) = Z_*\beta$ and $Var(y_*) = \sigma^2 V_*$, in which $V_* = U_*\Sigma_r U'_* + I_N$. The conditional distribution of $y_*|b_r$ is $N\left(\Sigma_r U'_* V_*^{-1}\left(y_* - Z_*\beta\right), \sigma^2\Sigma_r T_*\right)$, where $T_* = \left(I_q + U'_*U_*\Sigma_r\right)^{-1}$.

Now, we obtain the corrected log-likelihood of y_* and the conditional corrected log-likelihood of $b_r|y_*$ for model (3.4), respectively, as:

$$\ell^*\left(\beta, \sigma^2, \gamma_r, k, \omega; X, y, r\right) = \frac{-N}{2}\log\left(2\pi\sigma^2\right) - \frac{1}{2}\log|V_1| - \frac{1}{2}\log|V_2|$$
$$- \frac{1}{2\sigma^2}\left[(y - X\beta)'V_1^{-1}(y - X\beta) + \omega(r - R\beta)'V_2^{-1}(r - R\beta)\right]$$
$$- \frac{1}{2\sigma^2}\left[k\beta'\beta - tr\left(V_1^{-1}\right)\beta'\Lambda\beta\right],$$

and

$$\ell^*_{b_r}\left(\beta,\sigma^2,\gamma_r,k,\omega;X,y,r\right) = \frac{-q}{2}\log\left(2\pi\sigma^2\right) - \frac{1}{2}\log\left(|\Sigma_r T_*|\right)$$
$$-\frac{1}{2\sigma^2}\left[b_r - \Sigma_r U'_* V_*^{-1}\left(y_* - X_*\beta\right)\right]'\left(\Sigma_r T_*\right)^{-1}\left[b_r - \Sigma_r U'_* V_*^{-1}\left(y_* - X_*\beta\right)\right]$$
$$+\frac{1}{2\sigma^2} tr\left(I_{n_1} - V_1^{-1}\right)\beta'\Lambda\beta.$$

If the elements of γ_r are known, we can obtain the weighted mixed ridge estimates/predictor (WMRE/P) of parameters by differentiating $\ell^*\left(\beta,\sigma^2,\gamma_r,k,\omega;X,y,r\right)$, $\ell^*_1\left(\sigma^2,\gamma_r,k,\omega;X,y,r\right)$ and $\ell^*_{b_r}\left(\beta,\sigma^2,\gamma_r,k,\omega;X,y,r\right)$ with respect to β, σ^2 and b_r, as:

$$\hat{\beta}_{wmre} = \left[X'V_1^{-1}X + \omega R'V_2^{-1}R + kI_p - tr\left(V_1^{-1}\right)\Lambda\right]^{-1}\left(X'V_1^{-1}y + \omega R'V_2^{-1}r\right), \tag{3.5}$$

$$\hat{\sigma}^2_{wmre} = \frac{1}{N}\left(y'V_1^{-1}y - \hat{\beta}'_{wmre}X'V_1^{-1}y - \omega\hat{\beta}'_{wmre}R'V_2^{-1}r + \omega r'V_2^{-1}r\right),$$

$$\hat{b}_{wmrp} = \begin{bmatrix}\Sigma_1 U V_1^{-1}\left(y - X\hat{\beta}_{wmre}\right)\\ \Sigma_2 H'V_2^{-1}\sqrt{\omega}\left(r - R\hat{\beta}_{wmre}\right)\end{bmatrix} = \begin{bmatrix}\hat{b}_{1_{wmrp}}\\ \hat{b}_{2_{wmrp}}\end{bmatrix}.$$

Using lemma 1(presented in Appendix), we can rewrite $\hat{\beta}_{wmre}$ as:

$$\hat{\beta}_{wmre} = \hat{\beta}_{re}$$
$$+\omega\left[X'V_1^{-1}X - tr\left(V_1^{-1}\right)\Lambda + kI_p\right]^{-1}R'\left[V_2 + \omega R\left(X'V_1^{-1}X - tr\left(V_1^{-1}\right)\Lambda + kI_p\right)R'\right]^{-1}$$
$$\times\left(r - R\hat{\beta}_{re}\right),$$

in which $\hat{\beta}_{re} = \left[X'V_1^{-1}X - tr\left(V_1^{-1}\right)\Lambda + kI_p\right]^{-1}X'V_1^{-1}y$, is a ridge estimate of β. If the elements of γ_r are unknown, their WMRE are substituted back into Σ_r to obtain $\hat{\beta}_{wmre}$, $\hat{b}_{wmre}$ and $\hat{\sigma}^2_{wmre}$. For the WMRE of γ_{ji}'s, we obtain the WMRE of σ^2_{1i}'s and σ^2_{2i}'s as:

$$\hat{\sigma}^2_{1i_{wmre}} = \frac{\left[\hat{b}'_{1i_{wmrp}}\hat{b}_{1i_{wmrp}} - tr\left(\hat{D}'_{i_{wmre}}\hat{D}_{i_{wmre}}\right)\hat{\beta}'_{wmre}\Lambda\hat{\beta}_{wmre}\right]}{\left[q_{1i} - tr\left(T_{ii}\right)\right]}, i = 1,\ldots,c_1,$$
$$\hat{b}_{1i_{wmrp}} = \hat{D}_{i_{wmre}}\left(y - X\hat{\beta}_{wmre}\right), \hat{D}_{i_{wmre}} = \hat{\gamma}_{1i_{wmre}}U'_i\hat{V}_1^{-1},$$

and

$$\hat{\sigma}^2_{2i_{wmre}} = \frac{\hat{b}'_{2i_{wmrp}}\hat{b}_{2i_{wmrp}}}{\left[q_{2i} - tr\left(F_{ii}\right)\right]}, i = 1,\ldots,c_2,$$
$$\hat{b}_{2i_{wmrp}} = \hat{\gamma}_{2i_{wmre}}H'_i\hat{V}_2^{-1}\sqrt{\omega}\left(r - R\hat{\beta}_{wmre}\right).$$

We must use an iterative algorithm to derive the WMRE/P of the parameters.

3.4 Asymptotic normality properties of the proposed estimator

In this section, we will concentrate on the asymptotic properties of the fixed effects estimators. With this aim in mind, we shall assume that all the derivatives related to the log-likelihood exist and the parameter β is identifiable. It is also assumed that as n tends to infinity, the limits of $n^{-1}tr\left(V_1^{-1}\right)$, $n^{-1}\left(Z'V_1^{-2}Z\right)$, $n^{-1}\left(Z'V_1^{-1}Z+R'V_2^{-1}R\right)$, $n^{-1}\left(Z'V_1^{-1}Z+\omega R'V_2^{-1}R\right)$, $n^{-1}\left(Z'V_1^{-1}Z+kI_p\right)$ and $n^{-1}\left(Z'V_1^{-1}Z+\omega R'V_2^{-1}R+kI_p\right)$ exist, (see [69]).

Theorem 3.4.1. *$\hat{\beta}_{wmre}$ has asymptotic normal distribution with the mean vector $M_{(\omega,k)}^{-1}M_\omega\beta$ and covariance matrix*

$$M_{(\omega,k)}^{-1}\left[B+\sigma^2\left(Z'V_1^{-1}Z+\omega^2R'V_2^{-1}R\right)\right]M_{(\omega,k)}^{-1},$$

where $M_{(\omega,k)}=\left(Z'V_1^{-1}Z+\omega R'V_2^{-1}R+kI_p\right)$, $B=\left[\beta'Z'V_1^{-2}Z\beta+\sigma^2tr\left(V_1^{-1}\right)\right]\Lambda$ and $M_\omega=M_{(\omega,0)}$.

Proof. Since $E\left(X'V_1^{-1}X\right)=Z'V_1^{-1}Z+tr\left(V_1^{-1}\right)\Lambda$, (see [16]). we have

$$X'V_1^{-1}X=Z'V_1^{-1}Z+tr\left(V_1^{-1}\right)\Lambda+O_p\left(n^{1/2}\right), \tag{3.6}$$

It follows from (3.5) and (3.6) that

$$\begin{aligned}\hat{\beta}_{wmre}&=\left[n^{-1}\left(Z'V_1^{-1}Z+\omega R'V_2^{-1}R+kI_p\right)+O_p\left(n^{-1/2}\right)\right]^{-1}n^{-1}\left(X'V_1^{-1}y+\omega R'V_2^{-1}r\right)\\&=\left[I_p+O_p\left(n^{-1/2}\right)\right]^{-1}\left[n\left(Z'V_1^{-1}Z+\omega R'V_2^{-1}R+kI_p\right)\right]^{-1}n^{-1}\left(X'V_1^{-1}y+\omega R'V_2^{-1}r\right)\\&=\left[I_p+O_p\left(n^{-1/2}\right)\right]\left[n^{-1}\left(Z'V_1^{-1}Z+\omega R'V_2^{-1}R+kI_p\right)\right]^{-1}n^{-1}\left(X'V_1^{-1}y+\omega R'V_2^{-1}r\right),\end{aligned}$$

where $\left[I_p+O_p\left(n^{-1/2}\right)\right]^{-1}=\left[I_p+O_p\left(n^{-1/2}\right)\right]$ is obtained from Taylor series expansion. Therefore, we have

$$\begin{aligned}\sqrt{n}\hat{\beta}_{wmre}=&\left[I_p+O_p\left(n^{-1/2}\right)\right]\left[n^{-1}\left(Z'V_1^{-1}Z+\omega R'V_2^{-1}R+kI_p\right)\right]^{-1}\\&\times n^{-1/2}\left(X'V_1^{-1}y+\omega R'V_2^{-1}r\right).\end{aligned} \tag{3.7}$$

Moreover, since the limit of $G=n^{-1}\left(Z'V_1^{-1}Z+\omega R'V_2^{-1}R+kI_p\right)$ exists, then (3.7) can be written as:

$$\sqrt{n}\hat{\beta}_{wmre}=G^{-1}h+O_p\left(n^{-1/2}\right), \tag{3.8}$$

where $h=n^{-1/2}\left(X'V_1^{-1}y+\omega R'V_2^{-1}r\right)$ is asymptotically normal (see [16]). It thus follows from $E\left(X'V_1^{-1}y+\omega R'V_2^{-1}r\right)=\left(Z'V_1^{-1}Z+\omega R'V_2^{-1}R\right)\beta$ that $E(h)=n^{-1/2}M_\omega^{-1}\beta$. Consequently, we have $\sqrt{n}\left(\hat{\beta}_{wmre}\right)-M_{(\omega,k)}^{-1}M_\omega\beta=G^{-1}\left[h-E(h)\right]+O_p\left(n^{-1/2}\right)$, which indicates that

