Journal of Statistical Research *https://doi.org/10.3329/jsr.v58i1.75423*
2024, Vol. 58, No. 1, pp. 151-180 *ISSN 0256 - 422 X*

NONSTATIONARY LONGITUDINAL AUTOREGRESSIVE MIXED MODEL FOR COUNT DATA WITH MEASUREMENT ERROR IN COVARIATES: ESTIMATION AND ASYMPTOTICS

ALWELL OYET

Department of Mathematics and Statistics
Memorial University of Newfoundland and Labrador, St. John's, A1C 5S7, NL, Canada
Email: aoyet@mun.ca

SUMMARY

Several methods have been proposed in the literature for computing unbiased and efficient estimates of the parameters of generalized linear models when the covariates are measured with error. However, to our knowledge, no documented research on computational techniques for parameter estimation currently exist in the literature when the data is a longitudinal count data influenced by an unobservable latent variable and observable covariates that are measured with error. In this paper, we propose a nonstationary conditionally Poisson mixed model for such data and develop unbiased estimating equations with iterative methods for computing estimates of the effect of the covariates, variance of the latent variable and the correlation index parameter. The performance of the iterative methods are examined through extensive simulation studies. The results show that the methods performed well when the magnitude of the measurement error is not so large as to dominate or mask the effect of the true covariates. Using observed longitudinal count data on the number of patents awarded to 168 firms in the United States from 1974 to 1979 along with associated covariate information on the type of firm, log of the book value of capital in 1972 and research and development (R & D) expenditures we have demonstrated how the methods proposed in this paper can be applied to a real data. In addition, we derive the influence function of the estimator of the covariate effect and discuss the asymptotic properties of the estimator.

Keywords and phrases: Asymptotic normality, Generalized quasilikelihood, Generalized method of moments, Measurement error.

1 Introduction

Measurement errors usually occur when observed values of some or all variables in a study, namely response and covariates, are not measured or recorded accurately. This is sometimes due to human or sampling error, faulty instruments or a combination of issues associated with the data collection process. It has been well established in the literature that measurement errors in variables induce inconsistency and bias in parameter estimates of the model used in fitting the data. Several real life examples where measurement error are a concern can be found in Schneeweiss and Augustin

(2006), Cheng and Ness (1999) and Buonaccorsi (2010) and in their references. This has led to an extensive amount of work in the literature on methods for assessing the impact of measurement errors and the development of techniques for correcting bias in parameter estimates, in particular when the responses are assumed to be Gaussian and independent. Some of the techniques that have been proposed in the literature include corrected score estimators (Rosner, Willett and Spiegelmann 1989, Nakamura 1990), moment based approaches (Wansbeek 2001), regression calibration methods (Gleser 1990), likelihood based techniques (Schafer 1993, Rabe-Haskett, Pickles and Skrondal 2003), simulation based methods (Cook and Stefanski 1994) and techniques based on modifying estimating equations (Stefanski 1985, Buornaccosi 2010). These authors considered a variety of regression functions and measurement error models under a wide range of assumptions. For instance, some authors made assumptions on the distribution of the observed values of the covariate x given the true values z leading to the popular functional model. Some have assumed that the observed covariate is fixed but that the true covariate X is random. The measurement error model is then referred to as the Berkson model. The most common assumption on the response y is that it is measured exactly. See for instance, Jowaheer, Sutradhar and Fan (2013), Nakamura (1990) and Stefanski (1985). These studies were however limited to the independent setup.

The existence of a large body of computational methods for correcting bias in parameter estimates of generalized linear models with measurement error in covariates and/or the response may be attributed to the fact that the bias in the parameters of linear models, in general, are relatively easier to address than biases caused by measurement error in models for longitudinal count or binary data. In this paper, we consider a situation where the ith, $i = 1, \dots, K$ subject or experimental unit provided a small number $T \geq 2$ of repeated count responses y_{it} collected at equally spaced time points $t = 1, 2, \dots, T$ along with associated p-dimensional covariate information $\boldsymbol{x}_{it} = (x_{it1}, \dots . x_{itp})'$ that are measured with error. Clearly, in this longitudinal setup the repeated responses from each of the experimental units will not be independent but will be serially correlated. As an example, consider the data on longitudinal count responses on the number of patents awarded to 168 firms in the United States from 1974 to 1979 along with associated covariate information on the type of firm, log of the book value of capital in 1972 and the time dependent research and development (R & D) expenditures from 1971 to 1979. In this example, $K = 168$, $p = 6$ and $T = 6$. In addition to the influence of the covariates on y_{it}, we also assume that there exist latent variables which also influence the repeated responses. In this case, the longitudinal correlation structure induced by the repeated count responses will then be conditional on the subject specific random effect. To our knowledge, computational techniques for unbiased and efficient estimation of the parameters of longitudinal models with such mixed effects on count data with measurement error in time dependent covariates have not been developed in the literature due to the complexity of the problem. The complexity of the problem arise from the dependence of the basic properties of the responses such as mean, variance and covariance functions on the time dependent covariates which are measured with error and on the variance parameter of the unobservable subject specific random effect. Also, the correlation between observations must be taken into account when developing methods for efficient estimation of the model parameters. Furthermore, one has to also address the difficult problem of estimating the variance parameter of the random effect. One of the significant contributions of this paper, is the development of bias corrected estimating equations along with iterative methods for computing efficient estimates of the effects of the covariates, the variance of the random effect and the correlation index parameter of a dynamic longitudinal Poisson mixed model for count data when the observed covariate vector $\boldsymbol{x}_{it}$ is measured with error.

There is a long history of using random effects to account for the lagged correlation in repeated Gaussian responses. See for instance McCullagh and Nelder (1989) and Verbeke and Molenbergh (2000) and their references. This approach, which assumes that the unobservable random effect on subjects are the same at each time point was also adopted by some authors such as Carroll, Lin and Wang (1997) and Buonaccorsi, Demidenko and Tosteson (2000), amongst others, to address the problem of measurement errors in covariates in longitudinal Gaussian data. One limitation of such a model is that it does not account for the effect of time on the correlation between pairs of observations. It also limits the correlation between pairs of responses to the equicorrelation structure. Such an extended model is however more suitable for data in which measurements y_{ij} as well as covariate information are collected from the jth, $j = 1, \ldots, n_i$, member of the ith, $i = 1, \ldots, K$ family. The random effect γ_i, then represents the common unobservable effect of the ith family on the measurements. In this case, there is no time effect to consider but the model accounts for the effect of the familial correlation between responses from members of the same family when estimating the model parameters.

In order to account for the effects of time and covariates on the correlation between observed longitudinal responses, some authors have proposed and studied dynamic models for longitudinal Gaussian (see Schmid, Segal and Rosner 1994 and Bun and Carree 2005), count (McKenzie 1988, Oyet and Sutradhar 2013, Zhang and Oyet 2014) and binary (Kanter 1975, Qaqish 2003, Sutradhar 2011) data. For instance, in the context of a longitudinal branching process with immigration, Oyet and Sutradhar (2013) considered a situation where K communities are at risk of an infectious disease. At time $t = 1$, y_{i1} individuals in the ith community were observed to have developed the disease. Aside from the effect of the covariate vector $\boldsymbol{x}_{i1} = (x_{i11}, \ldots, x_{i1p})'$ on y_{i1}, they argued that other latent variables denoted by γ_i exist that may also influence the count at any time point t. Conditional on γ_i, let y_{i1} follow a Poisson distribution with mean $\mu^*_{i1} = \exp(\boldsymbol{x}'_{i1}\boldsymbol{\beta} + \gamma_i)$, where $\boldsymbol{\beta}$ is the covariate effect common to all communities. Then, in the more general case, one may model the count at $t = 2, \ldots, T$ as,

$$y_{it}\,|_{\gamma_i} = \sum_{j=1}^{y_{i,t-1}} B_j(n_t, \rho)|_{\gamma_i} + d_{it}|_{\gamma_i}, \tag{1.1}$$

where $B_j(n_t, \rho)$ is a binomial random variable, with parameters n_t and probability of success ρ, representing the number of offsprings reproduced by the jth individual infected at time $t-1$. The average number of offsprings reproduced by one individual is commonly referred to as the reproduction number in the infectious disease literature. Thus, the estimation of the model parameters is an important subject in branching processes with immigration. It was also assumed that (a) $\gamma_i \sim N(0, \sigma^2_\gamma)$, (b) $d_{it}\,|_{\gamma_i} \sim \text{Poi}(\mu^*_{it} - \rho n_t \mu^*_{i,t-1})$, for $t = 2, \ldots, T$, where $\mu^*_{it} = \exp(\mathbf{x}'_{it}\beta + \gamma_i)$, for all $t = 1, \ldots, T$; (c) $d_{it}\,|_{\gamma_i}$ and $y_{i,t-1}\,|_{\gamma_i}$ are independent for $t = 2, \ldots, T$.

