Journal of Sciences, Islamic Republic of Iran 34(4): 333 - 348 (2023) http://jsciences.ut.ac.ir
University of Tehran, ISSN 1016-1104

Maximum Approximated Likelihood Estimation in Generalized Linear Multilevel Model for Nominal Response with Covariates Subject to Measurement Error

M. Ahangari[1], M. Golalizadeh[1*], Z. Rezaei Ghahroodi[2]

[1] *Department of Statistics, Tarbiat Modares University, Tehran, Islamic Republic of Iran*
[2] *School of Mathematics, Statistics and Computer Science, University of Tehran, Tehran, Islamic Republic of Iran*

Received: 28 June 2023 / Revised: 30 December 2023 / Accepted: 22 April 2024

Abstract

Ignoring measurement errors in generalized linear mixed models (GLMMs) as well as other regression models will inevitably lead to significant biases, deviations, and incorrect inferences in the estimation of the parameters. In the presence of measurement error, numerous approaches have been proposed to rectify this issue. Furthermore, the application of the frequentist approach of GLMMs is intricate due to the emergence of an intractable numerical integration process. In this study, the complete likelihood inference of a multinomial logit random effects model with covariate measurement error in conjunction with replicate measures is suggested via the multivariate Gauss-Hermite quadrature approximation. To achieve this objective, the likelihood function will be evaluated and approximated for two distinct scenarios; the proposed method which incorporates the classical additive structural measurement error model for the error-prone covariate and the naive method. We will describe and compare different upshots of parameter estimation in these two different situations. The results of performing the proposed method, assessed through simulation, show that the proposed method performs well when correcting for measurement error in terms of bias, empirical standard error, root of mean squared error and coverage ratio. The application of the proposed method is further highlighted with real-world data based on a multilevel study concerning the prevalence of contraceptive methods used by women in Bangladesh.

Keywords: Mixed Models; Nominal Response; Multivariate Gauss-Hermite Quadrature; Multinomial Logit Model; Measurement Error.

Introduction

In numerous statistical scenarios, it is commonly assumed that the observations are independent of one another. However, there exist certain instances that violate this assumption. A prime example of this is evident in the realm of biological sciences, where the data collected often possesses a hierarchical or clustered structure. The presence of correlation among such data is not coincidental or negligible (1). T1o effectively analyze

* Corresponding Author: Tel: +98 21-82884705; Email: golalizadeh@modares.ac.ir

such correlated and longitudinally measured data, researchers employ mixed effects or multilevel (hierarchical) models, which also take into account the incorporation of inter-class correlation.

The purpose of mixed effects models is to model the relationship between the response variable and a set of covariates by considering the correlation between repeated measurements or clustered structure. For the continuous outcomes, linear mixed models (LMMs) are used (2). In the case of non-normally distributed responses, generalized linear mixed models (GLMMs) are widely applied, specifying a nonlinear link between the mean of the response and the predictors, and can model the correlation between responses by adding random effects (3). The maximum likelihood (ML) approach is most frequently used in estimating the parameters in LMMs, but its usage is limited in GLMMs, mainly because the conditional distribution of the response components is not normal, so the likelihood function may not have a closed form. To avoid computational problems of the ML method, a Monte Carlo EM (MCEM) approach has been described for binary responses which can handle both fixed and random effects structure (4). Hereof, a Bayesian method with flat priors has been applied to approximate ML estimates (see, e.g. (5) and (6)). It is noteworthy to emphasize that in instances where generalized linear latent and mixed models are concerned, (7) have devised innovative approaches for the purpose of model identification, estimation, prediction of latent variables, as well as model diagnostics. It is also important to note that other numerical algorithms such as Laplace approximation or Gaussian quadrature are also used to overcome the integrals in the likelihood function. One can consult (8) and (9), for more details.

One of the most challenging problems in regression analysis is the situation where some of the covariates are measured with error (10). There are few studies about the effects of covariate measurement error in estimating the parameters and analyzing the correlated data in the frequentist approach, only some approximation methods have been proposed, such as regression calibration (11), simulation extrapolation (12), corrected scores, and the Bayesian methods. The bootstrap approach is also a robust way to adjust for measurement error models (13). The non-informative Bayesian approach, is also employed in generalized linear mixed measurement error models, mainly because the Bayesian method is computationally convenient (10). Indeed, the inference about the parameters might change with different starting prior distributions. Lately a new statistical method called data cloning (DC), has been applied to compute the ML estimates and their standard errors (14). The DC method can also be applied in ML estimation of parameters and predicting the random effects, determining the estimability of parameters in mixed models (15). The data cloning approach, which avoids high dimensional numerical integration, has been applied to the generalized linear mixed measurement error models (16). With time-varying covariate measurement error, (17) proposed a joint modeling method in which a mechanistic nonlinear model is used for a longitudinal outcome that can be either discrete as binary and count or continuous.

An alternative to the aforementioned methods is using the Monte Carlo approach. Recently, it has been considered that the Monte Carlo Newton-Raphson (MCNR) method gives accurate estimates in the setting of a logistic regression model for analyzing longitudinal biomarker data, accounting for left-censoring and covariate measurement error (18). In this regard, the Monte Carlo approach has also been proposed to model the random effects covariance matrix in generalized linear mixed measurement error models for binary outcomes (19). In this method, the Monte Carlo Expectation Maximization (MCEM) algorithm is applied to estimate the parameters. Moreover, one can consult (20) on the problems of modern statistical regression modeling with the measurement error in the covariates, in order to conduct further research on the topic.

It is worth mentioning that a potential difficulty with the Monte Carlo method to approximate the integrals in generalized linear mixed measurement error models is that the integrals are computed many times and this might be undoubtfully expensive. The quadrature approach is another method of approximation that is computationally less expensive. This approach uses a weighted summation to approximate the integral and the integrated variable is assessed on a grid of quadrature points selected from the domain of the integration function (21). For the integrals involving the normal distribution, the gauss-Hermite quadrature approximation method is used to obtain the corresponding weights and quadrature points.

There is a paucity of research accounting for polytomous response variables in the analysis of multilevel data with covariates subject to measurement error. Most research have concentrated on mixed effects models for correlated binary responses. In this paper, we develop the multivariate Gauss-Hermite quadrature method for approximating the intractable integrals of the likelihood function for the analysis of multilevel data involving nominal correlated response components and measurement error in the covariate. Based on simulation results, this method of approximating the likelihood function can yield accurate estimates of the fixed effects for each category of the response variable as well as variance components of random effects and the measurement error distribution. At first, in Section 2, we will describe generalized linear mixed measurement error models for nominal outcomes when covariates are subject to measurement error. In Section 3, we describe how the likelihood function (which yields ML estimation) is evaluated in the presence of mixed models with measurement error in the covariate. We then describe the way by which the multivariate Gauss-Hermite quadrature methodology can be applied to approximate the likelihood function including the measurement error model. In Section 4, the performance of the multivariate Gauss-Hermite quadrature method to correct for induced bias will be studied in a simulation study while the amount of reliability ratio changes, i.e., small error in the covariate, tolerable and intensive error. We will check the efficiency of our proposed approach to correct for measurement error and evaluate how mis-specifying the measurement error distribution might result in high biases and standard errors in estimating the parameters. In Section 5, the improvement of the proposed method will be evaluated by analyzing a real dataset depending on a multilevel study on contraceptive methods utilized in Bangladesh. Concluding remarks are provided in the last Section.

Materials and Methods

The Generalized Linear Mixed Measurement Error Model, Basics and Notation

Suppose $Y_i = (Y_{i1}, \dots, Y_{in_i})^T$ denotes the observed outcomes for the ith subject, $1 \le i \le m$, where m is the total number of independent individuals, and n_i is the number of observations for individual i. Assume Y_i follows a generalized linear mixed model with a random intercept for each individual. The vectors of model covariates are indicated by $X_i = (X_{i1}, \dots, X_{in_i})^T$ and also through $Z_i = (Z_{i1}, \dots, Z_{in_i})^T$. Let Y_{ij} be the response variable for subject i at the j-th occasion $(j = 1, \dots, n_i)$, X_{ij} as the vector of true covariates which are error-prone and unobserved and hence latent, and Z_{ij} be the vector of error-free covariates. Furthermore, let W_{ij} be the vector of measurements that are observed in the absence of X_{ij}.