$\sqrt{n}\left(\hat{\beta}_{wmre} - M^{-1}_{(\omega,k)} M_{\omega}\beta\right)$ is asymptotically normal with mean zero. Furthermore, from (3.8) we have $AVar\left(\sqrt{n}\hat{\beta}_{wmre}\right) = G^{-1} Var(h)\, G^{-1}$. The variance of h can be obtained from

$$\begin{aligned} Var(h) &= E_{y_r}\left[Var(h|y_r)\right] + Var_{y_r}\left[E(h|y_r)\right] \\ &= n^{-1} E_{y_r}\left(y' V_1^{-2} y\right)\Lambda + n^{-1} Var_{y_r}\left(Z V_1^{-1} y + \omega R' V_2^{-1} r\right), \end{aligned}$$

where E_{y_r} and Var_{y_r} denote the expectation and variance with respect to the random vector $y_r' = (y', r')$. We obtain $E\left(y' V_1^{-2} y\right)$ and $Var\left(Z' V_1^{-1} y + \omega R' V_2^{-1} r\right)$ as:

$$E\left(y' V_1^{-2} y\right) = \beta' Z' V_1^{-2} Z\beta + \sigma^2 tr\left(V_1^{-1}\right),$$

$$Var\left(Z' V_1^{-1} y + \omega R' V_2^{-1} r\right) = \sigma^2\left(Z' V_1^{-1} Z + \omega^2 R' V_2^{-1} R\right),$$

therefore $Var(h) = n^{-1}\left[B + \sigma^2\left(Z' V_1^{-1} Z + \omega^2 R' V_2^{-1} R\right)\right]$, whose limit exists as n tends to infinity. Thus, $AVar\left(\hat{\beta}_{wmre}\right) = M^{-1}_{(\omega,k)}\left[B + \sigma^2\left(Z' V_1^{-1} Z + \omega^2 R' V_2^{-1} R\right)\right] M^{-1}_{(\omega,k)}$, which completes the proof. □

Corollary 3.4.1. *$\hat{\beta}_{wme}$ has an asymptotic normal distribution with mean vector β and covariance matrix $AVar\left(\hat{\beta}_{wme}\right) = M_{\omega}^{-1}\left[B + \sigma^2\left(Z' V_1^{-1} Z + \omega^2 R' V_2^{-1} R\right)\right] M_{\omega}^{-1}$, where $M_{\omega} = \left(Z' V_1^{-1} Z + \omega R' V_2^{-1} R\right)$.*

Corollary 3.4.2. *$\hat{\beta}_{me}$ has an asymptotic normal distribution with mean vector β and covariance matrix $AVar\left(\hat{\beta}_{me}\right) = M_R^{-1}\left(+\sigma^2 M_R\right) M_R^{-1}$, where $M_R = \left(Z' V_1^{-1} Z + R' V_2^{-1} R\right)$.*

Corollary 3.4.3. *$\hat{\beta}_{mre}$ has an asymptotic normal distribution with mean vector $M_k M^{-1}\beta$ and covariance matrix $AVar\left(\hat{\beta}_{mre}\right) = M_{Rk}^{-1}\left(B + \sigma^2 M_R\right) M_{Rk}^{-1}$, where $M_{Rk} = \left(Z' V_1^{-1} Z + R' V_2^{-1} R + kI_p\right)$.*

Corollary 3.4.4. *$\hat{\beta}_{re}$ has an asymptotic normal distribution with mean vector β and covariance matrix $AVar\left(\hat{\beta}_{re}\right) = M_k^{-1}\left(B + \sigma^2 M\right) M_k^{-1}$, where $M_k = \left(Z' V_1^{-1} Z + kI_p\right)$ and $M = \left(Z' V_1^{-1} Z\right)$.*

3.5 Mean square error comparison of the estimators

In this section, we will discuss the superiority of the $\hat{\beta}_{wmre}$ over the $\hat{\beta}_{wme}$ and $\hat{\beta}_{re}$, based on the mean square error matrix criterion. Firstly, we write the asymptotic mean square error matrix (AMSEM) of the $\hat{\beta}_{wmre}$, $\hat{\beta}_{wme}$ and $\hat{\beta}_{re}$ as:

$$AMSEM\left(\hat{\beta}_{wmre}\right) = M^{-1}_{(\omega,k)}\left[B + \sigma^2\left(Z' V_1^{-1} Z + \omega^2 R' V_2^{-1} R\right)\right] M^{-1}_{(\omega,k)} + k^2 M^{-1}_{(\omega,k)}\beta\beta' M^{-1}_{(\omega,k)},$$

$$AMSEM\left(\hat{\beta}_{wme}\right) = M_{\omega}^{-1}\left[B + \sigma^2\left(Z' V_1^{-1} Z + \omega^2 R' V_2^{-1} R\right)\right] M_{\omega}^{-1},$$

$$AMSEM\left(\hat{\beta}_{re}\right) = M_k^{-1}\left(B + \sigma^2 M\right) M_k^{-1} + k^2 M_k^{-1}\beta\beta' M_k^{-1}.$$

Theorem 3.5.1. *The estimator $\hat{\beta}_{wmre}$ is superior to the estimator $\hat{\beta}_{wme}$ using MSEM criterion if $d_1' D_1^{-1} d_1 \leqslant 1$.*

Proof. We consider the difference $AMSEM\left(\hat{\beta}_{wme}\right) - AMSEM\left(\hat{\beta}_{wmre}\right)$ as:

$$\begin{aligned}\Delta_1 = {} & M_\omega^{-1}\left[B + \sigma^2\left(Z'V_1^{-1}Z + \omega^2 R'V_2^{-1}R\right)\right]M_\omega^{-1} \\ & - M_{(\omega,k)}^{-1}\left[B + \sigma^2\left(Z'V_1^{-1}Z + \omega^2 R'V_2^{-1}R\right)\right]M_{(\omega,k)}^{-1} \\ & - k^2 M_{(\omega,k)}^{-1}\beta\beta' M_{(\omega,k)}^{-1} = D_1 - d_1 d_1',\end{aligned} \tag{3.9}$$

where,

$$\begin{aligned}D_1 = {} & kM_\omega^{-1}B + k\sigma^2 M_\omega^{-1}\left(Z'V_1^{-1}Z + \omega^2 R'V_2^{-1}R\right) + BkM_\omega^{-1} + k\sigma^2\left(Z'V_1^{-1}Z + \omega^2 R'V_2^{-1}R\right)M_\omega^{-1} \\ & + k^2 M_\omega^{-1}BM_\omega^{-1} + k^2\sigma^2 M_\omega^{-1}\left(Z'V_1^{-1}Z + \omega^2 R'V_2^{-1}R\right)M_\omega^{-1},\end{aligned}$$

and

$$d_1 = kM_{(\omega,k)}^{-1}\beta.$$

We note that D_1 is a p.d. matrix. Hence by applying lemma 2, we get $\Delta_1 \geqslant 0$ if $d_1' D_1^{-1} d_1 \leq 1$. From the above theorem, we conclude that the proposed estimator $\hat{\beta}_{wmre}$ can perform better than the estimator $\hat{\beta}_{wme}$ under certain conditions.

□

Based on the [71], we suggest a method to choose the parameter k, when ω is a fixed value. In this method, we derive k such that $\Delta_1 \geqslant 0$. The matrix difference $\Delta_1 \geqslant 0$ in (3.9) can be rewritten as

$$\begin{aligned}\Delta_1 = {} & M_{(\omega,k)}^{-1}\left[kM_\omega^{-1}B + k\sigma^2 M_\omega^{-1}\left(Z'V_1^{-1}Z + \omega^2 R'V_2^{-1}R\right) + BkM_\omega^{-1}\right. \\ & + k\sigma^2\left(Z'V_1^{-1}Z + \omega^2 R'V_2^{-1}R\right)M_\omega^{-1} + k^2 M_\omega^{-1}BM_\omega^{-1} \\ & \left. + k^2\sigma^2 M_\omega^{-1}\left(Z'V_1^{-1}Z + \omega^2 R'V_2^{-1}R\right)M_\omega^{-1} - k^2\beta\beta'\right]M_{(\omega,k)}^{-1}.\end{aligned}$$

We note that

$$\begin{aligned}& \left[kM_\omega^{-1}B + BkM_\omega^{-1} + k\sigma^2\left(Z'V_1^{-1}Z + \omega^2 R'V_2^{-1}R\right)M_\omega^{-1} + k^2 M_\omega^{-1}BM_\omega^{-1}\right. \\ & \quad \left. + k^2\sigma^2 M_\omega^{-1}\left(Z'V_1^{-1}Z + \omega^2 R'V_2^{-1}R\right)M_\omega^{-1}\right],\end{aligned}$$

is a p.d. matrix. Therefore, using lemma 1 (presented in Appendix), Δ_1 is p.d. if $k\sigma^2 M_\omega^{-1}\left(Z'V_1^{-1}Z + \omega^2 R'V_2^{-1}R\right) - k^2\beta\beta'$ is positive semi-definite (p.s.d.). Thus, a sufficient condition for $\hat{\beta}_{wmre}$ to be superior over $\hat{\beta}_{wme}$ is

$$k < \sigma^2\left[\beta' M_\omega\left(Z'V_1^{-1}Z + \omega^2 R'V_2^{-1}R\right)^{-1}\beta\right]^{-1}. \tag{3.10}$$

When $\omega = 1$, (3.10) reduces to the necessary and sufficient condition for the superiority of the estimator $\hat{\beta}_{mre}$ over the estimator $\hat{\beta}_{me}$. We replace the unknown parameters in k by unbiased estimators to obtain $\hat{k} = \hat{\sigma}^2\left[\hat{\beta}' M_\omega^{-1}\left(Z'\hat{V}_1^{-1}Z + \omega^2 R'\hat{V}_2^{-1}R\right)^{-1}\hat{\beta}\right]^{-1}$.

Theorem 3.5.2. *When* $\lambda_{\max}\left\{AVar\left(\hat{\beta}_{wmre}\right)\left[AVar\left(\hat{\beta}_{re}\right)\right]^{-1}\right\}<1$, $\hat{\beta}_{wmre}$ *is superior to the* $\hat{\beta}_{re}$ *in the MSEM sense, if* $d_1'\left(D_2+d_2d_2'\right)d_1\leq 1$.