Wang, Carroll and Liang (1996) considered a situation where the covariates in a generalized linear regression model are replicated and the measurement errors in the replicated covariates are correlated. They however assumed that the responses corresponding to the replicated covariates collected from a large number of individuals were independent. This, was therefore not a longitudinal measurement error problem. During the same period, Sutradhar and Rao (1996) considered the problem of measurement errors in the covariates of a generalized linear model (GLM) for cluster correlated data. Cluster correlated data is a set of independent multivariate responses with covariates associated with each response. They developed a bias corrected method for estimating the parameters under the assumption that the measurement error variance is known or estimable by extending the approach of Stefanski (1985). In a related development, Wansbeek (2001) studied and developed necessary conditions for obtaining consistent bias corrected generalized method of moments (BCGMM) estimate for the parameters of a linear model with serially correlated errors for panel data with measurement errors in covariates. Xiao et al (2007) then established the efficiency properties of the BCGMM estimator of Wansbeek (2001). Later, Fan, Sutradhar and Rao (2012) noted that the repeated continuous responses in the panel data with measurement error in covariates may also be influenced by some unobservable individual random effect. Thus, they considered a linear mixed model with serially correlated errors to fit the data and proposed a bias corrected generalized quasi likelihood (BCGQL) method for parameter estimation. Through a simulation study, they found that the BCGQL estimates were more efficient than the BCGMM estimates. A more detailed review of early developments in measurement error models and methods can be found in Schneeweiss and Augustin (2006). Recently, Sutradhar and Rao (2016) considered a longitudinal fixed effects model for count data with measurement errors in covariates. They introduced a bias corrected generalized quasilikelihood (BCGQL) method for estimating the regression parameter $\boldsymbol{\beta}$ of the model.

In what follows, we establish the notations for the dynamic mixed model we have considered in Section 2 and derive the unbiased estimators of the basic properties of the model. These unbiased estimators were then used in Section 3 to modify the naive generalized quasilikelihood (NGQL) estimating equation for the regression parameter and the naive generalized method of moments (NGMM) estimating equations for the variance and correlation index parameters of the model. Furthermore, in order to facilitate the study of the effect of the magnitude of the measurement error $\boldsymbol{v}_{it} = (v_{it1}, \ldots, v_{itp})'$ on the proposed methods, we have assumed that $\boldsymbol{v}_{it}$ consist of two components, one which is time independent and another that is time dependent. Specifically, we will assume that $v_{itu} = k_u + e_{itu}$, where $k_u \sim N(0, \sigma_u^{*2})$ and $e_{itu} \sim N(0, \sigma_{e(u)}^{*2})$ are independent, $u = 1, \ldots, p$. So that, $cov(v_{itu}, v_{iru}) = \sigma_u^{*2}$ and $\phi_u = corr(v_{itu}, v_{iru}) = \sigma_u^{*2}/\sigma_u^2 > 0$, where $\sigma_u^2 = var(v_{itu}) = \sigma_u^{*2} + \sigma_{e(u)}^{*2}$, $u = 1, \ldots, p$. It is clear that as a result of these assumptions, additional information on the measurement error variance σ_u^2 and correlation ϕ_u parameters are required for the unique and efficient estimation of the regression parameters. Some authors have assumed that these measurement error variances are constant or known. See for instance Staudenmayer and Buonaccorsi (2005). Others have proposed estimating the error variances. We note that this later approach is more realistic. In Section 3, we have discussed one approach for estimating the measurement error variance σ_u^2 and correlation ϕ_u parameters in the context of longitudinal count data. However, our focus is on estimating the model parameters of the longitudinal mixed model. We have also derived a measure of the influence of an observation on the regression parameter estimate in Section 3 and discussed the asymptotic properties of the estimate based on the influence function. The performance of the proposed modified estimating equations and iterative methods is examined through a simulation study in Section 4. The results of an application to real data are discussed in Section 5.

2 Nonstationary Conditionally Poisson Autoregressive Mixed Model Subject to Measurement Error

When $n_t = 1$, the model (1.1) can be written as,

$$y_{it}\,|_{\gamma_i} = \sum_{j=1}^{y_{i,t-1}} b_j(\rho)|_{\gamma_i} + d_{it}|_{\gamma_i}, \tag{2.1}$$

where $P[b_j(\rho) = 0] = 1 - \rho$ and $P[b_j(\rho) = 1] = \rho$. Let $\boldsymbol{x}_{it} = (x_{it1}, \ldots, x_{itp})'$ be the time dependent observed covariate vector with measurement error $\boldsymbol{v}_{it}$ and $\boldsymbol{z}_{it}$ be the time dependent true covariate vector, so that $\boldsymbol{x}_{it} = \boldsymbol{z}_{it} + \boldsymbol{v}_{it}$. If the true covariate vector were known, it can be shown that the count random variable Y_{it} in (2.1) is conditionally Poisson. That is, $\sigma^*_{iz,tt} = Var(Y_{it}|\gamma_i) = \mu^*_{iz,t}$. It can also be shown that conditional on γ_i, the lag k autocovariance function is given by $\sigma^*_{iz,t,t+k} = Cov(Y_{it}, Y_{i,t+k}|\gamma_i) = \rho^k \sigma^*_{iz,tt}$, $t = 1, 2, \ldots, T-1$; $k = 1, 2, \ldots, T-t$. The unconditional basic properties of the mixed model (2.1) are then obtained by averaging the conditional moments over the distribution of γ_i. The mixed model in (2.1) is nonstationary in the sense that the unconditional basic properties of the model are functions of the time dependent covariates $\boldsymbol{z}_{it}$ (see Fuller (1996, Section 1.2, Pg. 4 and Pg. 475)). More specifically, the unconditional mean, variance and autocovariance functions are given by

$$\mu_{iz,t} = E(Y_{it}) = \exp(\boldsymbol{z}'_{it}\boldsymbol{\beta} + \sigma^2_\gamma/2);\;\; \sigma_{iz,tt} = Var(Y_{it}) = \mu_{iz,t} + \mu^2_{iz,t}(e^{\sigma^2_\gamma} - 1),$$
$$\sigma_{iz,t,t+k} = Cov(Y_{it}, Y_{i,t+k}) = \rho^k \mu_{iz,t} + \mu_{iz,t}\mu_{iz,t+k}(e^{\sigma^2_\gamma} - 1), \tag{2.2}$$

respectively. Then, the lag k autocorrelation function of the count responses becomes

$$\rho_{iz,t,t+k} = \frac{\rho^k \mu_{iz,t} + \mu_{iz,t}\mu_{iz,t+k}(e^{\sigma^2_\gamma} - 1)}{\sqrt{\sigma_{iz,tt}}\sqrt{\sigma_{iz,t+k,t+k}}}.$$

In particular, if $\sigma^2_\gamma = 0$, we have that

$$\sigma_{iz,tt} = \mu_{iz,t} = \exp(\boldsymbol{z}'_{it}\boldsymbol{\beta}), \sigma_{iz,t,t+k} = \rho^k \mu_{iz,t},\; and\; \rho_{iz,t,t+k} = \rho^k \sqrt{\frac{\mu_{iz,t}}{\mu_{iz,t+k}}}$$

as in Sutradhar and Rao (2016). Thus, the longitudinal model for count data discussed by Sutradhar and Rao (2016) is a special case of the mixed model (2.1).

When the true covariates $\boldsymbol{z}_{it}$ are measured without error, Oyet and Sutradhar (2013) have shown, empirically, that a consistent and efficient estimate of the regression parameter vector β can be obtained by solving the GQL estimating equation,

$$\sum_{i=1}^{K} \frac{\partial \boldsymbol{\mu}_{iz}}{\partial \boldsymbol{\beta}} \Sigma_{iz}^{-1} (\boldsymbol{y}_i - \boldsymbol{\mu}_{iz}) = 0, \tag{2.3}$$

where $\boldsymbol{y}_i = (y_{i1}, \ldots, y_{iT})'$, $\boldsymbol{\mu}_{iz} = (\mu_{iz,1}, \ldots, \mu_{iz,T})'$ and Σ_{iz} is the covariance matrix of $\boldsymbol{y}_i$. Now, define $\boldsymbol{U}_{iz} = diag(\mu_{iz,1}, \ldots, \mu_{iz,T})$, $A_{iz} = diag(\sigma_{iz,11}, \ldots, \sigma_{iz,TT})'$, and $\Sigma_{iz} = A_{iz}^{1/2} \boldsymbol{C}_{iz} A_{iz}^{1/2}$ where,

$$\boldsymbol{C}_{iz} = \begin{pmatrix} 1 & \rho_{iz,1,2} & \rho_{iz,1,3} & \cdots & \rho_{iz,1,T} \\ & 1 & \rho_{iz,2,3} & \cdots & \rho_{iz,2,T} \\ & & 1 & \cdots & \rho_{iz,3,T} \\ & & \vdots & \vdots & \vdots \\ & & & \cdots & 1 \end{pmatrix},$$

is a $T \times T$ correlation matrix. Let $q_{iz,rj}$ be the rjth element of $\boldsymbol{Q}_{iz} = \boldsymbol{C}_{iz}^{-1}$. One can then verify that the GQL estimating equation (2.3) can be expressed as,

$$h_1(z, y; \mu, \sigma_\gamma, \rho) - h_2(z; \mu, \sigma_\gamma, \rho) = 0, \tag{2.4}$$

where

$$\begin{aligned} h_1(z, y; \mu, \sigma_\gamma, \rho) &= \sum_{i=1}^{K} \boldsymbol{Z}_i' \boldsymbol{U}_{iz} A_{iz}^{-1/2} \boldsymbol{Q}_{iz} A_{iz}^{-1/2} \boldsymbol{y}_i \\ &= \sum_{i=1}^{K} \sum_{j=1}^{T} \sum_{r=1}^{T} \exp[(\boldsymbol{z}_{ir} - \boldsymbol{z}_{ij})' \boldsymbol{\beta}/2] q_{iz,rj} d_{iz,rj}^{-1/2} \boldsymbol{z}_{ir} y_{ij}, \end{aligned} \tag{2.5}$$

and

$$\begin{aligned} h_2(z; \mu, \sigma_\gamma, \rho) &= \sum_{i=1}^{K} \boldsymbol{Z}_i' \boldsymbol{U}_{iz} A_{iz}^{-1/2} \boldsymbol{Q}_{iz} A_{iz}^{-1/2} \boldsymbol{\mu}_{iz} \\ &= \sum_{i=1}^{K} \sum_{j=1}^{T} \sum_{r=1}^{T} \exp[(\boldsymbol{z}_{ir} + \boldsymbol{z}_{ij})' \boldsymbol{\beta}/2] \exp(\sigma_\gamma^2/2) q_{iz,rj} d_{iz,rj}^{-1/2} \boldsymbol{z}_{ir}. \end{aligned} \tag{2.6}$$