The Outcome Model

Let Y_{ij} be a categorical response variable with B categories for subject i at occasion j. Multicategory logit model for Y_{ij} assuming within-subject variability can be defined as follows:

$$\log\left[\frac{P(Y_{ij} = b|x_{ij}, z_{ij}, \tau_i)}{P(Y_{ij} = B|x_{ij}, z_{ij}, \tau_i)}\right] = \beta_{0b} + \beta_{xb}x_{ij} + \beta_{zb}z_{ij} + \tau_i, \quad (2.1)$$

$$i = 1, \dots, m, \quad j = 1, \dots, n_i, \quad b = 1, \dots, B-1.$$

In (2.1), b is an index demonstrating different categories, as well as the baseline category is denoted by B in the notation. According to the outcome model (2.1), it can be seen that the fixed effects vary according to the response paired with the baseline. In (2.1), $\beta = (\beta_{0b}, \beta_{xb}, \beta_{zb})^T$ is the vector of fixed parameters. We assume that the random effects τ_i, $(i = 1, \dots, m)$ are independent, and also independent from the error-prone covariates X_{ij}. Here, we suppose $\tau_i \sim N(0, \sigma_\tau^2)$ with zero mean, where the variance is assumed to be constant for all subjects.

The Measurement Error Model

Assume the variable X_{ij} is subject to measurement error. For a continuous variable, any type of regression model can be employed to specify the relation between true values X_{ij} and the observed values W_{ij}. The most common model in the measurement error literature is based upon what is called classical measurement error, in which the true but latent variable is measured with additive error, usually assumed to have constant variance, i.e.;

$$W_{ij}|x_{ij} = x_{ij} + e_{ij}, \quad (2.2)$$

Where e_{ij} is the error term with $E(e_{ij}|x_{ij}) = 0$ and $Var(e_{ij}|x_{ij}) = \sigma_e^2$, assuming W_{ij} and X_{ij} are scalar. It is concluded that $E(W_{ij}|x_{ij}) = x_{ij}$, so W_{ij} is unbiased for the unobserved x_{ij}. We also consider that the measurement error in W_{ij} is non-differential. This means that Y_{ij} is conditionally independent of W_{ij}, given x_{ij}. In the following, it is assumed that we have homoscedastic measurement error, which refers to the case where the variance of W_{ij} given x_{ij} is constant.

We designate a fully structural case for the fallible covariate X_i, which is assumed i.i.d with

$$X_i = (X_{i1}^T, \dots, X_{in_i}^T)^T \sim N\left(\mu_x, \Sigma_x = diag(\sigma_{xj}^2)\right) \quad i = 1, \dots, m, \quad j = 1, \dots, n_i,$$

where

$$\mu_x = (\mu_{x1}, \dots, \mu_{xj})^T \quad , \quad \Sigma_x = \begin{bmatrix} \sigma_{x1}^2 & 0 & \dots & 0 \\ 0 & \sigma_{x2}^2 & \dots & 0 \\ \vdots & \vdots & \vdots & 0 \\ 0 & 0 & \dots & \sigma_{xn_i}^2 \end{bmatrix}.$$

The measurement error in the multivariate case, also has a normal distribution with the following structure:

$$e_i \sim N\left(0_{n_i}, \sigma_e^2 I_{n_i}\right).$$

It should be pointed out that to avoid complexity issues come to pass in estimating the parameters of interest in a simulation step, we will treat the mean together with the covariance matrix of the error-prone variable X_i to be fixed and develop a trajectory to correct for induced bias in the parameter estimation stage.

Correcting for Measurement Error Using Likelihood Inference Methodology

This section provides the way of constructing the likelihood function based on distributional assumptions recalled earlier. The likelihood formulation for the generalized linear mixed measurement error model involves up to four different parts: The model for Y_{ij} given x_{ij} for each of j (specific) occasions, the model for measurement error, the model for X_{ij} in the structural setting and the model for the random effects. We next exhibit how to build the likelihood function and later, delineate its proper approximation.

Likelihood Function

The method of maximum likelihood (ML) estimation is of course the most popular approach to identifying estimates of parameters in a statistical model. In this Section, we consider a parametric likelihood approach that allows for covariate measurement error in a multinomial logit model. According to (2.1) and (2.2), if θ is a vector of associated parameters, i.e., $\theta = \left(\beta, \sigma_e^2, \sigma_{xj}^2, \sigma_\tau^2\right)^T$, the likelihood function that incorporates the measurement error process in covariates, can be outlined as follows:

$$L(\theta; y, w, z) = \prod_{i=1}^{m} \int_R^{\infty} \int_{R^{n_i}}^{\infty} \prod_{j=1}^{n_i} f\left(y_{ij}|x_{ij}, w_{ij}, z_{ij}, \tau_i; \beta\right) f\left(w_{ij}|x_{ij}; \sigma_e^2\right) f\left(x_i; \mu_x, \sigma_{xj}^2\right) f(\tau_i; \sigma_\tau^2) dx_{ij} d\tau_i, \quad (3.1)$$

where $f\left(y_{ij}|x_{ij}, w_{ij}, z_{ij}, \tau_i; \beta\right)$ equaling to $f\left(y_{ij}|x_{ij}, z_{ij}, \tau_i; \beta\right)$ due to the assumption of non-differential measurement error, has a multinomial distribution given the random effects. This density function is defined as follows:

$$f(y_i|x_i, z_i, \tau_i; \beta) = \prod_{j=1}^{n_i} \prod_{b=1}^{B} \left(P_{ijb}\right)^{y_{ijb}}. \quad (3.2)$$

In this regard, we assume that y_{ijb} for $i = 1, \dots, m$ and $j = 1, \dots, n_i$, is a binary variable that takes the value one if $y_{ij} = b$ and zero, otherwise. The quantity P_{ijb} is a pointwise probability which is defined as

$$P\left(Y_{ij} = b|x_{ij}, z_{ij}, \tau_i\right) = \frac{exp\left(\beta_{0b} + \beta_{xb}x_{ij} + \beta_{zb}z_{ij} + \tau_i\right)}{1 + \sum_{h=1}^{B-1} exp\left(\beta_{0h} + \beta_{xh}x_{ij} + \beta_{zh}z_{ij} + \tau_i\right)}. \quad (3.3)$$

The likelihood conveyed by (3.1), reflects the fact that the occurrence of measurement error in the covariate typically has more intricate consequences in modifying the arrangement of the likelihood function when compared with a situation dismissing measurement error in the covariate. As a result, the likelihood function (3.1), which encompasses measurement error in the covariate, can be considered as the likelihood function corresponding to our proposed methodology. The suggested approach will be titled as the measurement error method (ME in abbreviation) during the succeeding sections. To have a comparative study about the results obtained from the ME methodology, we will contemplate an alternative scenario, subsequently.

With random covariate X_i subject to error, we assume distributional assumption on error-prone covariate X_i , i.e., structural approach in which covariate X_i has a normal distribution with mean μ_x and variance Σ_x , and is independent of τ_i . Furthermore, let us consider W_{ij} as a surrogate for X_{ij} with classical additive measurement error (2.2), where $e_{ij} \sim N(0, \sigma_e^2)$. Then, the conditional distribution of $W_{ij}|x_{ij}$ can be written as $W_{ij}|x_{ij} \sim N\left(x_{ij}, \sigma_e^2\right)$. Moreover, τ_i has a normal density with zero mean and variance σ_τ^2 . Consequently, the probability density function (p.d.f) for random effects in the single setting can be written as follows:

$$f(\tau_i; \sigma_\tau^2) = (2\pi\sigma_\tau^2)^{-\frac{1}{2}} exp\left(-\frac{1}{2}\frac{\tau_i^2}{\sigma_\tau^2}\right). \quad (3.4)$$

In the comparative scenario, the model which naively disregards the difference between observed and true values of the covariates is nominated as the "naive" model. The likelihood function for this method assuming θ^* is a vector of associated parameters, i.e., $\theta^* = (\beta, \sigma_\tau^2)^T$, is as follows:

$$L(\theta^*; y, w) = \prod_{i=1}^{m} \int_R^{\infty} \prod_{j=1}^{n_i} f\left(y_{ij}|w_{ij}, z_{ij}, \tau_i; \beta\right) f(\tau_i; \sigma_\tau^2) d\tau_i. \quad (3.5)$$

Unfavorably, the integrals in equations (3.1) and (3.5) take place in dimensions such that their values depend on the random effects along with or without the latent true covariate X_{ij} on n_i occasions. Consequently, the observed data likelihood function for this model cannot be expressed in a tractable manner. This means that finding the ML estimates for the parameters will not be straightforward in application. For this case, standard numerical integration methods such as Multivariate Gauss-Hermite quadrature (MGHQ) methods can be imposed to approximate the integral (22).