Proof. Let us consider the difference

$$\Delta_2=AMSEM\left(\hat{\beta}_{re}\right)-AMSEM\left(\hat{\beta}_{wmre}\right)=D_2+d_2d_2'-d_1d_1',$$

where $D_2=AVar\left(\hat{\beta}_{re}\right)-AVar\left(\hat{\beta}_{wmre}\right)$, $d_1=bias\left(\hat{\beta}_{wmre}\right)=kM_{(\omega,k)}^{-1}\beta$ and $d_2=bias\left(\hat{\beta}_{re}\right)=kM_k^{-1}\beta$. It is obvious that $AVar\left(\hat{\beta}_{re}\right)>0$ and $AVar\left(\hat{\beta}_{wmre}\right)>0$, therefore, we get $D_2>0$, when $\lambda_{\max}\left\{AVar\left(\hat{\beta}_{wmre}\right)\left[AVar\left(\hat{\beta}_{re}\right)\right]^{-1}\right\}<1$ by applying lemma4 (presented in Appendix). Furthermore, from lemma 3 (presented in appendix), we have $\Delta_2\geq 0$ if $d_1'\left(D_2+d_2d_2'\right)d_1\leq 1$. □

3.6 Mean square error comparison of the predictors

Prediction of linear combination of β and b_1 can be expressed as $\mu=L\beta+Sb_1$, for specific matrix $L\in{}^{p\times s}$ and $S\in{}^{q\times s}$. This type of prediction problem was investigated by [46] and [62] for the situation $s=1$. We assume $b_2=0$ and obtain the predictor of μ under the WMRP and WMP. Following [62], The mean square error matrix (MSEM) of any estimator or predictor is defined as:

$$\begin{aligned}MSEM(\hat{\mu})=E\left[(\hat{\mu}-\mu)(\hat{\mu}-\mu)'\right]&=Var(\hat{\mu})+Var(\mu)\\&+bias(\hat{\mu})\,bias(\hat{\mu})'-\operatorname{cov}(\hat{\mu},\mu)-\operatorname{cov}(\mu,\hat{\mu}),\end{aligned}\tag{3.11}$$

where $bias\left(\hat{\beta}\right)=E\left(\hat{\beta}\right)-\beta$. To get asymptotic $MSEM(\hat{\mu}_{wmrp})$ from equation (3.11), we obtain $Var(\hat{\mu}_{wmrp})$, $Var(\mu)$, $bias(\hat{\mu}_{wmrp})$ and $\operatorname{cov}(\hat{\mu}_{wmrp},\mu)$ as:

$$\begin{aligned}Var(\hat{\mu}_{wmrp})&=QVar\left(\hat{\beta}_{wmre}\right)Q'+\sigma^2S'\Sigma_1U'V_1^{-1}U\Sigma_1S+Q\operatorname{cov}\left(\hat{\beta}_{wmre},y\right)V_1^{-1}U\Sigma_1S\\&+S'\Sigma_1U'V_1^{-1}\operatorname{cov}\left(y,\hat{\beta}_{wmre}\right)Q'\\&=QVar\left(\hat{\beta}_{wmre}\right)Q'\\&+\sigma^2\left[S'\Sigma_1U'V_1^{-1}U\Sigma_1S+QA_wX'V_1^{-1}U\Sigma_1S+S'\Sigma_1U'V_1^{-1}XA'_wQ'\right],\end{aligned}\tag{3.12}$$

where $\operatorname{cov}\left(\hat{\beta}_{wmre},y\right)=\operatorname{cov}\left(A_w\left(X'V_1^{-1}y+R'V_2^{-1}r\right),y\right)=\operatorname{cov}\left(A_w\left(X'V_1^{-1}y\right),y\right)=\sigma^2A_wX'$ and $A_w=\left[X'V_1^{-1}X+\omega R'V_2^{-1}R-tr\left(V_1^{-1}\right)\Lambda\right]^{-1}$.

$$Var(\mu)=Var\left(L'\beta+S'b\right)=\sigma^2S'\Sigma_1S,\tag{3.13}$$

$$\begin{aligned}bias(\hat{\mu}_{wmrp})=E(\hat{\mu}_{wmrp}-\mu)&=E\left(Q\hat{\beta}_{wmre}+S'\Sigma_1U'V_1^{-1}y-L'\beta-S'b\right)\\&=QE\left(\hat{\beta}_{wmre}-\beta\right)=Qbias\left(\hat{\beta}_{wmre}\right),\end{aligned}\tag{3.14}$$

$$\begin{aligned}\operatorname{cov}(\hat{\mu}_{wmrp},\mu) &= \operatorname{cov}\left(Q\hat{\beta}_{wmre}+S'\Sigma_1U'V_1^{-1}y, L'\beta+S'b\right)\\ &= Q\operatorname{cov}\left(\hat{\beta}_{wmre},b\right)S+S'\Sigma_1U'V_1^{-1}\operatorname{cov}(y,b)S\\ &= \sigma^2QAX'V_1^{-1}U\Sigma_1S+\sigma^2S'\Sigma_1U'V_1^{-1}U\Sigma_1S, \end{aligned} \quad (3.15)$$

where $\operatorname{cov}\left(\hat{\beta}_{wmre},b\right)=\operatorname{cov}\left(A_w\left(X'V_1^{-1}y+R'V_2^{-1}r\right),b\right)$. Then equations (3.12), (3.13), (3.14) and (3.15) are put into equation (3.11) to get

$$AMSEM(\hat{\mu}_{wmrp})=QAMSEM\left(\hat{\beta}_{wmre}\right)Q'+\sigma^2S'\left[\Sigma_1-\Sigma_1U'V_1^{-1}U\Sigma_1\right]S. \quad (3.16)$$

Similarly, we can obtain the predictor of μ under the WMP as:

$$AMSEM(\hat{\mu}_{wmp})=QAMSEM\left(\hat{\beta}_{wme}\right)Q'+\sigma^2S'\left[\Sigma_1-\Sigma_1U'V_1^{-1}U\Sigma_1\right]S. \quad (3.17)$$

We are now interested in knowing under which conditions $\hat{\mu}_{wmrp}$ is better than $\hat{\mu}_{wmp}$. For this, we investigate the difference

$$\Delta=AMSEM(\hat{\mu}_{wmp})-AMSEM(\hat{\mu}_{wmrp}), \quad (3.18)$$

and from equations (3.16) and (3.17), Δ in equation (3.17) can equivalently be written as $\Delta=Q\Delta_1Q'$, where $\Delta_1=AMSEM\left(\hat{\beta}_{wme}\right)-AMSEM\left(\hat{\beta}_{wmre}\right)$. Similarly, we derive that the superiority of $\hat{\mu}_{wmrp}$ over $\hat{\mu}_{rp}$ is equivalent to the superiority of $\hat{\beta}_{wmre}$ over $\hat{\beta}_{re}$. This means that if we want to compare two different predictors, we just need to compare the corresponding estimators under the MSEM criterion, which is discussed in previous sections.

3.7 Data example and simulation study

To evaluate the performance of the proposed estimator, we consider a sample of real data, Boston housing data taken from [24]. [69] used this data set and considered the data of 132 census tracts in the 15 districts of Boston (as a part of 506 observations on census tracts in the Boston Standard Metropolitan Statistical Area [SMSA] in 1970). The census tracts within each district are taken as repeated measurements. The pollution variable NOX is taken to have measurement errors. Therefore, they employed a linear mixed model with the measurement error. We use the same data in this example (as presented in [16]).

For this data set, the condition number of $\hat{\mathbf{Z}}'\hat{\mathbf{V}}_1^{-1}\hat{\mathbf{Z}}$ is equal to 159, which indicates that there is collinearity among the fixed effects variables. We can improve the CSE by adding stochastic linear restrictions $\sqrt{\omega}r=\sqrt{\omega}R\beta+Hb_2+\varepsilon_2$, $\varepsilon_2\sim N(0,I_7)$ to model (3.1), where the rows are the 133th-139th observations of the historical data, (see [66]).

Let ω be a fixed value, selected as 0.1, 0.5 and 1, and fit model (3.3) to the data. Then, obtain the estimate of k as $\hat{k}=\hat{\sigma}^2\left[\hat{\beta}'M_{\omega}^{-1}\left(\hat{Z}'\hat{V}_1^{-1}\hat{Z}+\omega^2R'\hat{V}_2^{-1}R\right)^{-1}\hat{\beta}'\right]^{-1}$ and write the model given by (3.4). The estimated MSE values of the estimators are obtained by replacing the corresponding theoretical MSE expressions. Table 3.1 presents the result. As can be observed in Table 3.1, the MSE values of the $\hat{\beta}_{wmre}$ are smaller than those of the $\hat{\beta}_{wme}$, $\hat{\beta}_{re}$ and $\hat{\beta}$. We can see that with increasing ω values, the estimated MSE values of the $\hat{\beta}_{wmre}$ and the $\hat{\beta}_{wme}$ become smaller, which indicates that we can get a more exact estimator of the parameter, with more depended prior information. Moreover, for all cases we have $\hat{d}_1'\hat{D}_1^{-1}\hat{d}_1 \leqslant 1$ and $d_1'\left(D_2+d_2d_2'\right)d_1 \leq 1$, which agrees with our theoretical findings in Theorems 3.5.1 and 3.5.2 which implies that the $\hat{\beta}_{wmre}$ can perform better than the $\hat{\beta}_{wme}$ and $\hat{\beta}_{re}$ for our selected k.

	$\hat{\beta}$	$\hat{\beta}_{re}$	$\hat{\beta}_{wme}$	$\hat{\beta}_{wrme}$	$d_1'D_1^{-1}d_1$	$d_1'\left(D_2+d_2d_2'\right)d_1 \leq 1$
$\omega = 0.1$	0.1759	0.16809	0.1624	0.1566	0.00007	0.000027
$\omega = 0.5$	0.1759	0.16808	0.1561	0.1530	0.00001	0.000024
$\omega = 1$	0.1759	0.16807	0.1547	0.1506	0.00013	0.000021

Table 3.1: The Estimated MSE values of $\hat{\beta}$, $\hat{\beta}_{re}$, $\hat{\beta}_{wme}$, $\hat{\beta}_{wrme}$

For convenience, plots of the estimated MSE values of the $\hat{\beta}_{wmre}$ and $\hat{\beta}_{wme}$ versus k are depicted in the interval $[0.002, 0.01]$ in Figures 3.1 and 3.2, for $\omega = 0.1$ and $\omega = 1$, respectively. Since the WME does not depend on k, its estimated MSE values are the same for all values. As ω increases from 0.1 to 1, the estimated MSE values of the WMRE and WME simultaneously decrease.