In (2.5) and (2.6), $\boldsymbol{Z}_i$ is a $T \times p$ matrix with rows $\boldsymbol{z}_{ir}'$, $r = 1, \ldots, T$ and

$$d_{iz,rj} = 1 + (\mu_{iz,r} + \mu_{iz,j})[\exp(\sigma_\gamma^2) - 1] + \mu_{iz,r}\mu_{iz,j}[\exp(\sigma_\gamma^2) - 1]^2. \tag{2.7}$$

2.1 Unbiased estimators of basic properties of the mixed model

It is clear that in practical computations the estimating equations have to be expressed in terms of the observed covariates $\boldsymbol{x}_{it}$. Now, such a naive estimating equation, for example $h_1(x, y; \mu, \sigma_\gamma, \rho) -$

$h_2(x;\mu,\sigma_\gamma,\rho)=0$, will clearly lead to a biased estimate of the model parameters since $\boldsymbol{x}_{it}$ was measured with error. However, if we can find a modified estimating equation, say $H(x,y;\mu,\sigma_\gamma,\rho) = h^a(x,y;\mu,\sigma_\gamma,\rho) - h^b(x;\mu,\sigma_\gamma,\rho)$, such that $E[H(x,y;\cdots)] \approx h_1(z,y;\mu,\sigma_\gamma,\rho) - h_2(z;\mu,\sigma_\gamma,\rho) = 0$, it is possible to develop unbiased estimating equations for consistent and eficient estimation of the model parameters. It turns out that in order to find $H(x,y;\mu,\sigma_\gamma,\rho)$, unbiased estimates of moments of Y_{it} expressed in terms of the unobserved true covariates $\boldsymbol{z}_{it}$ are required. Therefore, we will begin by deriving unbiased estimates of some moments of the count response random variable $\boldsymbol{Y}_{it}$.

Now, in terms of the observed covariate vector $\boldsymbol{x}_{it}$, the conditional and unconditional means of the count response variable Y_{it} are $\mu^*_{ix,t} = \exp(\mathbf{x}'_{it}\boldsymbol{\beta} + \gamma_i) = \exp(\mathbf{z}'_{it}\boldsymbol{\beta} + \mathbf{v}'_{it}\boldsymbol{\beta} + \gamma_i)$ and $\mu_{ix,t} = \exp(\mathbf{x}'_{it}\boldsymbol{\beta} + \sigma^2_\gamma/2)$, respectively. Let $\Lambda = diag(\sigma^2_1,\ldots,\sigma^2_p)$ and $\kappa = \boldsymbol{\beta}'\Lambda\boldsymbol{\beta}$. Averaging the unconditional mean over the distribution of the measurement error vector $\boldsymbol{v}_{it}$, then yields

$$E[\mu_{ix,t}] = \exp(\mathbf{z}'_{it}\boldsymbol{\beta} + \kappa/2 + \sigma^2_\gamma/2) = \mu_{iz,t}\exp(\kappa/2).$$

Therefore, in terms of the observed covariates $\boldsymbol{x}_{it}$, an unbiased estimator of $\mu_{iz,t}$ is

$$\hat{\mu}_{iz,t} = \mu_{ix,t}\exp(-\kappa/2), \;\; t = 1,\ldots,T. \tag{2.8}$$

Similarly, it can be easily verified that

$$\widehat{\mu^2_{iz,t}} = \mu^2_{ix,t}\exp(-2\kappa) = \{\hat{\mu}_{iz,t}\}^2\exp(-\kappa), \;\; t = 1,\ldots,T, \tag{2.9}$$

is an unbiased estimator of $\mu^2_{iz,t}$. Then, for fixed σ^2_γ, it follows from (2.2) that

$$\hat{\sigma}_{iz,tt} = \mu_{ix,t}\exp(-\kappa/2) + \mu^2_{ix,t}\exp(-2\kappa)(e^{\sigma^2_\gamma} - 1), \tag{2.10}$$

is an unbiased estimator of $\sigma_{iz,tt}$.

Next we obtain the estimator of the product term $\mu_{iz,t}\mu_{iz,u}$. First, we note that $\mu_{iz,t}\mu_{iz,u} = \exp\{(\mathbf{z}_{it} + \mathbf{z}_{iu})'\boldsymbol{\beta} + \sigma^2_\gamma\}$. So, we explore the expectation, over the measurement error vector,

$$E_v[\exp\{(\mathbf{x}_{it} + \mathbf{x}_{iu})'\boldsymbol{\beta} + \sigma^2_\gamma\}] = \mu_{iz,t}\mu_{iz,u}\exp(\kappa + \kappa_\phi),$$

where $\kappa_\phi = \boldsymbol{\beta}'\Lambda_\phi\boldsymbol{\beta}$, with $\Lambda_\phi = diag(\sigma^2_1\phi_1,\ldots,\sigma^2_p\phi_p)$. Thus,

$$\widehat{\mu_{iz,t}\mu_{iz,u}} = \mu_{ix,t}\mu_{ix,u}\exp\{-(\kappa + \kappa_\phi)\} = \hat{\mu}_{iz,t}\hat{\mu}_{iz,u}\exp(-\kappa_\phi), \tag{2.11}$$

is the unbiased estimator of the product term $\mu_{iz,t}\mu_{iz,u}$ in the expression for the covariance (2.2). Therefore, it follows from (2.2), that

$$\hat{\sigma}_{iz,t,u} = \rho^{|t-u|}\hat{\mu}_{iz,t} + \widehat{\mu_{iz,t}\mu_{iz,u}}(e^{\sigma^2_\gamma} - 1), \tag{2.12}$$

is the unbiased estimator of the unknown covariance function $\sigma_{iz,t,u}$.

We now turn our attention to other components of $h_1(\cdot)$ and $h_2(\cdot)$ in (2.5) and (2.6), respectively. We note that $h_1(\cdot)$ and $h_2(\cdot)$ contain the expressions

$$\boldsymbol{z}_{ir}\exp[(\boldsymbol{z}_{ir} - \boldsymbol{z}_{ij})'\boldsymbol{\beta}/2] \;\text{ and }\; \boldsymbol{z}_{ir}\exp[(\boldsymbol{z}_{ir} + \boldsymbol{z}_{ij})'\boldsymbol{\beta}/2],$$

respectively, which cannot be computed since the covariates $\boldsymbol{z}_{ij}$ are unobserved. Now, define the functions

$$g_1{}^*(\boldsymbol{x}_{ir}, \boldsymbol{x}_{ij}; \boldsymbol{\beta}, \Lambda, \phi) = \mu_{ix,r}^{1/2}\mu_{ix,j}^{-1/2}[\boldsymbol{x}_{ir} - (\Lambda - \Lambda_\phi)\boldsymbol{\beta}/2]\exp[-\boldsymbol{\beta}'(\Lambda - \Lambda_\phi)\boldsymbol{\beta}/4]$$
$$g_2^*(\boldsymbol{x}_{ir}, \boldsymbol{x}_{ij}; \boldsymbol{\beta}, \Lambda, \phi) = \mu_{ix,r}^{1/2}\mu_{ix,j}^{1/2}\exp(-\sigma_\gamma^2/2)[\boldsymbol{x}_{ir} - (\Lambda + \Lambda_\phi)\boldsymbol{\beta}/2]\exp[-\boldsymbol{\beta}'(\Lambda + \Lambda_\phi)\boldsymbol{\beta}/4]. \tag{2.13}$$

Then, following Lemma 3 of Sutradhar and Rao (2016), we have that

$$\begin{aligned} E[g_1^*(\boldsymbol{x}_{ir}, \boldsymbol{x}_{ij}; \boldsymbol{\beta}, \Lambda, \phi)] &= \boldsymbol{z}_{ir}\exp[(\boldsymbol{z}_{ir} - \boldsymbol{z}_{ij})'\boldsymbol{\beta}/2], \text{ and} \\ E[g_2^*(\boldsymbol{x}_{ir}, \boldsymbol{x}_{ij}; \boldsymbol{\beta}, \Lambda, \phi)] &= \boldsymbol{z}_{ir}\exp[(\boldsymbol{z}_{ir} + \boldsymbol{z}_{ij})'\boldsymbol{\beta}/2]. \end{aligned} \tag{2.14}$$

We remark that if $\sigma_\gamma^2 = 0$, $d_{iz,rj} = 1$ in (2.5) and (2.6). We also note that $d_{iz,rj}$ can be estimated by

$$\hat{d}_{iz,rj} = 1 + (\hat{\mu}_{iz,r} + \hat{\mu}_{iz,j})[\exp(\sigma_\gamma^2) - 1] + \widehat{\mu_{iz,r}\mu_{iz,j}}[\exp(\sigma_\gamma^2) - 1]^2. \tag{2.15}$$

3 Estimating Equations and Iterative Procedures for Computing Estimates of β, σ_γ^2 and ρ

In this section we develop equations for estimating the parameters of model (2.1) with measurement error in $\boldsymbol{x}_{it}$. Since the equations are nonlinear, we have also proposed Newton-Raphson-type iterative methods for obtaining numerical solutions to the equations.