In a simple one-dimensional setting, this means that the likelihood function is approximated using the Gauss-Hermite quadrature (GHQ) method and optimized with common approaches. The GHQ technique approximates the integral by a weighted sum of the integrand which is evaluated at a number of quadrature points in the domain of the integration function (23). The location of the quadrature points and weights depends on the integrand and also on the domain of the integration function. For the integrals where the domain of the integration is $(-\infty, \infty)$ or the entire real line, and the integrand is the product of a specific function with a normal density, the locations of the quadrature points are the solutions to the Hermite polynomial function. See, for example (24) and (25), for more details.

The approximated likelihood function will then be maximized by invoking well-known standard algorithms such as the method known as Nelder and Mead or Newton-Raphson (27), yielding ML estimates for the parameters. To obtain variance components for the ML estimates, one can calculate the approximated observed information matrix.

Approximation of the Log-Likelihood Function

According to the likelihood function for the generalized linear mixed measurement error model (3.1), the log-likelihood function based on the observed data incorporating the measurement error distribution can be written as follows:

$$l(\theta; y, w, z) = \sum_{i=1}^{m} log\left[\int_R^\infty \int_{R^{n_i}}^\infty \prod_{j=1}^{n_i} \begin{matrix} f(y_{ij}|x_{ij}, w_{ij}, z_{ij}, \tau_i; \beta) f(w_{ij}|x_{ij}; \sigma_e^2) \\ f(x_i; \mu_x, \sigma_{xj}^2) f(\tau_i; \sigma_\tau^2) dx_{ij} d\tau_i \end{matrix}\right]. \quad (3.6)$$

Furthermore, the log-likelihood function for the naive method corresponding to (3.5), mis-specifying the measurement error in the covariate X_i is

$$l(\theta^*; y, w, z) = \sum_{i=1}^{m} log\left[\int_R^\infty \prod_{j=1}^{n_i} f(y_{ij}|w_{ij}, z_{ij}, \tau_i; \beta) f(\tau_i; \sigma_\tau^2) d\tau_i\right]. \quad (3.7)$$

For the approximation of the integral (3.6), multivariate forms of the Gauss-Hermite quadrature approximation method are needed to handle both unobserved covariate x_{ij} at n_i occasions and the random effects τ_i. Because both x_{ij} and τ_i are independent of each other, the product of their density functions can be defined as a multivariate normal distribution.

Multivariate Gauss-Hermite quadrature method (MGHQ) uses the idea of univariate GHQ rule for each coordinate of the integrated variables. For the approximation of the integrals in (3.6) which are $(n_i + 1)$ dimensional, a matrix of GHQ nodes is produced for latent variables X_{ij} at occasion j and also for the random effects τ_i according to the mean vector and variance-covariance matrix of the multivariate normal distribution $f(x_i, \tau_i;\ \sigma_{xj}^2, \sigma_\tau^2)$, which the integral is calculated. Furthermore, a vector of weights is calculated in the replacement with the multivariate normal distribution defined formerly.

For the integrals in (3.6), the approximation process will be done by choosing sets of GHQ nodes

$$\left\{x_{kj} = \left(x_{k_1 j}^{(1)}, x_{k_2 j}^{(2)}, \ldots, x_{k_{(n_i+1)}j}^{(n_i+1)}\right)' : 1 \le k_1 \le q_1; 1 \le k_2 \le q_2, \ldots, 1 \le k_{(n_i+1)} \le q_{(n_i+1)}\right\}$$

and

$$\left\{\tau_k = \left(\tau_{k_1}^{(1)}, \tau_{k_2}^{(2)}, \ldots, \tau_{k_{(n_i+1)}}^{(n_i+1)}\right)' : 1 \le k_1 \le q_1; 1 \le k_2 \le q_2, \ldots, 1 \le k_{(n_i+1)} \le q_{(n_i+1)}\right\}$$

And weights

$$\left\{w_k = \left(w_{k_1}^{(1)}, w_{k_2}^{(2)}, \ldots, w_{k_{(n_i+1)}}^{(n_i+1)}\right)' : 1 \le k_1 \le q_1; 1 \le k_2 \le q_2, \ldots, 1 \le k_{(n_i+1)} \le q_{(n_i+1)}\right\}.$$

In this setting, k refers to the indices $k_1, k_2, \ldots, k_{(n_i+1)}$, each index shows different sets of quadrature points assigned, in order to approximate each of $(n_i + 1)$ integrals (26). Moreover, $x_{k_1 j}^{(t)}$ and $\tau_{k_t}^{(t)}$ $\left(1 \le t \le (n_i + 1)\right)$ are the t-th root of the multivariate Hermite polynomial having degree t, i.e., $H_{q_t}(x_j, \tau)$ at occasion j $(j = 1, \ldots, n_i)$.

It is important to note that the total number of the integration quadrature points is $q = q_1 q_2 \ldots q_{(n_i+1)}$. But, if $q_1 = q_2 = \cdots = q_{(n_i+1)} = q_0$, i.e., the number of quadrature points for each dimension are assumed equal, then $q = q_0^{(n_i+1)}$. Adequate approximation usually needs a larger number of grids for standard errors. It is recommended to increase the number of grids basically until the changes are negligible and inconsequential in both estimates and standard errors.

Thereupon, the MGHQ approximation to the log-likelihood function (3.6) will have the following form:

$$l(\theta; y, w, z) = \sum_{i=1}^{m} log\left[\sum_{k_1=1}^{q_1}\sum_{k_2=1}^{q_2} \dots \sum_{k_{(n_i+1)}=1}^{q_{(n_i+1)}} w_{k_1}^{(1)} w_{k_2}^{(2)} \dots w_{k_{(n_i+1)}}^{(n_i+1)} \prod_{j=1}^{n_i} f\left(y_{ij}|x_{k_j}, z_{ij}, \tau_k; \beta\right) f\left(w_{ij}|x_{k_j}; \sigma_e^2\right)\right] \quad (3.8)$$

Under the assumption of naive methodology, to evaluate the integral (3.7) which includes only random effects τ_i as an integrated variable, a q-length vector comprised of quadrature points and the associated weights should be calculated from a Hermite polynomial with degree q. So, in this case, the set of GHQ nodes which are replaced to the random effects τ_i, with the corresponding weights, are defined in the following way:

$$\left\{\tau_k = (\tau_1, \dots, \tau_q)', w_k = (w_1, \dots, w_q)'\right\} \quad (3.9)$$

As a notation, τ_t $(t = 1, \dots, q)$ is the t-th zero of the Hermite polynomial, and w_t is the quadrature weight related with the quadrature point τ_t. With the associated nodes and weights in (3.9), the GHQ approximation of the log-likelihood function (3.7) will be of the following form:

$$l(\theta^*; y, w, z) = \sum_{i=1}^{m} log\left[\sum_{k=1}^{q}\prod_{j=1}^{n_i} f(y_{ij}|w_{ij}, z_{ij}, \tau_k; \beta)\, w_k\right]. \quad (3.10)$$

Simulations and Inferences

Now, we are going to conduct a simulation study to compare and appraise the efficiency of the Multivariate (Multidimensional) Gauss-Hermite Quadrature (MGHQ) approximation method in the analysis of multilevel data with measurement error in the covariate, i.e., the measurement error (ME) method to the aspect where the covariate measurement error has been mis-specified (Naive method). We aim to investigate how measurement error in the covariate might result in bias, the root of mean squared error (RMSE) and coverage probability/ratio (CR). We are interested in modeling the relationship between categorical nominal responses and two continuous covariates. The response variables are generated from a multinomial distribution. Following is the multinomial logit random effects model, used for data simulation:

$$log\left[\frac{P(Y_{ij} = b|x_{ij}, z_{ij}, \tau_i)}{P(Y_{ij} = B|x_{ij}, z_{ij}, \tau_i)}\right] = \beta_{0b} + \beta_{xb}x_{ij} + \beta_{zb}z_{ij} + \tau_i \quad b = 1, \dots, B-1,$$

for $i = 1, \dots, m$ individuals and $j = 1, \dots, n_i$ occasions, with Y_{ij} considered as a nominal response with B categories for the i-th subject at level-2, associated with level-1 measurement at occasion j. The probability that $Y_{ij} = b$ or the response corresponding to the i-th subject at occasion j, occurs in category b is given by:

$$P_{ijb} = P(Y_{ij} = b|x_{ij}, z_{ij}, \tau_i) = \frac{exp(\beta_{0b} + \beta_{xb}x_{ij} + \beta_{zb}z_{ij} + \tau_i)}{1 + \sum_{h=1}^{B-1} exp(\beta_{0h} + \beta_{xh}x_{ij} + \beta_{zh}z_{ij} + \tau_i)},$$

for $b = 1, \dots, B$ different categories of the outcome variable.