In order to further illustrate the behavior of the proposed estimator, we will perform a Monte Carlo simulation study by considering the different degrees of multicollinearity and different degrees of scalar weight. To achieve different degrees of multicollinearity, following [31], the fixed effect variables are generated as:

$$z_{kt} = \left(1-\rho^2\right)^{\frac{1}{2}} u_{kt} + \rho u_{kp+1},\ \ k = 1,\ldots,n_1,\ \ t = 1,\ldots,p,$$

where u_{kt}'s are independent standard normal pseudo-random numbers and ρ is specified so that the correlation between any two fixed effect variables is given by ρ^2. Three different sets of correlation were considered, corresponding to $\rho = 0.85, 0.9$, and 0.95. In this study, q_1 can be interpreted as the number of independent clusters and m_1 is the cluster size in a longitudinal study, so that $n_1 = m_1 \times q_1$ is the total size. We generate the sth set of simulated data as:

$$\begin{aligned} y_s &= Z\beta + Ub_{1s} + \varepsilon_{1s}, \qquad X_s = Z + L_s, \\ \sqrt{\omega}r_s &= \sqrt{\omega}R\beta + Hb_{2s} + \varepsilon_{2s},\ \ s = 1,2,\ldots,1000, \end{aligned}$$

where $y_s = \left(y_{11s},\ldots,y_{1m_1s},y_{21s},\ldots,y_{2m_1s},\ldots,y_{q_11s},\ldots,y_{q_1m_1s}\right)'$ and $b_{1s} = \left(b_{11s},\ldots,b_{1q_1s}\right)'$. we assume that $c_1 = 1$ and $Z^{(t)} = \left(Z_{11}^{(t)},\ldots,Z_{1m_1}^{(t)},Z_{21}^{(t)},\ldots,Z_{2m_1}^{(t)},\ldots,Z_{q_11}^{(t)},\ldots,Z_{q_1m_1}^{(t)}\right)$; $t = 1,2,3$, such that $Z = \left(Z^{(1)},Z^{(2)},Z^{(3)}\right)$. The vector β is chosen as the eigenvector corresponding to the largest eigenvalues of the $X'V_1^{-1}X$ matrix. We consider $U = I_{q_1} \otimes 1_{m_1}$, in which 1_{m_1} is an $m_1 \times 1$ vector where all elements are 1 and $\varepsilon_{1s} \sim N\left(0,\sigma^2\right)$ is rewritten in accordance with y_s. $R = \left(R^{(1)},R^{(2)},R^{(3)}\right)$, where $R^{(t)} = \left(R_{11}^{(t)},\ldots,R_{1m_2}^{(t)},R_{21}^{(t)},\ldots,R_{2m_2}^{(t)},\ldots,R_{q_11}^{(t)},\ldots,R_{q_1m_2}^{(t)}\right)$; $t = 1,2,3$, $R_{lt} \sim N(0,1)$; $l = 1,\ldots,n_2$ and $r_s = (r_{1s},r_{2s})'$. Also, $L = I_{q_2} \otimes 1_{m_2}$, in which 1_{m_2} is a $m_2 \times 1$ vector where all elements are 1 and $b_{2s} = \left(b_{21s},\ldots,b_{2q_1s}\right)'$. Also, $\varepsilon_{2s} \sim N(0,\sigma^2)$ is rewritten in

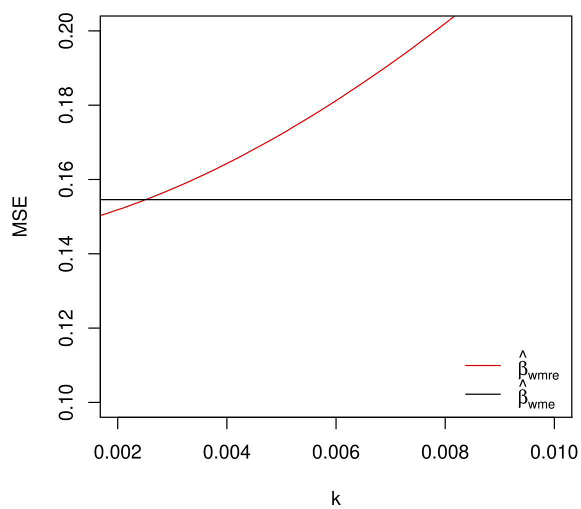

Figure 3.1: Estimated MSE values of the WME and WMRE with $\omega = 0.1$

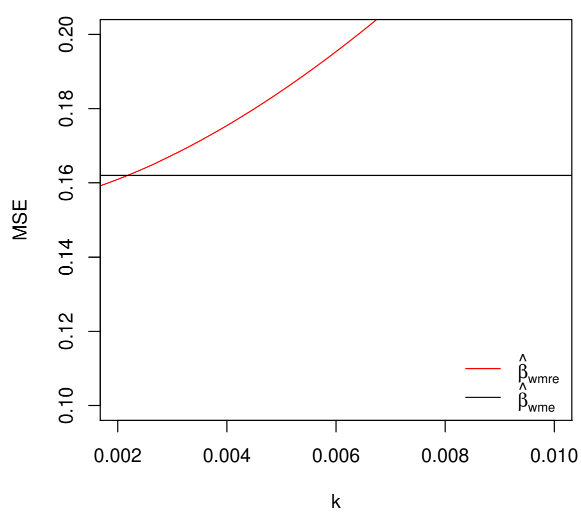

Figure 3.2: Estimated MSE values of the WME and WMRE with $\omega = 1$

accordance with r_s. We assume the following combinations for simulation:
$p=3$, $q_1=10$, $m_1=3$ or $m_1=7$ and so $n_1=30$ or 70, $m_2=1$, $q_2=2$, $\Lambda=diag(0.01,0.01,0.01)$, $diag(0.007,0.007,0.007)$ or $diag(0.001,0.001,0.001)$, $(\sigma^2,\sigma_1^2,\sigma_2^2)=(1,0.1,0.1)$ or $(6,0.6,0.6)$.

The experiment is replicated 1000 times and then we assess the performance of the estimators, via its scaler mean squared error (MSE) values, defined as:

$$MSE\left(\hat{\beta}\right)=\frac{1}{1000}\sum_{s=1}^{1000}\left(\hat{\beta}_s-\beta\right)'\left(\hat{\beta}_s-\beta\right),$$

where $\hat{\beta}_s$ is any estimator considered in this study in the sth repetition.

For relative comparison, we consider $\hat{\beta}$ as the baseline estimator. For any estimator $\hat{\beta}^*$ we compute the relative mean square (RMSE) as:

$$RMSE(\hat{\beta}:\hat{\beta}^*)=\frac{MSE\left(\hat{\beta}\right)}{MSE\left(\hat{\beta}^*\right)}.$$

It is clear that if RMSE is greater than one, it indicates the superiority of the estimator $\hat{\beta}^*$ over $\hat{\beta}$. The simulation study was conducted using R software. For four different levels of multicollinearity, the results are presented in Tables 3.2 to 3.10. As can be seen from the results, the performance of the estimators depends on the degree of correlation ρ, the degree of weight ω, the variance of random variables $\left(\sigma^2,\sigma_1^2,\sigma_2^2\right)$, the variance of measurement error Λ, and sample size n_1. We can derive the following results from Tables 3.2 to 3.10:

- With the increase of multicollinearity, MSE values of the WMRE, WME and RE increase, for fixed Λ, ω, n_1, and $\left(\sigma^2,\sigma_1^2,\sigma_2^2\right)$ values.
- With the increase of Λ, the simulated MSE values of the WMRE, WME and RE also increase, for fixed ρ, ω, n_1, and $\left(\sigma^2,\sigma_1^2,\sigma_2^2\right)$ values.
- As $\left(\sigma^2,\sigma_1^2,\sigma_2^2\right)$ increases from $(1,0.1,0.1)$ to $(6,0.6,0.6)$, the simulated MSE values of the WMRE, WME and RE increase, for fixed ρ, Λ, ω, and n_1.
- While n_1 increases from 30 to 70 for fixed ρ, Λ, ω, and $\left(\sigma^2,\sigma_1^2,\sigma_2^2\right)$ values, the simulated MSE values of the WMRE, WME and RE decrease.
- The estimated MSE values of the WMRE and WME become smaller when the value of ω, the level of the weight to the sample information and prior information, increases. It can therefore be concluded that we get a more exact estimator of the parameter with more depended prior information.
- For all cases, the WMRE has smaller MSE values than the WME and RE. In addition, after the 1000 times we consider the maximum values of $d_1'D_1^{-1}d_1$, and $d_1'\left(D_2+d_2d_2'\right)d_1$, we can see that for all cases the values of them are less than 1. Therefore, WMRE is superior to the WME and RE, which agrees with our theoretical finding in Theorems 3.5.1 and 3.5.2.

3.8 Conclusion

In this chapter, we proposed the WMRE for the vector of parameters using Nakamura's corrected score function in a linear mixed model with measurement error on fixed effects to overcome multicollinearity. We considered that prior information and sample information are not equally important; prior information is available on fixed and random effects. The estimator we proposed is a combination of the weighted mixed estimator (WME) and ridge estimator (RE). We investigated the asymptotic property of the new estimator, and proved that the WMRE is superior to the WME in terms of MSEM criterion under certain conditions. We also obtained the optimal value of parameter k based on the difference of MSEM of estimators. Finally, we performed a simulation study and data example, the results of both of which suggest that WMRE has a smaller MSE value compared to WME. As the level of ω increases, the estimated MSE values of the WMRE and WME become increasingly smaller, which implies that we can get a more exact estimator of the parameter when we get more depended prior information.

Acknowledgments. "The research of this chapter has been accepted in the journal of Communications in Statistics-Theory and Methods".

	$\Lambda = diag(0.001, 0.001, 0.001)$					
$(\sigma^2, \sigma_1^2, \sigma_2^2)$	$(1, 0.1, 0.1)$					
	$n_1 = 30$			$n_1 = 70$		
	w=0.1	w=0.5	w=1	w=0.1	w=0.5	w=1
$\hat{\beta}$	0.3402	0.3402	0.3402	0.1364	0.1364	0.1364
$\hat{\beta}_{re}$	0.2903	0.2889	0.2865	0.124595	0.124524	0.1243
$\hat{\beta}_{wme}$	0.3314	0.3119	0.2939	0.1346	0.1297	0.1249
$\hat{\beta}_{wmre}$	0.2838	0.2692	0.2549	0.1231	0.1189	0.1147
$RMSE(\hat{\beta}_{re}; \hat{\beta}_{wmre})$	1.023	1.073	1.123	1.012	1.046	1.083
$RMSE(\hat{\beta}_{wme}; \hat{\beta}_{wmre})$	1.16	1.15	1.15	1.09	1.09	1.08
$d_1' (D_2 + d_2 d_2') d_1$	0.00037	0.00011	0.00012	0.00033	0.000042	0.000044
$d_1' D_1^{-1} d_1$	0.1334	0.37	0.1725	0.1588	0.1581	0.1581
$(\sigma^2, \sigma_1^2, \sigma_2^2)$	$(6, 0.6, 0.6)$					
$\hat{\beta}$	2.0420	2.0420	2.0420	0.8186	0.8186	0.8186
$\hat{\beta}_{re}$	1.4455	1.4169	1.3705	0.6165	0.6131	0.6066
$\hat{\beta}_{wme}$	1.9886	1.8718	1.7638	0.8079	0.7785	0.7494
$\hat{\beta}_{wmre}$	1.4174	1.3457	1.2666	0.6106	0.5918	0.5694
$RMSE(\hat{\beta}_{re}; \hat{\beta}_{wmre})$	1.019	1.052	1.081	1.009	1.036	1.065
$RMSE(\hat{\beta}_{wme}; \hat{\beta}_{wmre})$	1.40	1.39	1.39	1.32	1.31	1.31
$d_1' (D_2 + d_2 d_2') d_1$	0.0050	0.0048	0.0046	0.00021	0.00022	0.00023
$d_1' D_1^{-1} d_1$	0.2032	0.2061	0.2062	0.2938	0.2932	0.2939

Table 3.2: The estimated MSE values and corresponding RMSE values of $\hat{\beta}$, $\hat{\beta}_{re}$, $\hat{\beta}_{wme}$ and $\hat{\beta}_{wmre}$ at $\rho = 0.85$ with $\Lambda = diag(0.001, 0.001, 0.001)$.