3.1 Modified GQL estimating equation for covariate effect

Using the unbiased estimating equations (2.8) - (2.15) it is clear that, for known values of $q_{iz,rj}$, we have

$$E[g_1^*(\boldsymbol{x}_{ir}, \boldsymbol{x}_{ij}; \boldsymbol{\beta}, \Lambda, \phi)q_{iz,rj}\hat{d}_{iz,rj}^{-1/2}] \approx \boldsymbol{z}_{ir}\exp[(\boldsymbol{z}_{ir} - \boldsymbol{z}_{ij})'\boldsymbol{\beta}/2]q_{iz,rj}\hat{d}_{iz,rj}^{-1/2}.$$

It follows that an approximate unbiased estimator of $h_1(\cdot)$ in (2.5) can be written as

$$\widehat{h}_1 = \sum_{i=1}^{K}\sum_{j=1}^{T}\sum_{r=1}^{T}\mu_{ix,r}^{1/2}\mu_{ix,j}^{-1/2}[\boldsymbol{x}_{ir} - (\Lambda - \Lambda_\phi)\boldsymbol{\beta}/2]\exp[-\boldsymbol{\beta}'(\Lambda - \Lambda_\phi)\boldsymbol{\beta}/4]q_{iz,rj}\hat{d}_{iz,rj}^{-1/2}y_{ij}.$$

In developing the R codes for actual computations in our simulation studies and applications to real data, we found that it was computationally convenient to express $\widehat{h}_1$ in matrix notations. For this purpose, we let $m_1 = \exp[-\boldsymbol{\beta}'(\Lambda - \Lambda_\phi)\boldsymbol{\beta}/4]$, $\boldsymbol{B}_{1\phi} = (\Lambda - \Lambda_\phi)/2$ and define the following matrices $\boldsymbol{D}_{iz,k} = diag(\hat{d}_{iz,k1}, \hat{d}_{iz,k2}, \ldots, \hat{d}_{iz,kT}) : T \times T$, and $\boldsymbol{M}_{1\phi} = diag(m_1, \ldots, m_1) : p \times p$. Also, let $\boldsymbol{J}_k$ be the $T \times T$ matrix with 1 as the kth diagonal element and zeros elsewhere. Then, the matrix $\boldsymbol{J}_k\boldsymbol{U}_{ix}^{1/2}\boldsymbol{Q}_{iz}\boldsymbol{U}_{ix}^{-1/2}$ will only have nonzero elements in the kth row. The expression for $\widehat{h_1}$ can then be written in matrix notations as,

$$\widehat{h}_1 = \sum_{i=1}^{K}\boldsymbol{M}_{1\phi}[\boldsymbol{X}_i' - \boldsymbol{B}_{1\phi}(\boldsymbol{\beta} \otimes \boldsymbol{1}_T')]\sum_{k=1}^{T}\boldsymbol{J}_k\boldsymbol{U}_{ix}^{1/2}\boldsymbol{Q}_{iz}\boldsymbol{U}_{ix}^{-1/2}\boldsymbol{D}_{iz,k}^{-1/2}\boldsymbol{y}_i, \tag{3.1}$$

where $\mathbf{1}_T$ is a $T \times 1$ vector with elements 1 and $\otimes$ is the kronecker product. Furthermore, for known values of $q_{iz,rj}$, we have that

$$E[g_2^*(\boldsymbol{x}_{ir}, \boldsymbol{x}_{ij}; \boldsymbol{\beta}, \Lambda, \phi) q_{iz,rj} \hat{d}_{iz,rj}^{-1/2}] \approx \boldsymbol{z}_{ir} \exp[(\boldsymbol{z}_{ir} + \boldsymbol{z}_{ij})' \boldsymbol{\beta}/2] q_{iz,rj} \hat{d}_{iz,rj}^{-1/2},$$

Then, an approximate unbiased estimator of $h_2(\cdot)$ in (2.6) can be written as

$$\widehat{h}_2 = \sum_{i=1}^{K} \sum_{j=1}^{T} \sum_{r=1}^{T} \mu_{ix,r}^{1/2} \mu_{ix,j}^{1/2} [\boldsymbol{x}_{ir} - (\Lambda + \Lambda_\phi)\boldsymbol{\beta}/2] \exp[-\boldsymbol{\beta}'(\Lambda + \Lambda_\phi)\boldsymbol{\beta}/4] q_{iz,rj} \hat{d}_{iz,rj}^{-1/2}.$$

Proceeding as before, it can be verified that one can rewrite $\widehat{h}_2$ in matrix notations as

$$\widehat{h}_2 = \sum_{i=1}^{K} \boldsymbol{M}_{2\phi} [\boldsymbol{X}_i' - \boldsymbol{B}_{2\phi}(\boldsymbol{\beta} \otimes \mathbf{1}_T')] \sum_{k=1}^{T} \boldsymbol{J}_k \boldsymbol{U}_{ix}^{1/2} \boldsymbol{Q}_{iz} \boldsymbol{U}_{ix}^{-1/2} \boldsymbol{D}_{iz,k}^{-1/2} \boldsymbol{\mu}_{ix}. \tag{3.2}$$

In (3.2), $\boldsymbol{B}_{2\phi} = (\Lambda + \Lambda_\phi)/2$ and $\boldsymbol{M}_{2\phi} = diag(m_2, \ldots, m_2)$, where $m_2 = \exp[-\boldsymbol{\beta}'(\Lambda + \Lambda_\phi)\boldsymbol{\beta}/4]$. It follows that the bias corrected GQL estimating equation for the regression parameter $\boldsymbol{\beta}$ is

$$f(x, y; \hat{\boldsymbol{\beta}}, \sigma_\gamma, \rho) = \widehat{h_1}(x, y; \hat{\boldsymbol{\beta}}, \sigma_\gamma, \rho) - \widehat{h_2}(x; \hat{\boldsymbol{\beta}}, \sigma_\gamma, \rho) = 0, \tag{3.3}$$

where $\widehat{h_1}(x, y; \mu, \sigma_\gamma, \rho)$ and $\widehat{h_2}(x; \mu, \sigma_\gamma, \rho)$ are given by (3.1) and (3.2), respectively.

We note that the estimating equation (3.3) is a function of the measurement error variance parameters ϕ_u, and σ_u^2, $u = 1, \ldots, p$, the regression parameter $\boldsymbol{\beta}$, variance of the random effect σ_γ^2 and the correlation index parameter ρ. It is clear that the measurement error variance parameters ϕ_u, and σ_u^2 are parameters of the observed covariates $\boldsymbol{x}_{it}$ and not that of the repeated count responses y_{it}. Thus, the estimation of ϕ_u, and σ_u^2 can be based entirely on $\boldsymbol{x}_{it}$. Following Sutradhar and Rao (2016, eqns (5.10) and (5.11)), σ_u^2 and ϕ_u, $u = 1, \ldots, p$ can be estimated as,

$$\begin{aligned} \hat{\sigma}_u^2 &= \frac{1}{KT} \sum_{i=1}^{K} \sum_{t=1}^{T} (x_{itu} - \bar{x}_{iu})^2, \quad \text{and} \\ \hat{\phi}_u &= \frac{1}{KT(T-1)} \sum_{i=1}^{K} \sum_{t=1}^{T} \sum_{r \neq t}^{T} \left(\frac{x_{itu} - \bar{x}_{iu}}{\hat{\sigma}_u} \right) \left(\frac{x_{iru} - \bar{x}_{iu}}{\hat{\sigma}_u} \right), \end{aligned}$$

respectively, where $\bar{x}_{iu} = \sum_{t=1}^{T} x_{itu}/T$.