In this section, we assume a multinomial response with three different categories $(B = 3)$, where $(\beta_{01}, \beta_{x1}, \beta_{z1})$ are the intercept and covariates' fixed effects coefficients for the first category, and $(\beta_{02}, \beta_{x2}, \beta_{z2})$ are the fixed effects based on the second category, and the third category is regarded as the baseline (reference).

We assume that z_{ij} is an error-free or exactly measured covariate generated as a normally distributed variable following $N(0, 4^2)$ distribution, which is treated fixed during the simulation study. The measurement error model that we have contemplated is the classical additive model. In this case, we generate a surrogate variable W_{ij} for the error-prone covariate X_{ij}, as in (2.2), where e_{ij} is the measurement error variable for X_{ij}, independently and identically distributed following $N(0, \sigma_e^2)$, albeit we will change the magnitude of σ_e^2 during the simulation study. This is because we are going to check the impact of considering measurement error on the estimation procedure. For this purpose, we have nominated three different levels of measurement error variation: σ_e^2 =0.5, 1.3 and 2 corresponding to small, moderate and intensive error scenarios, respectively.

The true but latent covariate X_{ij} is assumed to have a structural modeling, generated from a parametric homogenous normal distribution, so that:

$$X_{ij} = \mu_x + V_{ij},$$

where the V_{ij} variables are independent following $N(0, \sigma_{xj}^2)$ for $j = 1,2,3$, and $\sigma_{xj}^2 = 1$. The covariance matrix regarding to X_i at different occasions can be expressed as:

$$\Sigma_x = \begin{bmatrix} \sigma_{x1}^2 = 1 & 0 & 0 \\ 0 & \sigma_{x2}^2 = 1 & 0 \\ 0 & 0 & \sigma_{x3}^2 = 1 \end{bmatrix}.$$

Moreover, we consider random effects τ_i to have a homogenous variance for all individuals, that is $\tau_i \sim N(0, \sigma_\tau^2)$. Here, we consider m=100 subjects with $n_i = 3$, as the number of follow-up for each individual. The vector of initial values including the fixed effects for both categories and the random effects variance are set to the following values:

$$\theta = (\beta_{01}, \beta_{x1}, \beta_{z1}, \beta_{02}, \beta_{x2}, \beta_{z2}, \sigma_\tau^2)' = (0.02, 0.6, 0.5, 0.05, 0.4, 0.5, 2.4)'.$$

Due to the fact that there are latent true covariates X_{ij} for each of three different occasions and random effects τ_i needed to be integrated out to approximate the likelihood function via the MGHQ method, the number of quadrature points used in each of the dimensions was set to 3, note that this restriction leads to generate eighty-one nodes for both X_{ij} and τ_i integrands, along with a vector of weights with a length equal to eighty-one.

It is important to note that the likelihood function (3.1) is approximated with the MGHQ approximation method and then the Newton-Raphson algorithm can be applied to calculate the MLEs of both fixed and random effects. Afterwards numerically approximating the likelihood function, the score function and the observed information matrix can be determined subsequently. Therefore, the variance components of the MGHQ estimates are derived by inverting the negative of the Hessian matrix.

With the intention of checking how covariate measurement error might influence the estimation procedure, we generate R=1000 different data sets, we then fit two distinct scenarios. The first one is defining a model ignoring the covariate measurement error (Naive method) and using the observed values of X_i. The second scenario is the ME method that incorporates covariate measurement error in the likelihood function, also considering homogenous variance for the random effects.

Results

Within the framework previously reported, we express the simulation results, and evaluate the performance of the two aforementioned methods. The results are presented in terms of the mean of absolute bias,

$$Bias_{\theta_0} = \frac{1}{R}\sum_{r=1}^{R}\left|\hat{\theta}_0^{(r)} - \theta_0\right|,$$

where θ_0 is considered as an initial and true value of a particular parameter, and $\hat{\theta}_0^{(r)}$ is the estimated value of θ_0 in the r-th simulation run, empirical standard error (SE) or standard deviation of the estimates over the simulations and root of mean squared error, i.e.,

$$RMSE = \sqrt{{Bias_{\theta_0}}^2 + Var_{\theta_0}},$$

where Var_{θ_0} is the average of $Var(\hat{\theta}_0)$ over R simulation runs. Furthermore, it is important to point out that the coverage probability is calculated as the proportion of times that the true value of a parameter is covered by the 95% confidence interval during the simulations.

In Table 1, the results of the two methods (Naive and ME approach) for R=1000 simulated data sets based on different magnitudes of covariate measurement error variance are presented. The simulation results for the fixed effects of the first category are denoted by $(\beta_{01}, \beta_{x1}, \beta_{z1})$, as well as the results based on the second category are shown as $(\beta_{02}, \beta_{x2}, \beta_{z2})$. Furthermore, the estimation of random effect variance and the measurement error variation are also indexed as σ_τ^2 and σ_e^2 , respectively.

From the results of Table 1 with small error $(\sigma_e^2 = 0.5)$, it is completely perspicuous that considering and correcting for measurement error leads to fewer biases in estimating the fixed effects of the first and second category in comparison with the Naive approach (0.0141 for β_{01}, 0.0151 for β_{x1} 0.0090 for β_{z1} and 0.0110 for β_{02}, 0.0087 for β_{x2} and 0.0120 for β_{z2}). The empirical standard errors for the ME approach are also smaller than the corresponding values of the Naive approach. In addition, the ME approach shows smaller RMSEs and better CRs for the nominal value confidence interval in estimating the fixed effects for both categories unlike the Naive approach. In the case at hand, the estimation of measurement error variance was found acceptable concerning the nominal approximate confidence interval.

In the tolerable measurement error case $(\sigma_e^2 = 1.3)$, we can discern considerable biases in estimating the fixed effects of both categories in the Naive approach (0.0166 for β_{01} , 0.0169 for β_{x1} 0.0140 for β_{z1}, also for the parameters of the second category: 0.0287 for β_{02}, 0.0087 for β_{x2} and 0.0132

Table 1. Parameter estimation (Est), absolute of bias (Bias), standard error (SE), root of mean squared error (RMSE), and coverage rate (CR) of the parameter estimates with $n_i = 3$ replicates and 1000 simulation run for the Naive and ME approaches.