	$\Lambda = diag(0.007, 0.007, 0.007)$					
$(\sigma^2, \sigma_1^2, \sigma_2^2)$	$(1, 0.1, 0.1)$					
	$n_1 = 30$			$n_1 = 70$		
	w=0.1	w=0.5	w=1	w=0.1	w=0.5	w=1
$\hat{\beta}$	0.3603	0.3603	0.3603	0.1434	0.1434	0.1434
$\hat{\beta}_{re}$	0.3072	0.3056	0.3030	0.1309	0.1308	0.1306
$\hat{\beta}_{wme}$	0.3506	0.3293	0.3098	0.1415	0.1362	0.1310
$\hat{\beta}_{wmre}$	0.2999	0.2841	0.2686	0.1293	0.1248	0.1202
$RMSE(\hat{\beta}_{re}; \hat{\beta}_{wmre})$	1.024	1.075	1.128	1.012	1.048	1.089
$RMSE(\hat{\beta}_{wme}; \hat{\beta}_{wmre})$	1.16	1.15	1.15	1.09	1.09	1.08
$d_1'(D_2 + d_2 d_2')d_1$	0.00013	0.00013	0.00032	0.00035	0.00045	0.0004
$d_1' D_1^{-1} d_1$	0.1688	0.1678	0.1258	0.1540	0.1533	0.1534
$(\sigma^2, \sigma_1^2, \sigma_2^2)$	$(6, 0.6, 0.6)$					
$\hat{\beta}$	2.1617	2.1617	2.1617	0.8608	0.8608	0.8608
$\hat{\beta}_{RE}$	1.5348	1.5032	1.4530	0.6484	0.6447	0.6376
$\hat{\beta}_{WME}$	2.1031	1.9746	1.8566	0.8491	0.8172	0.7857
$\hat{\beta}_{WMRE}$	1.5038	1.4244	1.3374	0.6420	0.6212	0.5971
$RMSE(\hat{\beta}_{re}; \hat{\beta}_{wmre})$	1.0206	1.055	1.086	1.01	1.037	1.067
$RMSE(\hat{\beta}_{wme}; \hat{\beta}_{wmre})$	1.39	1.38	1.38	1.32	1.31	1.31
$d_1'(D_2 + d_2 d_2')d_1$	0.005	0.0049	0.0047	0.00022	0.00023	0.00024
$d_1' D_1^{-1} d_1$	0.1971	0.1999	0.2000	0.2870	0.2863	0.2871

Table 3.3: The estimated MSE values and corresponding RMSE values of $\hat{\beta}$, $\hat{\beta}_{re}$, $\hat{\beta}_{wme}$ and $\hat{\beta}_{wmre}$ at $\rho = 0.85$ with $\Lambda = diag(0.007, 0.007, 0.007)$.

	$\Lambda = diag(0.01, 0.01, 0.01)$					
$(\sigma^2, \sigma_1^2, \sigma_2^2)$	(1,0.1,0.1)					
	$n_1 = 30$			$n_1 = 70$		
	w=0.1	w=0.5	w=1	w=0.1	w=0.5	w=1
$\hat{\beta}$	0.3711	0.3711	0.3711	0.1472	0.1472	0.1472
$\hat{\beta}_{re}$	0.3163	0.3146	0.3119	0.1342	0.1341	0.1339
$\hat{\beta}_{wme}$	0.3609	0.3387	0.3183	0.1452	0.1396	0.1342
$\hat{\beta}_{wmre}$	0.3087	0.2921	0.2758	0.1325	0.1279	0.1231
$RMSE(\hat{\beta}_{re}; \hat{\beta}_{wmre})$	1.02	1.07	1.13	1.01	1.04	1.08
$RMSE(\hat{\beta}_{wme}; \hat{\beta}_{wmre})$	1.16	1.15	1.15	1,09	1.09	1.09
$d_1' (D_2 + d_2 d_2') d_1$	0.0035	0.0014	0.0033	0.0036	0.0047	0.0051
$d_1' D_1^{-1} d_1$	0.1654	0.1645	0.1236	0.1519	0.1512	0.1513
$(\sigma^2, \sigma_1^2, \sigma_2^2)$	(6,0.6,0.6)					
$\hat{\beta}$	2.2260	2.2260	2.2260	0.8832	0.8832	0.8832
$\hat{\beta}_{re}$	1.5831	1.5498	1.4978	0.6654	0.6614	0.6541
$\hat{\beta}_{wme}$	2.1646	2.0298	1.9072	0.8710	0.8376	0.8050
$\hat{\beta}_{wmre}$	1.5504	1.4666	1.3756	0.6587	0.6369	0.6117
$RMSE(\hat{\beta}_{re}; \hat{\beta}_{wmre})$	1.02	1.05	1.08	1.01	1.03	1.06
$RMSE(\hat{\beta}_{wme}; \hat{\beta}_{wmre})$	1.39	1.38	1.38	1.32	1.31	1.31
$d_1' (D_2 + d_2 d_2') d_1$	0.0051	0.0050	0.0048	0.0002	0.0024	0.0002
$d_1' D_1^{-1} d_1$	0.1941	0.1969	0.1969	0.2836	0.2829	0.2837

Table 3.4: The estimated MSE values and corresponding RMSE values of $\hat{\beta}$, $\hat{\beta}_{re}$, $\hat{\beta}_{wme}$ and $\hat{\beta}_{wmre}$ at $\rho = 0.85$ with $\Lambda = diag(0.01, 0.01, 0.01)$.

	$\Lambda = diag(0.001, 0.001, 0.001)$					
$(\sigma^2, \sigma_1^2, \sigma_2^2)$	$(1, 0.1, 0.1)$					
	$n_1 = 30$			$n_1 = 70$		
	w=0.1	w=0.5	w=1	w=0.1	w=0.5	w=1
$\hat{\beta}$	0.4916	0.4916	0.4916	0.1958	0.1958	0.1958
$\hat{\beta}_{re}$	0.4006	0.3965	0.3904	0.1729	0.1726	0.1721
$\hat{\beta}_{wme}$	0.4744	0.4370	0.4045	0.1923	0.1826	0.1735
$\hat{\beta}_{wmre}$	0.3891	0.3629	0.3371	0.1701	0.1622	0.1544
$RMSE(\hat{\beta}_{re}; \hat{\beta}_{wmre})$	1.0295	1.092	1.157	1.016	1.064	1.114
$RMSE(\hat{\beta}_{wme}; \hat{\beta}_{wmre})$	1.21	1.20	1.19	1.13	1.12	1.12
$d_1'(D_2 + d_2 d_2')d_1$	0.00029	0.00033	0.0007	0.00013	0.00018	0.00019
$d_1' D_1^{-1} d_1$	0.1471	0.14	0.1130	0.1538	0.1528	0.1531
$(\sigma^2, \sigma_1^2, \sigma_2^2)$	$(6, 0.6, 0.6)$					
$\hat{\beta}$	2.9505	2.9505	2.9505	1.1753	1.1753	1.1753
$\hat{\beta}_{re}$	2.0058	1.9420	1.8464	0.8379	0.8291	0.8133
$\hat{\beta}_{wme}$	2.8470	2.6230	2.4282	1.1541	1.0956	1.0411
$\hat{\beta}_{wmre}$	1.9574	1.8265	1.6820	0.8276	0.7909	0.7496
$RMSE(\hat{\beta}_{re}; \hat{\beta}_{wmre})$	1.024	1.063	1.0977	1.0125	1.048	1.085
$RMSE(\hat{\beta}_{wme}; \hat{\beta}_{wmre})$	1.45	1.43	1.44	1.39	1.38	1.38
$d_1'(D_2 + d_2 d_2')d_1$	0.0119	0.0115	0.0108	0.00068	0.00073	0.00073
$d_1' D_1^{-1} d_1$	0.1788	0.1793	0.1775	0.2615	0.2612	0.2624

Table 3.5: The estimated MSE values and corresponding RMSE values of $\hat{\beta}$, $\hat{\beta}_{re}$, $\hat{\beta}_{wme}$ and $\hat{\beta}_{wmre}$ at $\rho = 0.9$ with $\Lambda = diag(0.001, 0.001, 0.001)$.

	$\Lambda = diag(0.007, 0.007, 0.007)$					
$(\sigma^2, \sigma_1^2, \sigma_2^2)$	$(1, 0.1, 0.1)$					
	$n_1 = 30$			$n_1 = 70$		
	w=0.1	w=0.5	w=1	w=0.1	w=0.5	w=1
$\hat{\beta}$	0.5357	0.5357	0.5357	0.2111	0.2111	0.2111
$\hat{\beta}_{re}$	0.4363	0.4316	0.4245	0.1861	0.1858	0.1852
$\hat{\beta}_{wme}$	0.5159	0.4733	0.4365	0.2072	0.1962	0.1861
$\hat{\beta}_{wmre}$	0.42	0.3930	0.3640	0.1830	0.1741	0.1654
$RMSE(\hat{\beta}_{re}; \hat{\beta}_{wmre})$	1.031	1.098	1.166	1.016	1.067	1.119
$RMSE(\hat{\beta}_{wme}; \hat{\beta}_{wmre})$	1.21	1.20	1.19	1.13	1.12	1.12
$d_1' (D_2 + d_2 d_2') d_1$	0.00033	0.00042	0.00044	0.00015	0.0002	0.00022
$d_1' D_1^{-1} d_1$	0.1364	0.1367	0.1385	0.1472	0.1465	0.1470
$(\sigma^2, \sigma_1^2, \sigma_2^2)$	$(6, 0.6, 0.6)$					
$\hat{\beta}$	3.2138	3.2138	3.2138	1.2670	1.2670	1.2670
$\hat{\beta}_{re}$	2.1994	2.1265	2.0194	0.9053	0.8952	0.8775
$\hat{\beta}_{wme}$	3.0946	2.8384	2.6176	1.2431	1.1771	1.1163
$\hat{\beta}_{wmre}$	2.1429	1.9899	1.8259	0.8935	0.8513	0.8049
$RMSE(\hat{\beta}_{re}; \hat{\beta}_{wmre})$	1.026	1.068	1.105	1.013	1.051	1.09
$RMSE(\hat{\beta}_{wme}; \hat{\beta}_{wmre})$	1.44	1.42	1.43	1.39	1.38	1.38
$d_1' (D_2 + d_2 d_2') d_1$	0.0123	0.0119	0.0113	0.00072	0.00078	0.00078
$d_1' D_1^{-1} d_1$	0.1718	0.1721	0.1701	0.2514	0.2512	0.2525

Table 3.6: The estimated MSE values and corresponding RMSE values of $\hat{\beta}$, $\hat{\beta}_{re}$, $\hat{\beta}_{wme}$ and $\hat{\beta}_{wmre}$ at $\rho = 0.9$ with $\Lambda = diag(0.007, 0.007, 0.007)$.