Let $\boldsymbol{\theta} = (\phi_1, \ldots, \phi_p, \sigma_1^2, \ldots, \sigma_p^2, \boldsymbol{\beta}, \sigma_\gamma^2, \rho)'$. We observe that a closed form solution for $\boldsymbol{\beta}$ cannot be obtained from (3.3). Therefore, for fixed ϕ_u, σ_u^2, σ_γ^2 and ρ, we computed estimates of $\boldsymbol{\beta}$ by the Newton-Raphson iteration method

$$\boldsymbol{\beta}^{(r+1)} = \boldsymbol{\beta}^{(r)} - \left(\frac{\partial f}{\partial \boldsymbol{\beta}} \right)^{-1}_{\boldsymbol{\theta} = \boldsymbol{\theta}^{(r)}} f(\boldsymbol{\theta}^{(r)}), \tag{3.4}$$

where f is given by (3.3) and $\frac{\partial f}{\partial \boldsymbol{\beta}} = \frac{\partial \widehat{h}_1}{\partial \boldsymbol{\beta}} - \frac{\partial \widehat{h}_2}{\partial \boldsymbol{\beta}}$. Now, define

$$\begin{aligned}
\boldsymbol{H}_{1,irj} &= -m_1(\boldsymbol{x}_{ir} - \boldsymbol{B}_{1\phi}\boldsymbol{\beta})\boldsymbol{\beta}'\boldsymbol{B}_{1\phi}, \;\; \boldsymbol{W}_{1,irj} = -m_2(\boldsymbol{x}_{ir} - \boldsymbol{B}_{2\phi}\boldsymbol{\beta})\boldsymbol{\beta}'\boldsymbol{B}_{2\phi}, \\
\boldsymbol{H}_{2,irj} &= m_1\boldsymbol{B}_{1\phi}, \;\; \boldsymbol{W}_{2,irj} = m_2\boldsymbol{B}_{2\phi} \\
\boldsymbol{H}_{3,irj} &= \frac{m_1}{2}(\boldsymbol{x}_{ir} - \boldsymbol{B}_{1\phi}\boldsymbol{\beta})(\boldsymbol{x}_{ir} - \boldsymbol{x}_{ij})', \quad \boldsymbol{W}_{3,irj} = \frac{m_2}{2}(\boldsymbol{x}_{ir} - \boldsymbol{B}_{2\phi}\boldsymbol{\beta})(\boldsymbol{x}_{ir} + \boldsymbol{x}_{ij})'.
\end{aligned}$$

Then, it can be shown that

$$\begin{aligned}
\frac{\partial \widehat{h}_1}{\partial \boldsymbol{\beta}} &= \sum_{i=1}^{K}\sum_{j=1}^{T}\sum_{r=1}^{T} \mu_{ix,r}^{1/2}\mu_{ix,j}^{-1/2}[\boldsymbol{H}_{1,irj} - \boldsymbol{H}_{2,irj} + \boldsymbol{H}_{3,irj}]q_{iz,rj}\hat{d}_{iz,rj}^{-1/2}y_{ij}, \quad \text{and} \\
\frac{\partial \widehat{h}_2}{\partial \boldsymbol{\beta}} &= \sum_{i=1}^{K}\sum_{j=1}^{T}\sum_{r=1}^{T} \mu_{ix,r}^{1/2}\mu_{ix,j}^{1/2}[\boldsymbol{W}_{1,irj} - \boldsymbol{W}_{2,irj} + \boldsymbol{W}_{3,irj}]q_{iz,rj}\hat{d}_{iz,rj}^{-1/2}.
\end{aligned}$$

3.2 Modified GMM estimating equation for variance of random effect

The generalized method of moments estimating equation for computing σ_γ^2 we have constructed in this section is based on a function of squared and cross-product terms of all observations defined as

$$G(y) = \sum_{i=1}^{K}\sum_{t=1}^{T} Y_{it}^2 + 2\sum_{i=1}^{K}\sum_{t=1}^{T}\sum_{u<t}^{T} Y_{iu}Y_{it}. \tag{3.5}$$

Since the true covariates $\boldsymbol{z}_{it}$ are unknown, we will use the unbiased estimators we obtained in Section 2.1 to express the GMM estimating equation

$$\begin{aligned}
E[G(y)] - G(y) &= \exp(\sigma_\gamma^2)\left\{\sum_{i=1}^{K}\sum_{t=1}^{T}\mu_{iz,t}^2 + 2\sum_{i=1}^{K}\sum_{t=1}^{T}\sum_{u<t}^{T}\mu_{iz,u}\mu_{iz,t}\right\} &(3.6) \\
&+ \left\{\sum_{i=1}^{K}\sum_{t=1}^{T}\mu_{iz,t} + 2\sum_{i=1}^{K}\sum_{t=1}^{T}\sum_{u<t}^{T}\rho^{t-u}\mu_{iz,u}\right\} - G(y) = 0, &(3.7)
\end{aligned}$$

in terms of the observed covariates $\boldsymbol{x}_{it}$. Using equations (2.8), (2.9) and (2.11), it can be shown that an unbiased GMM estimating equation for σ_γ^2 can be written as $\mathcal{G} = \mathcal{G}_1 + \mathcal{G}_2 - G(y) = 0$, where

$$\begin{aligned}
\mathcal{G}_1 &= \exp(\sigma_\gamma^2)\Bigg\{\exp(-2\boldsymbol{\beta}'\Lambda\boldsymbol{\beta})\sum_{i=1}^{K}\sum_{t=1}^{T}\mu_{ix,t}^2 \\
&\quad + 2\exp[-\boldsymbol{\beta}'(\Lambda + \Lambda_\phi)\boldsymbol{\beta}]\sum_{i=1}^{K}\sum_{t=1}^{T}\sum_{u<t}^{T}\mu_{ix,u}\mu_{ix,t}\Bigg\}, \\
\mathcal{G}_2 &= \exp(-\boldsymbol{\beta}'\Lambda\boldsymbol{\beta}/2)\left\{\sum_{i=1}^{K}\sum_{t=1}^{T}\mu_{ix,t} + 2\sum_{i=1}^{K}\sum_{t=1}^{T}\sum_{u<t}^{T}\rho^{t-u}\mu_{ix,u}\right\}.
\end{aligned}$$

Again, in our simulation studies and in our application to real data we computed solutions to the GMM estimating equation for σ_γ^2 by iterating the Newton-Raphson procedure

$$\sigma_\gamma^{2(r+1)} = \sigma_\gamma^{2(r)} - \left(\frac{\partial \mathcal{G}}{\partial \sigma_\gamma^2}\right)^{-1}_{\boldsymbol{\theta}=\boldsymbol{\theta}^{(r)}} \mathcal{G}(\boldsymbol{\theta}^{(r)}), \tag{3.8}$$

to convergence (see Tables 2, 3 and 4). Using the terms

$$\begin{aligned}\frac{\partial\{\exp(\sigma_\gamma^2)\mu_{ix,t}^2\}}{\partial \sigma_\gamma^2} &= 2\mu_{ix,t}^2 \exp(\sigma_\gamma^2), \quad \frac{\partial \mu_{ix,t}}{\partial \sigma_\gamma^2} = \mu_{ix,t}/2, \quad \text{and} \\ \frac{\partial\{\exp(\sigma_\gamma^2)\mu_{ix,u}\mu_{ix,t}\}}{\partial \sigma_\gamma^2} &= 2\mu_{ix,u}\mu_{ix,t} \exp(\sigma_\gamma^2),\end{aligned}$$

it can be shown that $\partial\mathcal{G}/\partial\sigma_\gamma^2 = 2\mathcal{G}_1 + (1/2)\mathcal{G}_2$, in (3.8).

3.3 Modified GMM estimating equation for correlation index parameter

We begin by defining the lag-1 standardized autocovariance function of Y_{it} as,

$$S_{t,t+1} = \frac{1}{K(T-1)} \sum_{i=1}^{K}\sum_{t=1}^{T-1} \left(\frac{Y_{it} - \mu_{iz,t}}{\sigma_{iz,t}}\right)\left(\frac{Y_{i,t+1} - \mu_{iz,t+1}}{\sigma_{iz,t+1}}\right),$$

where $\sigma_{iz,t} = \sqrt{\sigma_{iz,t,t}}$. Then, using (2.2) we have that

$$E[S_{t,t+1}] = \frac{1}{K(T-1)} \sum_{i=1}^{K}\sum_{t=1}^{T-1} \frac{\rho\mu_{iz,t} + \mu_{iz,t}\mu_{iz,t+1}(e^{\sigma_\gamma^2} - 1)}{\sigma_{iz,t}\sigma_{iz,t+1}}.$$

Similarly, we can show that $E[S_{t,t}] = 1$, where

$$S_{t,t} = \frac{1}{KT} \sum_{i=1}^{K}\sum_{t=1}^{T} \left(\frac{Y_{it} - \mu_{iz,t}}{\sigma_{iz,t}}\right)^2.$$

Now, if $\boldsymbol{z}_{it}$ were known, an approximate GMM estimating equation for the correlation index parameter ρ can be written as

$$\frac{S_{t,t+1}}{S_{t,t}} - \frac{1}{K(T-1)} \sum_{i=1}^{K}\sum_{t=1}^{T-1} \mu_{iz,t}^{1/2}\mu_{iz,t+1}^{-1/2}[\rho + \mu_{iz,t+1}(e^{\sigma_\gamma^2} - 1)]d_{iz,t,t+1}^{-1/2} = 0, \tag{3.9}$$

where $\mu_{iz,t}$ and $d_{iz,t,t+1}$ are given by (2.2) and (2.7), respectively. Following the approach of Section 2.1, it can be shown that the unbiased estimators of $\mathcal{U}_{1,iz,t} = \mu_{iz,t}^{1/2}\mu_{iz,t+1}^{-1/2}, \mathcal{U}_{2,iz,t} = \mu_{iz,t}^{1/2}\mu_{iz,t+1}^{1/2}$ and $\mathcal{V}_{iz,t} = \sigma_{iz,t}\sigma_{iz,t+1}$ are, respectively given by

$$\hat{\mathcal{U}}_{1,iz,t} = m_1\mathcal{U}_{1,ix,t}, \quad \hat{\mathcal{U}}_{2,iz,t} = m_2\mathcal{U}_{2,ix,t}, \quad \text{and} \quad \hat{\mathcal{V}}_{iz,t} = \hat{\mathcal{U}}_{2,iz,t}\hat{d}_{iz,t,t+1}^{1/2},$$

where m_1 and m_2 are defined in §3.1. Using these unbiased estimators in (3.9) and solving for ρ, we obtained estimates of ρ by iterating