		Naive					**ME**				
Scale of Error	Real value	Est	Bias	SE	RMSE	CR	Est	Bias	SE	RMSE	CR
Small error ($\sigma_e^2 = 0.5$)	$\beta_{01} = 0.02$	0.0395	0.0195	0.2595	0.2605	0.951	0.0341	0.0141	0.2465	0.2469	0.953
	$\beta_{x1} = 0.6$	0.6206	0.0206	0.1814	0.1826	0.950	0.6151	0.0151	0.1696	0.1703	0.951
	$\beta_{z1} = 0.5$	0.5130	0.0130	0.0806	0.0816	0.947	0.5090	0.0090	0.0720	0.0725	0.947
	$\beta_{02} = 0.05$	0.0731	0.0231	0.2627	0.2637	0.942	0.0610	0.0110	0.2455	0.2458	0.948
	$\beta_{x2} = 0.4$	0.4089	0.0089	0.1771	0.1773	0.952	0.4087	0.0087	0.1728	0.1730	0.953
	$\beta_{z2} = 0.5$	0.5133	0.0133	0.0804	0.0815	0.949	0.5120	0.0120	0.0730	0.0739	0.952
	$\sigma_\tau^2 = 2.4$	2.5266	0.1266	1.0946	1.1019	0.944	2.4547	0.0547	0.9557	0.9573	0.952
	$\sigma_e^2 = 0.5$	-	-	-	-	-	0.6544	0.1544	0.0917	0.1796	0.643
AIC				-517.6758						-1484.882	
Tolerable error ($\sigma_e^2 = 1.3$)	$\beta_{01} = 0.02$	0.0366	0.0166	0.2570	0.2576	0.953	0.0205	0.0005	0.2541	0.2541	0.955
	$\beta_{x1} = 0.6$	0.6169	0.0169	0.1457	0.1466	0.938	0.6150	0.0150	0.1442	0.1450	0.952
	$\beta_{z1} = 0.5$	0.5140	0.0140	0.0800	0.0812	0.945	0.5074	0.0074	0.0759	0.0762	0.947
	$\beta_{02} = 0.05$	0.0787	0.0287	0.2550	0.2566	0.946	0.0472	0.0028	0.2437	0.2438	0.954
	$\beta_{x2} = 0.4$	0.4087	0.0087	0.1395	0.1397	0.946	0.4112	0.0112	0.1377	0.1381	0.949
	$\beta_{z2} = 0.5$	0.5132	0.0132	0.0790	0.0800	0.949	0.5091	0.0091	0.0756	0.0761	0.949
	$\sigma_\tau^2 = 2.4$	2.5207	0.1207	1.1187	1.1252	0.945	2.5019	0.1019	0.9714	0.9767	0.946
	$\sigma_e^2 = 1.3$	-	-	-	-	-	1.3279	0.0279	0.1886	0.1907	0.950
AIC				-528.2558						-1615.688	
Intensive error ($\sigma_e^2 = 2$)	$\beta_{01} = 0.02$	0.0080	0.0120	0.2640	0.2643	0.949	0.0216	0.0016	0.2442	0.2442	0.951
	$\beta_{x1} = 0.6$	0.6187	0.0187	0.1358	0.1370	0.949	0.6083	0.0083	0.1308	0.1311	0.957
	β_{z1}	0.5116	0.0116	0.0813	0.0821	0.940	0.5093	0.0093	0.0751	0.0757	0.953

for β_{z2}). The results also illustrate small standard errors, RMSEs and good CRs for the ME approach compared with the corresponding Naive method. The outcomes demonstrate that the estimated measurement error variance is reasonable and the reliability ratio at the j-th occasion can be scrutinized as the ratio of the variability of x_{ij} to that of w_{ij}, i.e., $\hat{\lambda}_j = \frac{\sigma_{xj}^2}{\sigma_{xj}^2 + \hat{\sigma}_e^2} = 0.43$.

From the consequences of intensive measurement error ($\sigma_e^2 = 2$), it can be concluded that the biases in fixed effects estimates are larger for the misspecification of the measurement error situation, and as expected, the ME approach provides fewer biases. In other words, the biases of parameter estimation in the ME method are fairly small approach (0.0016 for β_{01}, 0.0083 for β_{x1} 0.0093 for β_{z1}, and 0.0016 for β_{02}, 0.0069 for β_{x2} and 0.0118 for β_{z2}). Moreover, the RMSEs of the ME approach are much smaller, as well as good CRs. It is concluded from the estimation results that σ_e^2 has also become reasonable corresponding to the nominal approximate confidence interval, in this case.

To find the model that minimizes the information loss, we have calculated the Akaike Information Criterion (AIC) value as a goodness of fit test assessed by the logarithm of the likelihood function for the Naive and the ME models. The mean of the AIC value over simulations selected the

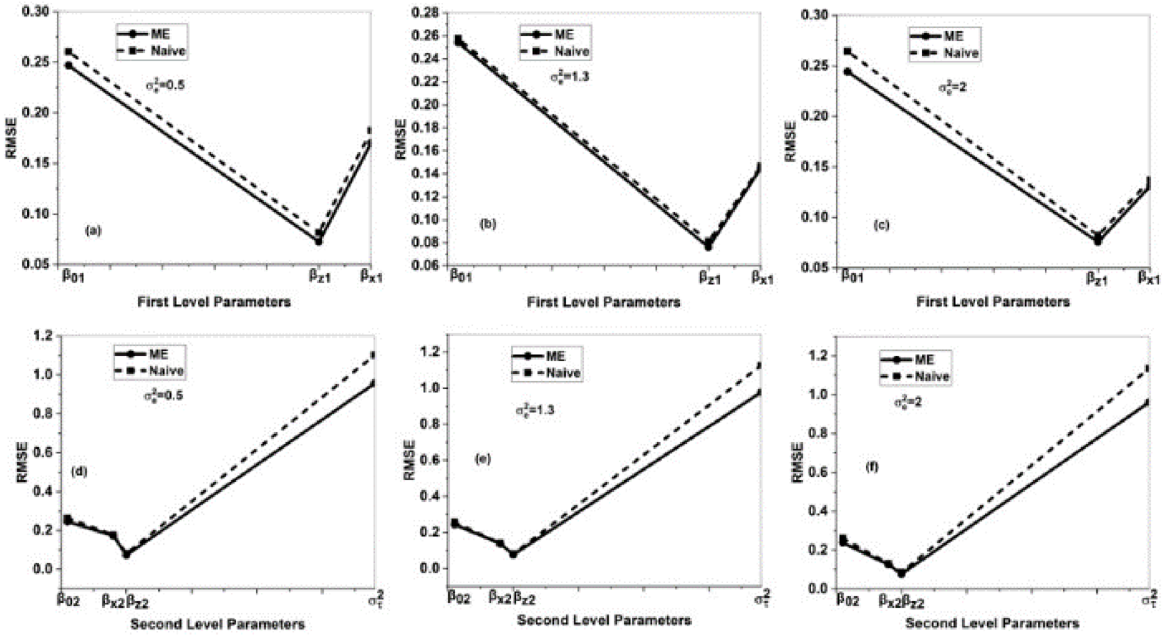

Figure 1. Comparison of RMSEs of Naive and measurement error parameters in multinomial logit random effect model based on simulation studies between two methods. The solid line corresponds to the RMSE estimates of the measurement error (ME) approach, while the dashed line, represents the Naive RMSE estimates. The three plots in (a), (b) and (c) corresponds to RMSEs against parameters of the first level of the response variable, with different levels of measurement error variance, while (d), (e) and (f) represents the RMSEs against the parameters of the second level of the response variable as well as the variance component of random effects.

correct model which illustrates the fact that the ME approach minimizes the information loss and provides a better fit than the Naive approach for all three different levels of measurement error variation.

In Figure 1, we have evaluated the RMSEs in Naive and ME approaches for the parameters of the first and second level of the response variable and also the random effect component based on three different levels of measurement error variation. It is important to note that the major feature of the plot is that the standard errors of the Naive method are larger than those of the ME method, and the ME method consistently gives smaller estimates of the standard errors for all fixed parameters as well as the variance component of the random effect involved in the simulation study.

In Table 2, the results of the estimation of fixed and random coefficients for the first and second categories under $n_i = 5$ follow-up and intensive measurement error variation over R=1000 simulation runs are presented. In this case, to calculate the likelihood function, six-dimensional integration must be approximated via the MGHQ method. From the results, it is obvious that the ME approach performs better than the corresponding Naive method in terms of Bias, SE, RMSE and CR for the 95% nominal approximate confidence interval. As expected, the RMSEs for both approaches tend to decrease with increasing the number of follow-up.

The RMSE estimates of the first and second level parameters of the response variable and the variance component of the random effect for the ME approach with $\sigma_e^2 = 2$ based on $n_i = 3$ and $n_i = 5$ replicates are plotted in Figure. 2. The particular component recognized in this plot is that just as the number of follow-up increases, the RMSEs tend to decrease.