	$\Lambda = diag(0.01, 0.01, 0.01)$					
$(\sigma^2, \sigma_1^2, \sigma_2^2)$	$(1, 0.1, 0.1)$					
	$n_1 = 30$			$n_1 = 70$		
	w=0.1	w=0.5	w=1	w=0.1	w=0.5	w=1
$\hat{\beta}$	0.5602	0.5602	0.5602	0.2195	0.2195	0.2195
$\hat{\beta}_{re}$	0.4563	0.4511	0.4435	0.1933	0.1930	0.1923
$\hat{\beta}_{wme}$	0.5389	0.4933	0.4542	0.2153	0.2036	0.1929
$\hat{\beta}_{wmre}$	0.4421	0.4098	0.3788	0.1900	0.1805	0.1714
$RMSE(\hat{\beta}_{re}; \hat{\beta}_{wmre})$	1.03	1.10	1.17	1.01	1.06	1.12
$RMSE(\hat{\beta}_{wme}; \hat{\beta}_{wmre})$	1.21	1.20	1.19	1.13	1.12	1.12
$d_1' (D_2 + d_2 d_2') d_1$	0.0036	0.0094	0.0045	0.0015	0.0021	0.0042
$d_1' D_1^{-1} d_1$	0.1316	0.1024	0.2672	0.1443	0.1436	0.1441
$(\sigma^2, \sigma_1^2, \sigma_2^2)$	$(6, 0.6, 0.6)$					
$\hat{\beta}$	3.3602	3.3602	3.3602	1.3170	1.3170	1.3170
$\hat{\beta}_{re}$	2.3082	2.2298	2.1155	0.9423	0.9314	0.9127
$\hat{\beta}_{wme}$	3.2326	2.9576	2.7240	1.2916	1.2213	1.1571
$\hat{\beta}_{wmre}$	2.2469	2.0807	1.9060	0.9296	0.8843	0.8352
$RMSE(\hat{\beta}_{re}; \hat{\beta}_{wmre})$	1.02	1.07	1.10	1.01	1.05	1.09
$RMSE(\hat{\beta}_{wme}; \hat{\beta}_{wmre})$	1.43	1.42	1.42	1.38	1.38	1.38
$d_1' (D_2 + d_2 d_2') d_1$	0.0124	0.0122	0.0153	0.0007	0.0008	0.0009
$d_1' D_1^{-1} d_1$	0.1684	0.1685	1.1516	0.2464	0.2462	0.2476

Table 3.7: The estimated MSE values and corresponding RMSE values of $\hat{\beta}$, $\hat{\beta}_{re}$, $\hat{\beta}_{wme}$ and $\hat{\beta}_{wmre}$ at $\rho = 0.9$ with $\Lambda = diag(0.01, 0.01, 0.01)$.

	$\Lambda = diag(0.01, 0.01, 0.01)$					
$(\sigma^2, \sigma_1^2, \sigma_2^2)$	(1,0.1,0.1)					
	$n_1 = 30$			$n_1 = 70$		
	w=0.1	w=0.5	w=1	w=0.1	w=0.5	w=1
$\hat{\beta}$	0.9578	0.9578	0.9578	0.3792	0.3792	0.3792
$\hat{\beta}_{re}$	0.7171	0.6953	0.6659	0.3099	0.3078	0.3044
$\hat{\beta}_{wme}$	0.9025	0.7870	0.6969	0.3671	0.3356	0.3097
$\hat{\beta}_{wmre}$	0.6867	0.6118	0.5447	0.3017	0.2778	0.2569
$RMSE(\hat{\beta}_{re}; \hat{\beta}_{wmre})$	1.044	1.136	1.222	1.027	1.107	1.185
$RMSE(\hat{\beta}_{wme}; \hat{\beta}_{wmre})$	1.31	1.28	1.27	1.21	1.20	1.20
$d_1'(D_2 + d_2d_2')d_1$	0.00056	0.00056	0.0007	1.4513	0.00023	0.00028
$d_1'D_1^{-1}d_1$	0.1594	0.1584	0.1311	0.1444	0.1459	0.1451
$(\sigma^2, \sigma_1^2, \sigma_2^2)$	(6,0.6,0.6)					
$\hat{\beta}$	5.7480	5.7480	5.7480	2.2757	2.2757	2.2757
$\hat{\beta}_{re}$	3.6988	3.4589	3.1320	1.4892	1.4488	1.3825
$\hat{\beta}_{wme}$	5.4154	4.7218	4.1848	2.2030	2.0138	1.8583
$\hat{\beta}_{wmre}$	3.5728	3.1765	2.7853	1.4599	1.3425	1.2251
$RMSE(\hat{\beta}_{re}; \hat{\beta}_{wmre})$	1.035	1.088	1.124	1.02	1.079	1.128
$RMSE(\hat{\beta}_{wme}; \hat{\beta}_{wmre})$	1.51	1.48	1.50	1.50	1.49	1.49
$d_1'(D_2 + d_2d_2')d_1$	0.0502	0.0466	0.0415	0.0037	0.004	0.0039
$d_1'D_1^{-1}d_1$	0.1476	0.1410	0.1336	0.2093	0.2112	0.2122

Table 3.8: The estimated MSE values and corresponding RMSE values of $\hat{\beta}$, $\hat{\beta}_{re}$, $\hat{\beta}_{wme}$ and $\hat{\beta}_{wmre}$ at $\rho = 0.95$ with $\Lambda = diag(0.001, 0.001, 0.001)$.

	$\Lambda = diag(0.007, 0.007, 0.007)$					
$(\sigma^2, \sigma_1^2, \sigma_2^2)$	$(1, 0.1, 0.1)$					
	$n_1 = 30$			$n_1 = 70$		
	w=0.1	w=0.5	w=1	w=0.1	w=0.5	w=1
$\hat{\beta}$	1.1419	1.1419	1.1419	0.4417	0.4417	0.4417
$\hat{\beta}_{re}$	0.8579	0.8286	0.7901	0.3600	0.3572	0.3528
$\hat{\beta}_{wme}$	1.0685	0.9203	0.8077	0.4261	0.3869	0.3556
$\hat{\beta}_{wmre}$	0.8176	0.7180	0.6323	0.3495	0.3195	0.2939
$RMSE(\hat{\beta}_{re}; \hat{\beta}_{wmre})$	1.049	1.153	1.249	1.029	1.117	1.2
$RMSE(\hat{\beta}_{wme}; \hat{\beta}_{wmre})$	1.30	1.28	1.27	1.21	1.21	1.20
$d_1' (D_2 + d_2 d_2') d_1$	0.00053	0.00101	0.000624	0.000018	0.00027	0.0003
$d_1' D_1^{-1} d_1$	0.2934	0.1168	0.1345	0.1302	0.1315	0.1309
$(\sigma^2, \sigma_1^2, \sigma_2^2)$	$(6, 0.6, 0.6)$					
$\hat{\beta}$	6.8503	6.8503	6.8503	2.6502	2.6502	2.6502
$\hat{\beta}_{re}$	4.4976	4.1869	3.7829	1.7560	1.7043	1.6237
$\hat{\beta}_{wme}$	6.4095	5.5136	4.8404	2.5569	2.3208	2.1328
$\hat{\beta}_{wmre}$	4.3204	3.7789	3.2783	1.7162	1.5641	1.4191
$RMSE(\hat{\beta}_{re}; \hat{\beta}_{wmre})$	1.041	1.107	1.153	1.023	1.089	1.144
$RMSE(\hat{\beta}_{wme}; \hat{\beta}_{wmre})$	1.48	1.45	1.47	1.48	1.48	1.50
$d_1' (D_2 + d_2 d_2') d_1$	0.0538	0.0518	0.0472	0.0041	0.0046	0.0045
$d_1' D_1^{-1} d_1$	0.1402	0.1333	0.1257	0.1914	0.1935	0.1949

Table 3.9: The estimated MSE values and corresponding RMSE values of $\hat{\beta}$, $\hat{\beta}_{re}$, $\hat{\beta}_{wme}$ and $\hat{\beta}_{wmre}$ at $\rho = 0.95$ with $\Lambda = diag(0.007, 0.007, 0.007)$.

	$\Lambda = diag(0.01, 0.01, 0.01)$					
$(\sigma^2, \sigma_1^2, \sigma_2^2)$	$(1, 0.1, 0.1)$					
	$n_1 = 30$			$n_1 = 70$		
	w=0.1	w=0.5	w=1	w=0.1	w=0.5	w=1
$\hat{\beta}$	1.2567	1.2567	1.2567	0.4790	0.4790	0.4790
$\hat{\beta}_{re}$	0.9467	0.9124	0.8686	0.3901	0.3867	0.3818
$\hat{\beta}_{wme}$	1.1719	1.0028	0.8777	0.4613	0.4173	0.3830
$\hat{\beta}_{wmre}$	0.8998	0.7841	0.6887	0.3781	0.3443	0.3160
$RMSE(\hat{\beta}_{re}; \hat{\beta}_{wmre})$	1.05	1.16	1.26	1.03	1.12	1.20
$RMSE(\hat{\beta}_{wme}; \hat{\beta}_{wmre})$	1.30	1.27	1.27	1.22	1.21	1.21
$d_1' (D_2 + d_2 d_2') d_1$	0.1281	0.2234	0.1261	0.1239	0.1252	0.1246
$d_1' D_1^{-1} d_1$	0.0006	0.0064	0.0069	0.0062	0.0031	0.0035
$(\sigma^2, \sigma_1^2, \sigma_2^2)$	$(6, 0.6, 0.6)$					
$\hat{\beta}$	7.5380	7.5380	7.5380	2.8741	2.8741	2.8741
$\hat{\beta}_{re}$	5.0045	4.6552	4.2031	1.9176	1.8588	1.7702
$\hat{\beta}_{wme}$	7.0270	6.0081	5.2589	2.7676	2.5027	2.2974
$\hat{\beta}_{wmre}$	4.7928	4.1561	3.5939	1.8710	1.6973	1.5365
$RMSE(\hat{\beta}_{re}; \hat{\beta}_{wmre})$	1.04	1.11	1.16	1.02	1.09	1.15
$RMSE(\hat{\beta}_{wme}; \hat{\beta}_{wmre})$	1.46	1.44	1.46	1.47	1.47	1.49
$d_1' (D_2 + d_2 d_2') d_1$	0.05518	0.0552	0.05019	0.0042	0.0048	0.0048
$d_1' D_1^{-1} d_1$	0.1368	0.1305	0.1223	0.1830	0.1852	0.1867

Table 3.10: The estimated MSE values and corresponding RMSE values of $\hat{\beta}$, $\hat{\beta}_{re}$, $\hat{\beta}_{wme}$ and $\hat{\beta}_{wmre}$ at $\rho = 0.95$ with $\Lambda = diag(0.001, 0.001, 0.001)$.