$$\hat{\rho} = \frac{\left[\frac{\hat{S}_{t,t+1}}{\hat{S}_{t,t}} - \frac{1}{K(T-1)}\sum_{i=1}^{K}\sum_{t=1}^{T-1}\hat{\mathcal{U}}_{2,iz,t}(e^{\sigma_\gamma^2}-1)\hat{d}_{iz,t,t+1}^{-1/2}\right]}{\frac{1}{K(T-1)}\sum_{i=1}^{K}\sum_{t=1}^{T-1}\hat{\mathcal{U}}_{1,iz,t}\hat{d}_{iz,t,t+1}^{-1/2}}, \tag{3.10}$$

to convergence, for fixed $\boldsymbol{\beta}$ and σ_γ^2, where

$$\hat{S}_{t,t} = \frac{1}{KT}\sum_{i=1}^{K}\sum_{t=1}^{T}\frac{(y_{it}^2 - 2y_{it}\hat{\mu}_{iz,t} + \widehat{\mu_{iz,t}^2})}{\hat{\sigma}_{iz,t,t}}, \text{ and}$$

$$\hat{S}_{t,t+1} = \frac{1}{K(T-1)}\sum_{i=1}^{K}\sum_{t=1}^{T-1}\frac{(y_{it}y_{i,t+1} - y_{it}\hat{\mu}_{iz,t+1} - y_{i,t+1}\hat{\mu}_{iz,t} + \widehat{\mu_{iz,t}\mu_{iz,t+1}})}{\hat{\mathcal{U}}_{2,iz,t}\hat{d}_{iz,t,t+1}^{1/2}}.$$

3.4 Asymptotic properties of regression parameter estimate based on the influence function

Plug-in estimates of the influence function of a statistic have been shown to be useful in computing estimates of the variance of the statistic and for measuring the influence of an observation or a set of observations on the statistic. See for instance Hampel (1974), Devlin, Gnanadesikan and Kettenring (1975) and Wasserman (2006, Section 2.3). More recently, Selvaratnam, Oyet, Yi and Gadag (2017) derived the asymptotic properties of the maximum likelihood estimators (MLE) of the parameters of a generalized linear mixed model for response adaptive designs based on the influence function of the MLE and introduced an influence function approach for parameter estimation. They found that estimates based on the influence function approach will in general have smaller bias. Given the influence of measurement errors on parameter estimates, we will in this section adopt the approach based on the influence function to discuss the asymptotic properties of the bias corrected GQL estimate of the regression parameter $\boldsymbol{\beta}$.

Let $\boldsymbol{Y}_1, \boldsymbol{Y}_2, \cdots, \boldsymbol{Y}_K$ be a sequence of independent random vectors with joint distribution function $F(\boldsymbol{t})$, $\boldsymbol{t} = (t_1, t_2, \ldots, t_T)$. Define the empirical distribution function of the observed responses $\boldsymbol{y}_1, \boldsymbol{y}_2, \cdots, \boldsymbol{y}_K$ as

$$F_K(\boldsymbol{t}) = \frac{1}{K}\sum_{i=1}^{K}\delta_{\boldsymbol{y}_i}(\boldsymbol{t}),$$

where $\delta_{\boldsymbol{y}_i}(\boldsymbol{t})$ is the indicator function

$$\delta_{\boldsymbol{y}_i}(\boldsymbol{t}) = \begin{cases} 1, & \text{if } y_{i1} \le t_1, y_{i2} \le t_2, \ldots, y_{iT} \le t_T \\ 0, & \text{otherwise.} \end{cases}$$

Then, using (3.1) and (3.2) the bias corrected estimating equation (3.3) for $\boldsymbol{\beta}$ can be written as

$$f(x, y; \hat{\boldsymbol{\beta}}, \sigma_\gamma, \rho) = \int [h_1^*(F_K) - h_2^*(F_K)]dF_K(\boldsymbol{y}) = 0, \tag{3.11}$$

where

$$h_1^*(F) = \boldsymbol{M}_{1\phi}(F)[\boldsymbol{X}' - \boldsymbol{B}_{1\phi}\{\boldsymbol{\beta}(F) \otimes \boldsymbol{1}_T'\}] \sum_{k=1}^{T} \boldsymbol{J}_k \boldsymbol{U}_x^{1/2}(F) \boldsymbol{Q}_z \boldsymbol{U}_x^{-1/2}(F) \boldsymbol{D}_{z,k}^{-1/2}(F)\boldsymbol{y},$$
$$h_2^*(F) = \boldsymbol{M}_{2\phi}(F)[\boldsymbol{X}' - \boldsymbol{B}_{2\phi}\{\boldsymbol{\beta}(F) \otimes \boldsymbol{1}_T'\}] \sum_{k=1}^{T} \boldsymbol{J}_k \boldsymbol{U}_x^{1/2}(F) \boldsymbol{Q}_z \boldsymbol{U}_x^{-1/2}(F) \boldsymbol{D}_{z,k}^{-1/2}(F)\boldsymbol{\mu}_x. \tag{3.12}$$

Clearly, any solution to (3.11) will be a function of $F_K(t)$ and hence can be written as $\boldsymbol{\beta}(F_K)$. Since the true value of $\boldsymbol{\beta}$ is also a solution to the estimating equation (3.3) we have that

$$\int [h_1^*(F) - h_2^*(F)] dF(\boldsymbol{y}) = 0, \tag{3.13}$$

and represent the true value as $\boldsymbol{\beta}(F)$. Using these results, we show that the covariate effect estimate obtained from (3.11) is consistent and asymptotically normal in distribution. The results are outlined in Theorem 1. The proof can be found in the Appendix.

Theorem 1. *Let $\boldsymbol{Y}_1, \boldsymbol{Y}_2, \cdots, \boldsymbol{Y}_K$ be a sequence of independent random vectors with distribution function $F(\boldsymbol{y})$. Let $F_K(\boldsymbol{y})$ be the empirical distribution function of the observed responses $\boldsymbol{y}_1, \boldsymbol{y}_2, \cdots, \boldsymbol{y}_K$. For an arbitrary distribution function G and $\epsilon > 0$, define $F_\epsilon = (1-\epsilon)F + \epsilon G$ to be the ϵ-contaminated distribution function of F. If $h_1^*(F)$, $h_2^*(F)$, $\partial h_1^*(F)/\partial\boldsymbol{\beta}$ and $\partial h_2^*(F)/\partial\boldsymbol{\beta}$ are given by (3.12) and (A.2) respectively, then,*

(a) the influence function of $\boldsymbol{\beta}$ at F is

$$IF(\boldsymbol{y}, \boldsymbol{\beta}, F)] = -\left\{E_F\left[\frac{\partial h_1^*(F)}{\partial\boldsymbol{\beta}} - \frac{\partial h_2^*(F)}{\partial\boldsymbol{\beta}}\right]\right\}^{-1} \left[h_1^*(F) - h_2^*(F)\right]. \tag{3.14}$$

(b) by the weak law of large numbers, $\boldsymbol{\beta}(F_K) \longrightarrow \boldsymbol{\beta}(F)$ as $K \longrightarrow \infty$.

(c) by Linderberg's central limit theorem, $\sqrt{K}(\boldsymbol{\beta}(F_K) - \boldsymbol{\beta}(F)) \xrightarrow{d} N\{\boldsymbol{0}, \Sigma\}$ as $K \longrightarrow \infty$, with $\Sigma = E_F[IF(\mathbf{y}, \boldsymbol{\beta}, F)IF(\mathbf{y}, \boldsymbol{\beta}, F)']$.

4 Simulation Studies

In this section, we examined the effect of measurement errors in covariates on the performance of the estimating equations and iterative methods we proposed for computing unbiased estimates of the parameters of the conditionally Poisson model (2.1). The results were obtained using an R code developed by the author for implementing the techniques.

4.1 Data generation

For the purpose of our simulation studies, we generated data from (2.1) using a combination of parameter values. First, we choose $K = 100$ subjects, $T = 4$ time points, $p = 2$ covariates and generated the true covariates from $z_{it1} \sim N(0,1)$ and $z_{it2} \sim (\chi^2_{(4)} - 4)/\sqrt{8}$. We note that in general, the true covariates can be time dependent. We then fixed $\mathbf{z}_{it} = (z_{it1}, z_{it2})'$ for each cycle of 1000 simulations. Next, for a fixed vector $\boldsymbol{\beta}$, chosen from the set $\{(0,0), (0.3, 0.1)\}$, we computed $\mu^*_{iz,t} = \exp(\mathbf{z}'_{it}\beta + \gamma_i)$ in (2.1), $i = 1, \ldots, K$, $t = 1, \ldots, T$ with γ_i generated from $\gamma_i \sim N(0, \sigma^2_\gamma)$ for a fixed value of σ^2_γ. The values of σ^2_γ we used were $0.25, 0.5, 0.75$. The next step was to generate $y_{i1} \sim Poi(\mu^*_{iz,1})$, for all i, using $\mu^*_{iz,1}$. To generate the remainder of the data from model (2.1) we note that the mean of the Poisson random variable $d_{it}|_{\gamma_i}$ must be non-negative. This implies that ρ must satisfy the condition $0 \leq \rho \leq \mu^*_{iz,t}/\mu^*_{iz,t-1}$, for all $t = 1, \ldots, T$. Thus, we first computed the ratio $\mu^*_{iz,t}/\mu^*_{iz,t-1}$ in the R code we developed for our simulation studies. Then, we used a value of ρ that is less than $min\{\mu^*_{iz,t}/\mu^*_{iz,t-1}\}_{t=1}^T - 0.005$ to generate y_{it}, $i = 1, \ldots, K$, $t = 2, 3, \ldots, T$. The values of ρ we used are shown in Tables 1, 2, 3 and 4. We note that when $\boldsymbol{\beta} = (0,0)'$, we have that $0 \leq \rho \leq 1$. In this case we used $\rho = 0.3$, 0.5 and 0.8 in our simulation. See Tables 3 and 4.