Application: Contraceptive Data in Bangladesh

Contraception methods or using birth control techniques have increased rapidly in all regions of Bangladesh since 1975. Due to the high prevalence rate of contraceptive use, Bangladesh has encountered a decline in fertility and population growth. Some researchers believe that most of the childbearing decline is due to the efforts of Bangladesh national

Table 2. Parameter estimation (Est), absolute of bias (Bias), standard error (SE), root of mean squared error (RMSE), and coverage rate (CR) of the parameter estimates with $n_i = 5$ replicates and 1000 simulation run for the Naive and ME approaches.

		Naive					ME				
Scale of Error	Real value	Est	Bias	SE	RMSE	CR	Est	Bias	SE	RMSE	CR
Intensive error ($\sigma_e^2 = 2$)	$\beta_{01} = 0.02$	0.0062	0.0138	0.2260	0.2264	0.944	0.0339	0.0139	0.2151	0.2156	0.947
	$\beta_{x1} = 0.6$	0.6076	0.0076	0.1024	0.1026	0.942	0.6057	0.0057	0.1004	0.1006	0.951
	$\beta_{z1} = 0.5$	0.5027	0.0027	0.0578	0.0579	0.941	0.5025	0.0025	0.0565	0.0565	0.953
	$\beta_{02} = 0.05$	0.0348	0.0152	0.2209	0.2214	0.952	0.0642	0.0142	0.2076	0.2081	0.956
	$\beta_{x2} = 0.4$	0.4029	0.0029	0.0990	0.0990	0.951	0.4017	0.0017	0.0973	0.0973	0.954
	$\beta_{z2} = 0.5$	0.5039	0.0039	0.0571	0.0573	0.948	0.5018	0.0018	0.0563	0.0563	0.956
	$\sigma_\tau^2 = 2.4$	2.2358	0.1642	0.6473	0.6678	0.939	2.2605	0.1395	0.6470	0.6619	0.944
	$\sigma_e^2 = 2$	-	-	-	-	-	1.9957	0.0043	0.1875	0.1875	0.940
AIC				-872.5785						-2784.728	

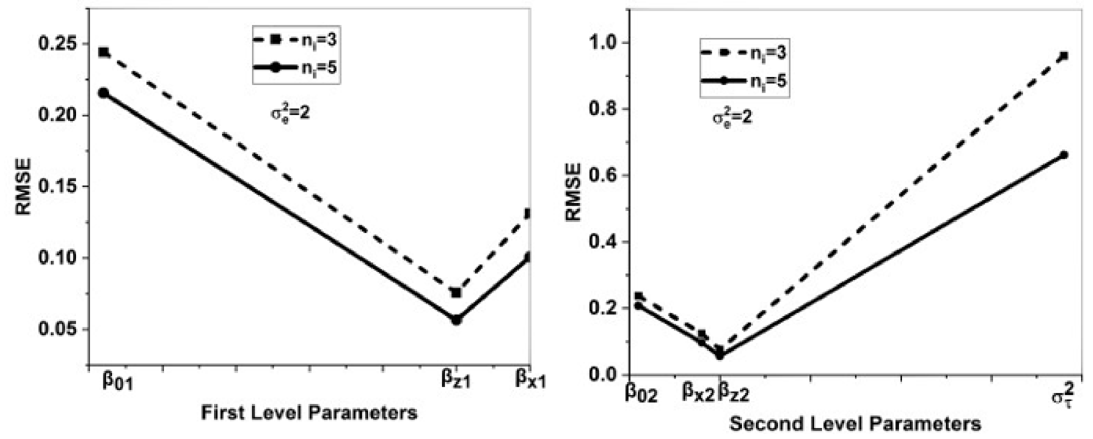

Figure 2. Comparison of RMSE estimates of measurement error approach parameters according to the first and second level of the response variable as well as the variance of the random effects in multinomial logit random effect model based on 1000 simulation studies with n_i= 3 and n_i = 5 replicates and intensive error. The solid line corresponds to RMSEs of measurement error (ME) parameters with n_i = 5, while the dashed line, represents the RMSEs based on n_i = 3 replicates.

family planning programs. Instead, others believe that changing social and economic situations also played an important role in the substantial decline in the fertility rate (28). Meanwhile, for the policy consequences purpose, it is also important to identify the factors associated with the use of different contraception methods.

In this Section, we use a sub-sample data from the 1989 Bangladesh Fertility Survey (29). The aim of the analysis is to identify the covariates associated with the contraceptive behavior, concurrently accounting for measurement error in some covariates. It is important to consider the fact that due to the deficiency of data associated with the analysis carried out in the ensuing article, i.e., scarcity of available data corresponding to correlated nominal response, in addition to validation data sources to overcome the issue of measurement error in the covariates, the proposed method has been applied to analyze a dataset from the contraception methods utilized in Bangladesh, as an illustration. Although the suggested methodology has been formulated for the multilevel data structure, it can also be assigned to data with a longitudinal design together with family studies, in

Table 3. Descriptive statistics for the variables associated with the analysis, Bangladesh, 1989. Abbreviation: Sd refers to standard deviation. No. is number of.

Variables	Frequency (%)	Mean (Sd)
Contraceptive methods		
Sterilization	302 (10.53)	-
Modern methods	555 (19.36)	-
Traditional methods	282 (9.84)	-
Not using contraception	1728 (60.27)	-
Age	-	29.3076 (8.6998)
No. children in the family	-	2.4405 (1.0528)
District		
Urban	804 (28.04)	-
Rural	2063 (71.96)	-
Religion		
Muslim	2480 (86.5)	-
Hindu	387 (13.49)	-
Education level		
Lower primary	1806 (62.99)	-
Upper primary	439 (15.31)	-
Secondary and above	265 (9.24)	-
None	357 (12.45)	-
Total	2867	

which the association of parameters is a subject of interest. In Table 3, we have cited information about some factors that might be incorporated in modelling fertility behavior in Bangladesh. According to the descriptive statistics of the data in Table 3, it can be concluded that 10.5% of women (or their husbands) were sterilized, 19.4% were using a modern reversible method (mainly pills in Bangladesh), 9.8% were using a traditional method, and 60.3% were not using any contraception (30). The multinomial response variable that we have used, consists of four categories, that distinguishes between different contraception methods, including sterilization, modern or efficient methods (e.g. pills and IUDs), traditional or inefficient methods (e.g. withdrawals) and not using contraception.

Since multilevel models are well qualified for recognizing hierarchical variation among regions, we will examine the geographical district influences by fitting a multinomial logit random effect model with 2867 women nested within 60 districts (30, 31). The main covariates incorporated in this analysis are women's age, region of residence (urban, and rural as the reference), religion (Hindu, and Muslim as the reference), having children and women's education. The main reason for taking these covariates into consideration is because they are found to have meaningful effects on contraceptive behavior among married women in some researches.

Based on the results of observed data in Table. 3, it has been concluded that the use of contraceptive prevalence involves the married women aged between 15-49 years who were currently using at least one method of contraception with a mean 29.3 and a standard deviation 8.7. Women's age plays an important role in using contraceptive methods. According to related researches (see, for example (31), middle aged women have a higher chance of using contraception and as women become older, the use of contraception methods decreases. The women in the analysis mostly have low level of education with 2.4 children. In the subsequent, we consider the situation where there has been inaccurate measurement occurrence in recording the covariate age for the women covered in the analysis, no matter of the region of residence.

In the contraceptive behavior dataset in Bangladesh, to have a precise and rigorous analysis, there are some important factors to be included in the model such as measurement problems in gathering data, which have been originated from the quality of family planning services in recording the covariates. According to (32), the development of family planning services would not be effective in increasing contraceptive use, which is principally due to social and economic conditions in Bangladesh.