Chapter 4

Conclusions and areas for further research

Linear mixed models (LMMs), (see [34] and [41]) are an extension to the linear models in which the models contain random effects in addition to the usual fixed effects. These models provide a flexible and powerful tool for the analysis of grouped data, which arise in many areas as diverse as agriculture, biology, economics, manufacturing, and geophysics. Examples of grouped data include longitudinal data, repeated measures, blocked designs, and multilevel data that Longitudinal data are very common in practice, either in observational studies or in experimental studies. In a longitudinal study, individuals in the study are followed over a period of time and for each individual, data are collected at multiple time points. LMMs have been widely used to model longitudinal and repeated measurements data during the last two decades, and have received much attention in the fields of medical and biological sciences.

In practice it often happens that some collected data are subject to measurement error. Sometimes covariates of interest may be difficult to observe precisely due to physical location or cost. Sometimes it is impossible to measure covariates accurately due to their nature. In other situations, a covariate may represent an average of a certain quantity over time, and any practical way of measuring such a quantity necessarily features measurement error. When carrying out statistical inference in such settings, it is important to account for the effects of mismeasured covariates; otherwise, erroneous or even misleading results may be produced. Measurement error may degrade the quality of inference and should be avoided whenever possible. In order to correct the effect of measurement error on parameters estimation, [45] suggested an approach based on the corrected score function.

The standard assumption in the linear regression analysis is that all the regressor variables are linearly independent. When this assumption is violated, the problem of multicollinearity enters into the data and it inflates the variance of an ordinary least squares estimator of the regression coefficient. With multicollinear data, some coefficients may be statistically insignificant and may have the wrong signs, see [36] for more details. Obtaining the estimators for multicollinear data is an important problem in the literature. To overcome this multicollinearity problem, different remedial methods have been proposed. One estimation technique designed to combat multicollinearity is using biased estimators, most notable of which are the Stein estimator by [56], the principal components regression (PCR) estimator by [44], the ordinary ridge regression (ORR) estimator by [29] and the Liu estimator by [39]. Another method to combat multicollinearity is through the collection and use of additional information, which can be either exact or stochastic restrictions [48]. However, exact restrictions are often discomforting in many applied work such as economic relations, industrial structures, production planning, and so on. While, as pointed out by [2] using stochastic linear restriction, one can accomplish an examination and analysis of one's own thoughts and feelings (prior information via introspection). In addition, one may also have prior information from a previous sample which usually makes some relations through stochastic subspace restrictions. when it comes to stochastic linear restrictions, [11], [58] and [57] proposed the ordinary mixed estimator (OME) by combining the sample model with stochastic restrictions. Some other important references on this subject are [35], [36], [61], [63], and so on.

As the third solution, the combination of ridge estimator and stochastic linear restrictions may inherit the advantages of both estimators. [22], [35], [71], [36], [35] and [62] considered stochastic ridge estimation in the linear regression models. [20], [19], and [18] studied the ridge estimator with stochastic linear restrictions in linear models with measurement error. [65] concentrated on the ridge estimates of fixed and random effects using Nakamura's approach for LMMs with the measurement error, then derived the stochastic restricted ridge estimates of fixed and random effects, when prior information is available in the form of the stochastic linear restrictions on fixed and random effects. Also, they assumed that the variance parameters are not known and got the estimation of variance parameters. When the prior information and the sample information are not equally important in practice, [52] introduced the method of weighted mixed regression and proposed the weighted mixed estimator (WME). Motivated by this, In chapter 2, we proposed a new ridge type estimator called the new mixed ridge estimator (NMRE) for the vector of parameters in LMMs with measurement error. Also, we investigated the performance of the NMRE using the MSEM criterion. Furthermore, We show that the new estimator is superior to the other estimator in the MSEM sense under certain conditions. Finally, we illustrated our findings with a data example and simulation study.

In chapter 3, we introduced the new ridge estimator called the weighted mixed ridge estimator (WMRE), which is a generalization of the WME and RE estimator, and shows that this estimator is superior to the WME and RE using MSEM, under certain conditions.

Now, we have some suggestions for future research.

- Obtaining the parameter's estimation in LMMs when both fixed and random effects are subject to measurement error.
- Using another biased estimator instead of a ridge estimator such as a Liu estimator
- Selecting another distribution other than the normal distribution for random error vector and random effects vector in LMMs with measurement error
- One of the assumptions in this thesis is that matrix Λ is known and matrix Σ is diagonal, ignoring these restrictions can be the conditions for the future research.

Appendix A

A.1 Corrected score method to variance component estimation

The corrected log-likelihood to the estimation of σ^2 and $\gamma_{1i}{}'s,\ i=1,...,c_1$, is

$$\ell_1^*\left(\sigma^2,\gamma_1;X,y\right)=\tfrac{-n_1}{2}\log\left(2\pi\sigma^2\right)-\tfrac{1}{2}\log|V_1|$$
$$-\tfrac{1}{2\sigma^2}\left[\left(y-X\hat{\beta}\right)'V_1^{-1}\left(y-X\hat{\beta}\right)-tr\left(V_1^{-1}\right)\hat{\beta}'\Lambda\hat{\beta}\right],$$

A corrected score estimator is a vector value of $\sigma^2,\ \gamma_{1i}\,(i=1,...,c_1)$, that maximizes ℓ_1^*, this is derived as an explicit solution to the following equations:

$$\partial\ell_1^*\left(\sigma^2,\gamma_1;X,y\right)/\partial\sigma^2=0, \tag{A.1}$$

$$\partial\ell_1^*\left(\sigma^2,\gamma_1;X,y\right)/\partial\gamma_{1i}=0,\ i=1,...,c_1. \tag{A.2}$$

From equation A.1, the CSE of σ^2 is given by:

$$\hat{\sigma}^2=\frac{1}{n_1}\left[\left(y-X\hat{\beta}\right)'V_1^{-1}\left(y-X\hat{\beta}\right)-tr\left(V_1^{-1}\right)\hat{\beta}'\Lambda\hat{\beta}\right],$$

In order to solve equation A.2, we note that $|V_1|=|I_{n_1}+U'U\Sigma_1|$, where $V_1=I_{n_1}+U\Sigma_1U'=I_{n_1}+\sum_{i=1}^{c_1}U_iU'_i$. Now using the following relations, $\partial V/\partial\gamma_i=U_iU_i'$, $\partial V^{-1}/\partial\gamma_i=-V^{-1}U_iU_i'V^{-1}$ and $\partial\sum_1/\partial\gamma_{1i}=diag\left(0,...,0,I_{q_{1i}},0,...,0\right)$. Then by solving the above equation, the CSE of $\hat{\sigma}_{11}^2,...,\hat{\sigma}_{1c_1}^2$ are given by:

$$\hat{\sigma}_{1i}^2=\frac{\left[\hat{b}_{1i}\hat{b}_{1i}-tr\left(\hat{D}_i'\hat{D}_i\right)\hat{\beta}'\Lambda\hat{\beta}\right]}{\left[q_{1i}-tr\left(T_{ii}\right)\right]},$$

An iterative algorithm is needed to compute the corrected score estimate of σ^2, $\gamma_i\,(i=1,...,c_1)$, among different algorithms. The following one which is regarded as an extension of [25] and [53], is selected to deal with the extra measurement errors:

Step 0: set $m=0$ and choose starting values $\sigma^{2(0)}$ and $\hat{\sigma}_i^{2(0)}$, $i=1,...,c_1$

Step 1: Calculate estimates $\hat{\beta}^{(m)}$ and $\hat{b}_1^{(m)},...,\hat{b}_{c_1}^{(m)}$.

Step 2: Calculate

$$\hat{\sigma}^{2(m+1)}=\frac{1}{n_1}\left[\left(y-X\hat{\beta}^{(m)}\right)'V_1^{-1(m)}\left(y-X\hat{\beta}^{(m)}\right)-tr\left(V_1^{-1(m)}\right)\hat{\beta}^{(m)\prime}\Lambda\hat{\beta}^{(m)}\right],$$

and

$$\hat{\sigma}_c^{2(m+1)}=\frac{1}{q_{1i}-tr\left(T_{ii}\right)}\left[\hat{b}_i^{(m)\prime}\hat{b}_i^{(m)}-tr\left(\hat{D}_i^{(m)\prime}\hat{D}_i^{(m)}\right)\hat{\beta}^{(m)\prime}\Lambda\hat{\beta}^{(m)}\right].$$

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

A.2 Several Lemmas

Lemma 1. *Assume that square matrices A and C are not singular and B and D are matrices with proper orders, then* $(A+BCD)^{-1} = A^{-1} - A^{-1}B\left(C^{-1} + DA^{-1}B\right)DA^{-1}$, *(see [48]).*

Lemma 2. *Let M be an* $n \times n$ *p.d matrix, namely* $M > 0$, *and* α *be a nonzero* $n \times 1$ *vector , then* $M - \alpha\alpha' \geqslant 0$, *If* $\alpha' M^{-1}\alpha \leqslant 1$, *(see [13]).*

Lemma 3. *Let* $\hat{\beta}_j = A_j y,\ j = 1,2$ *be two competing linear estimator of* β. *Suppose that* $D = Var\left(\hat{\beta}_1\right) - Var\left(\hat{\beta}_2\right) > 0$ *, then* $\Delta = MSE\left(\hat{\beta}_1\right) - MSE\left(\hat{\beta}_2\right) \geq 0$, *if* $d_2'\left(D + d_1' d_1\right) d_2 \leq 1$, *where* $MSE\left(\hat{\beta}_j\right)$, d_j *denote the mean square error matrix and bias vector of* $\hat{\beta}_j$, *respectively, (see [59]).*

Lemma 4. *Let* $n \times n$ *matrices* $M > 0$, $N \geq 0$ *, then* $M > N$, *if and only if* $\lambda_1 < 1$ *, where* λ *is the largest eigenvalue of the matrix* NM^{-1}, *(see [48]).*

Bibliography

[1] T.W Anderson. Estimating linear statistical relationships. *The Annals of Statistics*, 87:1–45, 1984.

[2] M. Arashi and S. M. M. Tabatabaey. Stein-type improvement under stochastic constraints: Use of multivariate student-t model in regression. *Statistics and probability letters*, 78:2142–2153, 2008.

[3] B. Armstrong. Measurement error in the generalized linear model. *Comm. Statist. Simulation Comput.*, 14:529–544, 1985.

[4] J. P. Buonaccorsi. *Measurement error: models, methods, and applications*. CRC press, 2010.

[5] R.J. Carrol, D. Ruppert, L.A Stefanski, and C.M Crainiceanu. *Measurement error in nonlinear models: a modern perspective*. CRC press, 2006.

[6] R.J Carroll and L.A. Stefanski. Asymptotic theory for the simex estimator in measurement error models. *In Advances in Statistical Decision Theory and Applications*, 87:151–164, 1997.

[7] H. Cui, K. W. Ng, and L. Zhu. Estimation in mixed effects model with errors in variables. *Journal of multivariate analysis.*, 91:53–73, 2004.