Now, in practice the longitudinal count data y_{it} and data on the covariates $\mathbf{x}_{it}$ with measurement errors will be available to the practitioner. However, data on the associated true covariates $\mathbf{z}_{it}$ will not be available. So, in our simulation studies our estimation was based on $\mathbf{x}_{it}$'s. We recall that $x_{itu} = z_{itu} + v_{itu}$, where $v_{itu} = k_u + e_{itu} \sim N(0, \sigma^2_u)$, $u = 1, 2, \ldots, p$. Consequently, we fixed $\phi_u = corr(v_{itu}, v_{iru})$ and σ^2_u. Then, for $\sigma^{*2}_{e(u)} = (1 - \phi_u)\sigma^2_u$, we generated $e_{itu} \sim N(0, \sigma^{*2}_{e(u)})$ and for $\sigma^{*2}_u = \phi_u \sigma^2_u$, we generated $k_u \sim N(0, \sigma^{*2}_u)$. We note that when $\phi_u = 1$, $\sigma^{*2}_{e(u)} = 0$. In this case, the measurement error in the covariate is contributed by only the component k_u and therefore may not be large. Therefore, in order to assess the impact of the magnitude of the measurement error on the parameter estimates, we considered $(\phi_1, \phi_2) = (1, 1), (0.25, 0.5)$ in our simulation studies. We also used $(\sigma^2_1, \sigma^2_2) = (0.1, 0.3), (0.3, 0.3), (0.5, 0.2)$.

Once a longitudinal data has been generated, we used $\boldsymbol{\beta} = (0,0)'$, $\sigma^2_\gamma = 0$ and $\rho = 0$ as starting values to iterate the Newton-Raphson procedure (3.4), for $\boldsymbol{\beta}$ estimation to convergence. Next, the improved estimate of $\boldsymbol{\beta}$ was used to iterate the procedure (3.8), for σ^2_γ estimation to convergence. The updated estimates of $\boldsymbol{\beta}$ and σ^2_γ were then used in (3.10) to compute a modified GMM estimate of ρ. These improved estimates of $\boldsymbol{\beta}$, σ^2_γ and ρ were then used to repeat the iterative process described above, until overall convergence to three decimal places was achieved. We then computed the mean and standard errors of the estimates obtained from 1000 simulations under various combinations of true parameter values. The results are reported in Tables 1 to 3. For the purpose of highlighting the need for bias correction we also estimated the model parameters using the naive GQL and GMM estimating equations in (2.4), (3.7) and (3.9). The simulated mean, absolute bias and mean squared errors of the naive and bias corrected estimates are shown side-by-side in Tables 4 and 5 for ease of comparison. We also examined the effect of small (T = 25) and large (T = 300) data on the estimates. Results from this comparison are shown in Table 6.

Table 1: Simulated means (SM) and standard errors (SSE) of modified GQL estimates of $\boldsymbol{\beta}$ obtained from 1000 simulations, for selected values of σ_γ^2, ρ, measurement error variances $\boldsymbol{\sigma}_u^2 = (\sigma_1^2, \sigma_2^2)$ and correlations ϕ_1, ϕ_2.

ϕ_1	ϕ_2	$\boldsymbol{\beta}'$	σ_γ^2	ρ	$\boldsymbol{\sigma}_u^2$		$\hat{\boldsymbol{\beta}}$
1	1	(0,0)	0.25	0	(0.1,0.3)	SM	(0.0010,-0.00006)
						SSE	(0.0496,0.0561)
					(0.3,0.3)	SM	(-0.00009,-0.0014)
						SSE	(0.0528,0.0535)
				0.3	(0.5,0.2)	SM	(0.0023,-0.0018)
						SSE	(0.0433,0.0492)
				0.5	(0.3,0.3)	SM	(0.00025,-0.0013)
						SSE	(0.0389,0.0374)
				0.8	(0.5,0.2)	SM	(-0.000068,-0.00082)
						SSE	(0.0194,0.0235)
		(0.3,0.1)	0.25	0	(0.1,0.3)	SM	(0.3042,0.1122)
						SSE	(0.0487,0.0514)
					(0.3,0.3)	SM	(0.3069,0.1055)
						SSE	(0.0653,0.0477)
				0.3	(0.1,0.3)	SM	(0.3005,0.1059)
						SSE	(0.0475,0.0461)
					(0.5,0.2)	SM	(0.3158,0.1081)
						SSE	(0.0713,0.0478)
			0.75	0	(0.1,0.3)	SM	(0.3001,0.1028)
						SSE	(0.0459,0.0422)
					(0.3,0.3)	SM	(0.3025,0.1054)
						SSE	(0.0507,0.0440)
				0.3	(0.5,0.2)	SM	(0.3023,0.1024)
						SSE	(0.0454,0.0376)
0.25	0.5	(0.3,0.1)	0.25	0	(0.1,0.3)	SM	(0.2810,0.0934)
						SSE	(0.0462,0.0419)
					(0.2,0.2)	SM	(0.2616,0.0996)
						SSE	(0.0486,0.0435)
			0.75	0	(0.1,0.3)	SM	(0.2800,0.0912)
						SSE	(0.0439,0.0371)
				0.2	(0.2,0.2)	SM	(0.2628,0.0936)
						SSE	(0.0420,0.0323)

Table 2: Simulated means (SM) and standard errors (SSE) of $\hat{\boldsymbol{\beta}}$ and $\hat{\sigma}_\gamma^2$ obtained from 1000 simulations, for selected values of ρ, measurement error variances $\boldsymbol{\sigma}_u^2 = (\sigma_1^2, \sigma_2^2)$ with correlations $\phi_1 = \phi_2 = 1$.

$\boldsymbol{\beta}'$	σ_γ^2	ρ	$\boldsymbol{\sigma}_u^2$		$\hat{\boldsymbol{\beta}}$	$\hat{\sigma_\gamma^2}$
(0,0)	0.25	0	(0.1,0.3)	SM	(0.0015,-0.0026)	0.233
				SSE	(0.049,0.053)	(0.084)
			(0.3,0.3)	SM	(0.0008,-0.0005)	0.239
				SSE	(0.049,0.055)	(0.087)
		0.3	(0.5,0.2)	SM	(0.0013,-0.0014)	0.230
				SSE	(0.044,0.043)	(0.097)
		0.5	(0.3,0.3)	SM	(0.0002,-0.0010)	0.223
				SSE	(0.038,0.036)	(0.111)
		0.7	(0.5,0.2)	SM	(-0.0000,-0.0025)	0.235
				SSE	(0.027,0.031)	(0.120)
	0.5	0	(0.3,0.3)	SM	(-0.0009,-0.0027)	0.477
				SSE	(0.049,0.047)	(0.128)
		0.3	(0.5,0.2)	SM	(-0.0007,-0.0036)	0.468
				SSE	(0.037,0.042)	(0.130)
	0.75	0	(0.1,0.3)	SM	(-0.0001,0.0003)	0.709
				SSE	(0.044,0.045)	(0.180)
(0.3,0.1)	0.25	0	(0.1,0.3)	SM	(0.284,0.089)	0.242
				SSE	(0.048,0.042)	(0.088)
			(0.3,0.3)	SM	(0.262,0.084)	0.257
				SSE	(0.050,0.050)	(0.093)
		0.3	(0.1,0.3)	SM	(0.280,0.083)	0.245
				SSE	(0.045,0.043)	(0.102)
			(0.5,0.2)	SM	(0.218,0.096)	0.260
				SSE	(0.042,0.042)	(0.103)
	0.5	0.3	(0.3,0.3)	SM	(0.230,0.087)	0.498
				SSE	(0.044,0.036)	(0.141)
	0.75	0.3	(0.3,0.3)	SM	(0.235,0.079)	0.737
				SSE	(0.039,0.034)	(0.189)

4.2 Performance of estimation procedures

The performance of the estimates computed using the iterative procedures were examined in stages. First, we computed only the MGQL estimate of the effect of the covariates $\boldsymbol{\beta}$, using (3.4), under the assumption that the true values of all other parameters are known. The simulated means and standard errors we obtained are shown in Table 1 under various combinations of the model parameters. The results in Table 1 show that the proposed iterative procedure for the MGQL approach performed well in estimating the effect of the covariates in the presence of measurement error in the covariates, in particular, when the magnitude of the measurement error is not too large (i.e. when $\phi_1 = \phi_2 = 1$ as in Sutradhar and Rao, 2016). For instance, when $\phi_1 = \phi_2 = 1$ which implies that $\sigma^{*2}_{e(u)} = 0$, $u = 1, 2$, with $\boldsymbol{\beta}' = (0.3, 0.1)'$, $\sigma^2_\gamma = 0.75$, $\rho = 0.3$, and $\boldsymbol{\sigma}^2_u = (0.5, 0.2)$, the MGQL estimates were found to be $\hat{\boldsymbol{\beta}} = (0.3023, 0.1024)'$ with standard errors $s_{\hat{\beta}} = (0.0454, 0.0376)'$. Also, when $\phi_1 = 0.25,\ \phi_2 = 0.5$ ($\sigma^{*2}_{e(u)} \neq 0$), with $\boldsymbol{\beta} = (0.3, 0.1)'$, $\sigma^2_\gamma = 0.75$, $\rho = 0$, and $\boldsymbol{\sigma}^2_u = (0.1, 0.3)$, we obtained $\hat{\boldsymbol{\beta}} = (0.2800, 0.0912)'$ with standard errors $s_{\hat{\beta}} = (0.0439, 0.0371)'$.