Bangladesh is predominantly rural and is economically reliant on agriculture; so desired family size is so high and children are valuable in the family for their beneficial role in production. Due to these facts, infant and child mortality rates are relatively high; and education levels are very low, mainly for women. These factors are consistently related to the decrease in the quality of family planning services and policy. In this regard, we consider the fact that there has been an inaccurate measurement episode in documenting the covariate age for the women involved in the analysis. Hence, we focus on the perception where there exists a variability in age measurement reported by the interviewers or informed by the women associated with the interview. Let Y_{ij} be a nominal response variable, with four categories comprising sterilization, modern reversible methods, traditional methods, and not using contraception, which has been determined as the reference category in the model. The model we have dealt with, is of the following form:

$$\log\left[\frac{P(Y_{ij}=b)}{P(Y_{ij}=B)}\right] = \beta_{0b} + \beta_{1b}Age_{ij} + \beta_{2b}Urban_{ij} + \beta_{3b}Hindu_{ij} + \beta_{4b}Lc_{ij} + \beta_{5b}Educ_{ij} + \tau_j. \quad (5.1)$$

for $i = 1, \dots, 2867$ women nested in $j = 1, \dots, 60$ districts, involving b = 1, 2, 3 categories and B =4 is the reference level (not using contraception). It is worth noting that $Urban_{ij}$ is a binary covariate indicating the residence area of individual i at district j, along with $Hindu_{ij}$ showing whether the individual's religion is Muslim or Hindu, Lc_{ij} is the number of children living in the family, and $Educ_{ij}$ is the education level of women i at district j. In this setting, Age_{ij} represents the true age of individual i at district j, which is latent and treated as the error-prone covariate. The Kolmogorov-Smirnov test has been implemented to test the normality of variable Age. The value of the test statistic is D = 0.16778 with p-value=0.3295, accordingly the test has accepted the null-hypothesis. It is assumed that W_{ij} is the age recorded for woman i at district j, following the classical additive structural measurement error model $W_{ij} = Age_{ij} + e_{ij}$. It is assumed that $Age_{ij} \sim N(\mu_x, \sigma_x^2)$ and the measurement error variable follows independent normal distribution with mean 0 and variance σ_τ^2. Here, to specify the district effects on the probability of using each of the contraception methods defined earlier, it is assumed that the distribution of district effects τ_j is normal with mean 0 and variance σ_τ^2.

We have employed the Naive and ME approaches to analyze the data. These two approaches are used for studying the effect of ignoring measurement error in regression coefficients and variance components. The results of parameter estimation and corresponding standard errors based on the Naive and ME likelihood approaches are displayed in Table 4 and Table 5, respectively.

From the results of Table 4, it can be concluded that age has an inverse association with the use of sterilization than not using contraception. The effect of age covariate also shows that with the increase in women's age, the log odds of using modern and traditional contraception methods decrease concerning not using any contraception. No significant variation can be found in using sterilization than not using contraception in urban areas compared with rural regions. The results show that the log odds of using modern or traditional contraceptive methods rather than not using contraception increases in urban areas than women living in rural regions. Due to the fact that by using the sterilization method, there is no chance of fertility but in modern or traditional methods, there is still an attempt for women to be pregnant, this matter clearly shows families' demand for having additional children.

From Table 4, it can be concluded that there is no significant variation in contraceptive behavior according to religion. The results show that with the increase in the number of living children in the family, there is a positive effect on using each of the contraception methods than not using contraception. Based on the results, it can be seen that an increase in the women's education level does not have much impact on the choice of contraception methods.

To establish the significant determinants of contraceptive behavior in Bangladesh, producing true estimates, and taking the intrinsic measurement error on the covariate age into account, we have applied the proposed model in this paper to the Bangladesh data set. From the results of Table 5, it can be concluded that as women's age increases, the log odds of using any of the contraception methods decrease compared with not using contraception. For the women in urban areas, the log odds of using the sterilization method decrease compared with not using contraception, and it can be seen that married women in urban areas use mostly modern and traditional methods rather than not using any ways of contraception.

Moreover, Table 5 shows that there is a significant variation in terms of religious attitude. Non-Muslim women are using more of each of contraceptive behaviors than Muslims. Muslim women are less likely to use each of contraceptive method compared with non-Muslim women. This is mainly due to the

Table 4. Coefficients and standard errors (SE) based on the Naive approach for the contraceptive status other than not using contraception, Bangladesh, 1989. Abbreviation: * refers to significant at 0.05, - shows the reference level.

Explanatory Variables	Sterilization		Modern Methods		Traditional Methods	
	Coefficient	SE	Coefficient	SE	Coefficient	SE
Constant	1.1381	0.8622	-1.4692	0.8337	1.8807*	0.8922
Age	-0.0696*	0.0260	-0.1001*	0.0267	-0.1258*	0.0286
Residence						
Rural	-	-	-	-	-	-
Urban	0.1614	0.2824	1.8275*	0.2621	0.5727*	0.2854
Religion						
Muslim	-	-	-	-	-	-
Hindu	0.2337	0.2080	-0.3578	0.2060	-0.0327	0.2285
No. children in the family	0.6751*	0.1184	0.6328*	0.1131	0.2885*	0.1227
Education level	-1.3422*	0.22	-0.0485	0.1928	-0.6435*	0.2116
Random effects variance	Coefficient	SE				
σ_τ^2	1.0379*	0.2215				
			AIC= -6461.77			

Table 5. Coefficients and standard errors (SE) based on the ME approach for the contraceptive status other than not using contraception, Bangladesh, 1989. Abbreviation: * refers to significant at 0.05, - shows the reference level.

Explanatory Variables	Sterilization		Modern Methods		Traditional Methods	
	Coefficient	SE	Coefficient	SE	Coefficient	SE
Constant	3.8691*	0.5570	1.8591*	0.6025	4.0592*	0.5548
Age	-0.0721*	0.0082	-0.1340*	0.0089	-0.1855*	0.0087
Residence						
Rural	-	-	-	-	-	-
Urban	-0.5806*	0.1561	0.8440*	0.1712	0.6799*	0.1462
Religion						
Muslim	-	-	-	-	-	-
Hindu	-0.9716*	0.1813	-0.6290*	0.1874	-0.5711*	0.1784
No. children in the family	0.0517	0.0760	0.2972*	0.0858	-0.1007	0.0741
Education level	-0.7727*	0.0874	-0.0419	0.1001	-0.1807*	0.0832
Other model and measurement error parameters	Coefficient	SE				
μ_x	29.5705*	0.1984				
σ_x^2	46.6062*	1.7705				
σ_τ^2	2.8092*	0.2215				
σ_e^2	26.6204*	1.4112				
AIC= -27479.1						

fact that religious beliefs can decrease the contraceptive behavior. The results show that there is not any significant variation between the increase of living children in the family and using sterilization or traditional methods rather than not using contraception. However, it is realized that with the increase in the number of living children, the log odds of using modern reversible methods increases than not using any contraception. According to the results derived in Table 5, it can be seen that there is an inverse association between the increase in women's level of education and using sterilization or traditional methods compared with not using any contraception, and there is no significant variation between women's education and using modern methods rather than not using any of contraceptive behavior.

According to the results of Table 4 and Table 5, it can be concluded that using the observed and mis-measured covariates and estimating the parameters based on the Naive approach will inevitably misspecify the structure of fixed and random effects. This is why the results of the Naive and ME approaches are different. To avoid misleading results, the ME approach has been employed to correct for additive error using multivariate Gauss-Hermite quadrature approximation technique in a multilevel setting. For the Bangladesh data, the estimated reliability ratio is 0.6364, in other words $\frac{\hat{\sigma}_x^2}{\hat{\sigma}_x^2+\hat{\sigma}_e^2} = 0.6364$, which expresses that there is about 36% error joined with the covariate age. It's worth mentioning that the ME approach provides a better fit compared with the Naive method, using the Akaike Information Criterion value.

Discussion

Longitudinal and hierarchical studies are effective statistical inference methods in evaluating time-varying and multi-level structured response variables with the relevant covariates. A plethora of methods are available for inference and parameter estimation in the context of multilevel studies. The application of these methods relies on the assumption that the variables used in the study are precisely measured. However, there are many situations where this assumption is violated for some of the variables. Measurement error in any of the covariates leads to biased coefficients and incorrect inferences in estimating the parameters. Measurement errors may arise for different reasons and from various sources (10, 33). In the presence of covariate measurement error, an approximate based method called regression calibration or a simulation-based method called simulation extrapolation is frequently used, and the common difficulty with the likelihood-based method for the estimation of model parameters under covariate measurement error is intractable numerical integration.