[8] M. Davidian and D. M. Giltinan. *Nonlinear models for repeated measurements data*. CRC press, London, 1995.

[9] P. Diggle, P. Heagerty, and and S. Zeger K. Y. Liang. *Analysis of longitudinal data*. Oxford University Press, 2002.

[10] N.R Draperand and H. Smith. *Applied regression analysis*. John Wiley and Sons., 1998.

[11] J. Durbin. A note on regression when there is extraneous information about one of the coefficients. *Journal of the American Statistical Association*, 48:799–808, 1953.

[12] M. Eliot, J. Ferguson, and M. P. Foulkes. Ridge regression for longitudinal biomarker data. *The international journal of biostatistic.*, 7:1–11, 2011.

[13] R. W. Farebrother. Further results on the mean square error of ridge regression. *Journal of the Royal Statistical Society*, 38:248–250, 1976.

[14] G. Fitzmaurice, M. Davidian, G. Verbeke, and G. Molenberghs. *Longitudinal data analysis*. CRC press, 2008.

[15] W. A. Fuller. *Measurement error models*. Wiley, New York, 1987.

[16] W. K. Fung, X. P. Zhong, and B. C. Wei. On estimation and influence diagnostics in linear mixed measurement error models. *Journal of Mathematical and Management Sciences,*, 23:37–59, 2003.

[17] F. Ghapani. Stochastic restricted liu estimator in linear mixed measurement error models. *Communications in Statistics-Simulation and Computation*, pages 1–14, 2019.

[18] F. Ghapani and B. Babadi. A new ridge estimator in linear measurement error model with stochastic linear restrictions. *Journal of The Iranian Statistical Society*, 15:87–103, 2016.

[19] F. Ghapani, A. R. Rasekh, , and B. Babadi. The weighted ridge estimator in stochastic restricted linear measurement error models. *Statistical Papers*, 59:709–723, 2018.

[20] F. Ghapani, A. R. Rasekh, M. R. Akhoond, and B. Babadi. Detection of outliers and influential observations in linear ridge measurement error models with stochastic linear restrictions. *Journal of Sciences, Islamic Republic of Iran.*, 26:355–366, 2015.

[21] A. Gilmour, B.Cullis, S.Welham, and B.Gogel. An efficient computing strategy for prediction in mixed linear models. *Computational statistics and data analysis*, 44:571–586, 2004.

[22] J. Groß. Restricted ridge estimation. *Statistics and probability letters*, 65:57–64, 2003.

[23] F. N. Gumedze and T. T. Dunne. Parameter estimation and inference in the linear mixed model. *Linear Algebra and its Applications.*, 435:1920–1944, 2011.

[24] Jr. D. Harrison and D. L. Rubinfeld. Hedonic housing prices and the demand for clean air. *Journal of environmental economics and management*, 5:81–102, 1987.

[25] D. A. Harvill. Maximum likelihood approaches to variance component estimation and related problems (with Discussion. *J. Amer. Statist. Assoc.*, 72:320–340, 1977.

[26] D. He and Y. Wu. A stochastic restricted principal components regression estimator in the linear model. *The Scientific World Journal*, 2014.

[27] D. Hedeker and R. D. Gibbons. *Longitudinal data analysis*. John Wiley and Sons, 2006.

[28] C. R. Henderson, O. Kempthorne, S. R. Searle, and C. M. Von Krosigk. The estimation of environmental and genetic trends from records subject to culling. *Biometrics*, 2:192–218, 1959.

[29] A. E. Hoerl and R. W. Kennard. Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics*, 12:55–67, 1970.

[30] M. H. Hubert and P. Wijekoon. Improvement of the liu estimator in linear regression model. *Statistical Papers*, 47:471, 2006.

[31] B. G. Kibria. Performance of some new ridge regression estimators. *Communications in Statistics-Simulation and Computation*, 32:419–435, 2003.

[32] Ö. kuran and M. R. Özkale. Gilmour's approach to mixed and stochastic restricted ridge predictions in linear mixed models. *Linear Algebra and its Applications*, 508:22–47, 2016.

[33] Ö. kuran and M. R. Özkale. Improvement of mixed predictors in linear mixed models. *Journal of Applied Statistics*, 12:1–19, 2020.

[34] N. M. Laird and J. H. Ware. Random-effects models for longitudinal data. *Biometrics*, pages 963–974, 1982.

[35] Y. Li and H. Yang. A new stochastic mixed ridge estimator in linear regression model. *Statistical Papers*, 51:315–323, 2010.

[36] Y. Li and H. Yang. A new ridge-type estimator in stochastic restricted linear regression. *Statistics*, 42:123–130, 2011.

[37] Z. Li. Estimation in linear mixed models for longitudinal data under linear restricted conditions. *Journal of statistical planning and inference*, 2:869–876, 2011.

[38] M. J. Lindstrom and D. M. Bates. Newton—raphson and em algorithms for linear mixed-effects models for repeated-measures data. *Journal of the American Statistical Association*, 83:1014–1022, 1988.

[39] K. Liu. A new class of biased estimate in linear regression. Communications in Statistics. *Theory and Methods.*, 22:393–402, 1993.

[40] X. Q. Liu and P. Hu. General ridge predictors in a mixed linear model. *Statistics*, 47:363–378, 2013.

[41] N. T. Longford. *Random Coefficient Models*. Oxford University Press, Oxford, 1993.

[42] C. E. MaCulloch and S. R. Searle. *Generalized, linear, and mixed models*. Jon Wiley and Sons, New York, 2001.

[43] D. W. Marquardt and R. D. Snee. Ridge regression in practice. *The American Statistician*, 29:3–20, 1975.

[44] W. F. Massy. Principal components regression in exploratory statistical research. *Journal of the American Statistical Association*, 60:234–256, 1965.

[45] T. Nakamura. Corrected score function for errors-in-variables models. *Methodology and application to generalized linear models*, 77:127–137, 1990.

[46] L. N. Pereira and P. S. Coelho. A small area predictor under area-level linear mixed models with restrictions. *Communications in Statistics-Theory and Methods*, 41:2524–2544, 2012.

[47] D.A Pierce, D.O Stram, M. Vaeth, and D. Schafer. Some insights into the errors in variables problem provided by consideration of radiation dose-response analyses for the a-bomb survivors. *Journal of the American Statistical Association,*, 87:351–359, 1992.

[48] C. R. Rao, H. Toutenburg, and H. C. Shalabh. *Linear models and generalizations. Least Squares and Alternatives*. Springer: Berlin Heidelberg, New York, 2008.

[49] M. Riquelme and M. Galea H. Bolfarine. Robust linear functional mixed models. *Journal of Multivariate Analysis*, 132:82–98, 2015.

[50] B. Rosner, D. Spiegelman, and W.C. Willett. Correction of logistic regression relative risk estimates and confidence intervals for measurement error: the case of multiple covariates measured with error. *American journal of epidemiology*, 132:734–745, 1990.

[51] S.S Sandra. *Collinearity in mixed models*. Ph.D. Thesis, Department of Biostatistics, University of North Carolina, 1993.

[52] B. Schaffrin and H. Toutenburg. Weighted mixed regression. *Zeitschrift fur Angewandte Mathematik und Mechanik*, 78:T735–T738, 1990.

[53] R. Schall. Estimation in generalized linear models with random effects. *Biometrika*, 78:719–727, 1991.

[54] S. R. Searle, G. Casella, and C. E. MaCulloch. *Variance components*. Wiley, New York, 2006.

[55] L. A. Stefanski and R. J. Carroll. Conditional scores and optimal scores for generalized linear measurement error models. *Biometrika*, 74:703–716, 1987.

[56] C. Stein. *Inadmissibility of the usual estimator for the mean of a multivariate normal distribution=*. Stanford University Stanford United States, 1956.

[57] H. Theil. On the use of incomplete prior information in regression analysis. *Journal of the American Statistical Association*, 58:401–414, 1963.

[58] H. Theil and A.S. Goldberger. On pure and mixed statistical estimation in economics. *International Economic Review*, 2:65–78, 1961.

[59] G. Trenkler and H. Toutenburg. Mean squared error matrix comparisons between biased estimators—an overview of recent results. *Statistical Papers*, 31:165–179, 1990.

[60] L. Wu. *Mixed effect model for complex data*. Chapman and Hall, New York, 2010.

[61] H. Yang and J. Cui. A stochastic restricted two-parameter estimator in linear regression model. *Communications in Statistics-Theory and Methods*, 40:2318–2325, 2011.

[62] H. Yang and K. Xue H. Ye. A further study of predictions in linear mixed models. *Communications in Statistics-Theory and Methods*, 43:4241–4252, 2014.

[63] H. Yang and J. Xu. An alternative stochastic restricted liu estimator in linear regression. *Statistical Papers*, 50:639–647, 2009.

[64] B. Yavarizadeh and S.E. Ahmed. The new mixed ridge estimator in a linear mixed model with measurement error under stochastic linear mixed restrictions. *Communications in Statistics-Simulation and Computation*, pages 1–17, 2021.

[65] B. Yavarizadeh, A. R. Rasekh, S. E. Ahmed, and B. Babadi. Ridge estimation in linear mixed measurement error models with stochastic linear mixed restrictions. *Communications in Statistics-Simulation and Computation*, pages 1–17, 2019.

[66] B. Yavarizadeh, A. R. Rasekh, and B. Babadi. Estimation of parameters in linear mixed measurement error models with stochastic linear restrictions. *Communications in Statistics-Theory and Methods*, 23:5853–5865, 2020.

[67] N. Yildiz. On the weighted mixed Liu-type estimator under unbiased stochastic restrictions. *Communications in Statistics-Simulation and Computation.*, 9:7238–7248, 2017.

[68] K. Zare, A.R. Rasekh, and A. Rasekhi. Estimation of variance components in linear mixed measurement error models. *Statistical Papers*, 53:849–863, 2012.

[69] X. P. Zhong, W. K. Fung, and B. C. Wei. Estimation in linear models with random effects and errors in variables. *Ann. Inst. Math. Statist.*, 54:595–606, 2002.

[70] N. Özbay and S. Kaçıranlar. Estimation in a linear regression model with stochastic linear restrictions: a new two-parameter-weighted mixed estimator. *Journal of Statistical Computation and Simulation*, 88:1669–1683, 2018.

[71] M. R. Özkale. A stochastic restricted ridge regression estimator. *Journal of Multivariate Analysis*, 100:1706–1716, 2009.

[72] M. R. Özkale. The relative efficiency of the restricted estimators in linear regression models. *Journal of Applied Statistics.*, 41:998–1027, 2014.

[73] M. R. Özkale and F. Can. An evaluation of ridge estimator in linear mixed models: an example from kidney failure data. *Journal of Applied Statistics*, 44:2251–2269, 2017.

[74] M. R. Özkale and Ö. Kuran. A further prediction method in linear mixed models: Liu prediction. *Communications in Statistics-Simulation and Computation*, pages 1–25, 2019.