In this paper, we proposed multivariate Gauss-Hermite quadrature approximation method for the likelihood inference of the generalized linear mixed model with repeatedly measured covariates subject to measurement error. We analytically derived the likelihood function for the case of multinomial outcome, with structural covariate subject to classical additive measurement error. To emphasize on the importance of involving measurement error in the covariate in the analysis of longitudinal data with nominal response, we compared the measurement error approach to the method where the covariate measurement error has been mis-specified. We then illustrated the measurement error effects on parameter estimation. For this purpose, we showed how the multivariate Gauss-Hermite quadrature ML method can be applied to approximate and to optimize the observed data likelihood. We investigated the procedure of multivariate Gauss-Hermite ML, a method which directly approximates the likelihood-product of the integrals on the error-prone covariates at different occasions and also on the random effects, together with a matrix of quadrature points and a vector of weights associated with the quadrature points.

Simulation studies indicated that mis-specifying the covariate measurement error distribution in the likelihood function, causes larger biases and standard errors in the parameter estimation. Comparisons of the ME and the Naive approach in terms of absolute bias, empirical standard error, root of mean squared error and coverage rate show that the multivariate Gauss-Hermite quadrature ML performs well in handling measurement error in the likelihood approximation and also corrects for the induced bias. Based on real dataset analysis, we scrutinized that in the presence of covariate measurement error, using the Naive approach to estimating the parameters will inaccurately specify the fixed and random effects components. To correct for mis-measured covariates, the ME approach has been applied using the multivariate Gauss-Hermite quadrature technique.

It is perceptible to enlighten that a general concern with covariate measurement error problems is whether the model is identifiable or not. In the presence of measurement error, usually all the parameters are not identifiable. To overcome this, it is often assumed that some additional assumptions or data being in the form of validation data, replication data, or instrumental ones are available to accomplish a measurement error analysis. As stated in (19), in longitudinal studies, repeated measurements are collected for error-prone covariates, and for identifying model parameters, these measurements can be used as replicates. In the present article, the aforementioned numerical adventure did not suggest that there is a matter with non-identifiability for the models considered (simulation study together with the real data application).

An exclusive type of measurement error for the discrete variables is called misclassification. As seen in our simulation studies, Naive analysis ignoring measurement error leads to inaccurate results, this is an imperative matter to consider that discounting for misclassification also brings about biased estimates of model parameters. In our future work, we aim to investigate the analysis of correlated data with measurement error in the covariate and handle categorical response misclassification, simultaneously. As there are many statistical models for analyzing correlated data originating from longitudinal and heirarachical studies, extending existing methods to transition or marginal models containing measurement error in the covariate and misclassification in the categorical response variable, will be the subject of our forthcoming endeavor.

Acknowledgments

The authors would like to thank the reviewers and editor for providing constructive comments leading to improve the quality of our manuscript.

Mathematics Subject Classification (2010): 62J12, 62F10.

References

1. Goldstein H. Multilevel Statistical Models. 4th ed. London: John Wiley & Sons, 2011.
2. Searle SR, Casella G and McCullach CE. Variance Components. New Jersey: John Wiley & Sons, 1992.
3. McCullach CE, Searle SR and Neuhaus JM. Generalized, Linear, and Mixed Models. 2nd ed. London: John Wiley & Sons, 2008.
4. McCulloch CE. Maximum likelihood variance components estimation for binary data. Journal of the American Statistical Association. 1994 Mar 1;89(425):330-335.
5. Tanner MA. Tools for Statistical inference: Observed Data and Data Augmentation. 2nd ed. New York: Springer Science and Business Media, 1993.
6. Diggle PJ, Liang KY and Zeger SL. Analysis of Longitudinal Data. Oxford: Oxford University Press, 1994.
7. Skrondal, A., and Rabe-Hesketh, S. Generalized Latent Variable Modeling: Multilevel, Longitudinal, and Structural Equation Models. Boca Raton: CRC Press, 2004.
8. Wu L. Mixed Effects Models for Complex Data. Boca Raton: CRC Press, 2009.
9. Stroup WW. Generalized Linear Mixed Models: Modern Concepts, Methods and Applications. Boca Raton: CRC Press, 2012.
10. Carroll RJ, Ruppert D, Stefanski LA and Crainiceanu CM. Measurement Error in Nonlinear Models: A Modern Perspective. Boca Raton: CRC Press, 2006.
11. Wang N, Lin X, Guttierrez RG. A bias correction regression calibration approach in generalized linear mixed measurement error models. Communications in Statistics-Theory and Methods. 1999 Jan 1;28(1):217-32.
12. Wang N, Lin X, Gutierrez RG, Carroll RJ. Bias analysis and SIMEX approach in generalized linear mixed measurement error models. Journal of the American Statistical Association. 1998 Mar 1;93(441):249-61.
13. Buonaccorsi JP, Romeo G, Thoresen M. Model-based bootstrapping when correcting for measurement error with application to logistic regression. Biometrics. 2018 Mar;74(1):135-44.
14. Lele SR, Dennis B, Lutscher F. Data cloning: easy maximum likelihood estimation for complex ecological models using Bayesian Markov chain Monte Carlo methods. Ecology letters. 2007 Jul;10(7):551-63.
15. Lele SR, Nadeem K, Schmuland B. Estimability and likelihood inference for generalized linear mixed models using data cloning. Journal of the American Statistical Association. 2010 Dec 1;105(492):1617-25.
16. Torabi M. Likelihood inference in generalized linear mixed measurement error models. Computational statistics & data analysis. 2013 Jan 1;57(1):549-57.
17. Zhang H, Wu L. An approximate method for generalized linear and nonlinear mixed effects models with a mechanistic nonlinear covariate measurement error model. Metrika. 2019 May 1;82(4):471-99.
18. Xie X, Xue X, Strickler HD. Generalized linear mixed model for binary outcomes when covariates are subject to measurement errors and detection limits. Statistics in medicine. 2018 Jan 15;37(1):119-36.
19. Hoque ME, Torabi M. Modeling the random effects covariance matrix for longitudinal data with covariates measurement error. Statistics in Medicine. 2018 Dec 10;37(28):4167-84.
20. Lipovetsky, S. Handbook of Measurement Error Models. Boca Raton: Chapman & Hall, 2023.
21. Hartzel J, Agresti A, Caffo B. Multinomial logit random effects models. Statistical Modelling. 2001 Jul;1(2):81-102.
22. Jaeckel P. A Note on Multivariate Gauss-Hermite Quadrature. London: ABN Amro, 2005. Retrieved from:
23. http://www.pjaeckel.webspace.virginmedia.com/ANoteOnMultivariateGaussHermiteQuadrature.pdf
24. 23. Agresti A. Categorical Data Analysis. New York: John Wiley & Sons, 2002.
25. Skrondal A and Rabe-Hesketh S. Generalized Latent Variable Modeling. Boca Raton: CRC Press, 2004.
26. Molenberghs G and Verbeke G. Models for Discrete Longitudinal Data. New York: Springer Science & Business Media, 2006.
27. Pan JX, Thompson R. Generalized linear mixed models: An improved estimating procedure. InCOMPSTAT: Proceedings in Computational Statistics 14th Symposium held in Utrecht, The Netherlands, 2000 2000 (pp. 373-378). Physica-Verlag HD.
28. Nelder JA and Mead RA. A simplex algorithm for function minimization. Comput J. 1965;7:308-13.
29. Steele F, Diamond I. Contraceptive switching in Bangladesh. Studies in Family Planning. 1999 Dec;30(4):315-28.
30. Huq NM and Cleland J. Bangladesh fertility survey, National Institute of Population Research and Training (NIPORT), 1990.
31. Rasbash J, Steele F, Browne WJ, Goldstein H. A user's guide to MLwiN, v2. 10. Centre for multilevel modelling, University of Bristol. 2009;192.
32. Hossain MB. Analysing the relationship between family planning workers' contact and contraceptive switching in rural Bangladesh using multilevel modelling. Journal of biosocial science. 2005 Sep 1;37(5):529.

33. DeGraff DS. Increasing contraceptive use in Bangladesh: The role of demand and supply factors. Demography. 1991 Feb;28:65-81.
34. Yi G and Cook RJ. Errors in the Measurement of Covariates. The Encyclopedia of Biostatistics. 2005; 3:1741-1748.