Next, we examined the performance of the proposed methods when the effect of the covariates $\boldsymbol{\beta}$ and the variance of the random effect σ^2_γ are estimated. In this case, the MGQL and MGMM estimates we obtained are shown in Table 2. Again, we see that the MGQL and MGMM methods performed well in estimating $\boldsymbol{\beta}$ and σ^2_γ, respectively. As an example, when the true values of the model parameters are $\boldsymbol{\beta} = (0, 0)'$, $\sigma^2_\gamma = 0.5$, $\rho = 0.3$, and $\boldsymbol{\sigma}^2_u = (0.5, 0.2)$, the MGQL estimate of $\boldsymbol{\beta}$ was $\hat{\boldsymbol{\beta}} = (-0.0007, -0.0036)'$ with standard errors $s_{\hat{\beta}} = (0.0370, 0.0419)'$, whereas, the MGMM estimate of σ^2_γ was 0.468 with standard error 0.130. Furthermore, when $\boldsymbol{\beta} = (0.3, 0.1)'$, $\sigma^2_\gamma = 0.25$, $\rho = 0.3$, and $\boldsymbol{\sigma}^2_u = (0.1, 0.3)$, the MGQL estimate of $\boldsymbol{\beta}$ was $\hat{\boldsymbol{\beta}} = (0.280, 0.083)'$ with standard errors $s_{\hat{\beta}} = (0.0448, 0.0429)'$, while the MGMM estimate of σ^2_γ was 0.245 with standard error 0.102.

In Table 3, we show the results obtained when all 3 parameters, namely, $\boldsymbol{\beta}$, σ^2_γ, and ρ, are estimated for $(\phi_1, \phi_2) = (1, 1)$ and $(\phi_1, \phi_2) = (0.25, 0.5)$. We see that when $\boldsymbol{\beta} = (0, 0)'$, increasing the bias in the covariates does not appear to affect the performance of the estimates. As an example, when $\boldsymbol{\sigma}^2_u = (0.3, 0.3)$, $\boldsymbol{\beta} = (0, 0)'$, $\sigma^2_\gamma = 0.25$ and $\rho = 0.8$, the MGQL estimate was $\hat{\boldsymbol{\beta}} = (0.0001, -0.0008)'$ with standard errors $s_{\hat{\beta}} = (0.024, 0.024)'$, while, the MGMM estimates of σ^2_γ was 0.245 and of ρ was 0.795 with standard errors 0.113 and 0.040, respectively. Also, when $\boldsymbol{\sigma}^2_u = (0.5, 0.2)$, $\boldsymbol{\beta} = (0, 0)'$, $\sigma^2_\gamma = 0.25$ and $\rho = 0.5$, the MGQL estimate was $\hat{\boldsymbol{\beta}} = (-0.0003, -0.0014)'$ with standard errors $s_{\hat{\beta}} = (0.036, 0.038)'$, while, the MGMM estimates of σ^2_γ was 0.204 and of ρ was 0.522 with standard errors 0.106 and 0.068, respectively. On the contrary, the results in Table 2 appear to suggest that when $\boldsymbol{\beta} \neq (0, 0)'$ the bias in the estimates will increase as the magnitude of the error in the covariates increases. For instance, if $\boldsymbol{\beta} = (0.3, 0.1)'$, $\sigma^2_\gamma = 0.25$ and $\rho = 0.3$, but $\boldsymbol{\sigma}^2_u$ is increased from $(0.1, 0.3)$ to $(0.5, 0.2)$, the MGQL estimate of β_1 ($\hat{\beta}_1 = 0.218$) and the MGMM estimate of σ^2_γ ($\hat{\sigma}^2_\gamma = 0.260$) become more biased. The increase in bias as the magnitude of the error in the covariates increases becomes even more evident when the results in Table 4 are compared to the results in Table 5. Recall that the magnitude of the error in the covariates based on $(\phi_1, \phi_2) = (1, 1)$ used in computing the results in Table 4 is less than that based on $(\phi_1, \phi_2) = (0.25, 0.5)$ used for Table 5. As an example, when $\boldsymbol{\sigma}^2_u = (0.1, 0.3)$,

Table 3: Simulated means (SM) and standard errors (SSE) of modified GQL and GMM estimates of $\boldsymbol{\beta}$, σ^2_γ and ρ respectively, from 1000 simulations, for selected values of measurement error variances $\boldsymbol{\sigma}^2_u = (\sigma^2_1, \sigma^2_2)'$ and correlations ϕ_1, ϕ_2.

			$\phi_1 = \phi_2 = 1$				
$\boldsymbol{\beta}'$	$\boldsymbol{\sigma}^2_u$	σ^2_γ	ρ		$\hat{\boldsymbol{\beta}}'$	$\hat{\sigma}^2_\gamma$	$\hat{\rho}$
(0,0)	(0.3,0.3)	0.25	0.8	SM	(-0.0001,0.0009)	0.255	0.794
				SSE	(0.027,0.028)	(0.116)	(0.037)
		0.5	0.8	SM	(-0.0005,-0.0012)	0.458	0.805
				SSE	(0.024,0.022)	(0.139)	(0.046)
		0.75	0.8	SM	(-0.0002,-0.0001)	0.664	0.809
				SSE	(0.023,0.026)	(0.172)	(0.053)
	(0.5,0.2)	0.25	0.3	SM	(0.0016,-0.005)	0.182	0.336
				SSE	(0.054,0.047)	(0.108)	(0.087)
			0.5	SM	(-0.0011,-0.0029)	0.195	0.518
				SSE	(0.041,0.040)	(0.113)	(0.069)
			0.8	SM	(0.0000,0.0009)	0.245	0.790
				SSE	(0.025,0.027)	(0.120)	(0.038)
			$\phi_1 = 0.25, \phi_2 = 0.5$				
$\boldsymbol{\beta}'$	$\boldsymbol{\sigma}^2_u$	σ^2_γ	ρ		$\hat{\boldsymbol{\beta}}'$	$\hat{\sigma}^2_\gamma$	$\hat{\rho}$
(0,0)	(0.3,0.3)	0.25	0.8	SM	(0.0001,-0.0008)	0.245	0.795
				SSE	(0.024,0.024)	(0.113)	(0.040)
			0.995	SM	(0.0002,0.0004)	0.269	0.994
				SSE	(0.009,0.009)	(0.112)	(0.005)
		0.5	0.8	SM	(0.0001,-0.0005)	0.451	0.806
				SSE	(0.023,0.022)	(0.135)	(0.045)
	(0.5,0.2)	0.25	0.3	SM	(-0.0016,-0.0017)	0.179	0.346
				SSE	(0.040,0.051)	(0.102)	(0.085)
			0.5	SM	(-0.0003,-0.0014)	0.204	0.522
				SSE	(0.036,0.038)	(0.106)	(0.068)
			0.8	SM	(0.0011,-0.0011)	0.242	0.795
				SSE	(0.022,0.029)	(0.114)	(0.039)

Table 4: Simulated means (SM) and Mean square errors (MSE) of naive and bias corrected GQL and GMM estimates of $\boldsymbol{\beta}$, σ_γ^2 and ρ respectively, from 1000 simulations, for $\boldsymbol{\beta} = (0.3, 0.1)$ and selected values of measurement error variances $\boldsymbol{\sigma}_u^2 = (\sigma_1^2, \sigma_2^2)'$ and correlations $\phi_1 = \phi_2 = 1$.

				Naive Estimates			Bias Corrected Estimates		
$\boldsymbol{\sigma}_u^2$	σ_γ^2	ρ		$\hat{\boldsymbol{\beta}}'$	$\hat{\sigma}_\gamma^2$	$\hat{\rho}$	$\hat{\boldsymbol{\beta}}'$	$\hat{\sigma}_\gamma^2$	$\hat{\rho}$
(0.1,0.3)	0.25	0.302	SM	(0.299,0.093)	0.137	0.371	(0.304,0.111)	0.253	0.283
			MSE	(.0019,.0014)	(0.022)	(0.017)	(0.0019,0.0019)	(0.016)	(0.011)
	0.5	0.348	SM	(0.299,0.098)	0.391	0.400	(0.301,0.109)	0.481	0.319
			MSE	(0.0022,0.0015)	(0.030)	(0.012)	(0.0020,0.0016)	(0.029)	(0.015)
	0.75	0.261	SM	(0.295,0.096)	0.540	0.374	(0.301,0.105)	0.683	0.252
			MSE	(0.0017,0.0012)	(0.079)	(0.025)	(0.0019,0.0014)	(0.047)	(0.016)
(0.3,0.3)	0.25	0.200	SM	(0.277,0.083)	0.045	0.356	(0.324,0.110)	0.236	0.190
			MSE	(0.0021,0.0024)	(0.046)	(0.030)	(0.0037,0.0020)	(0.019)	(0.012)
	0.5	0.292	SM	(0.263,0.077)	0.173	0.479	(0.310,0.109)	0.460	0.269
			MSE	(0.0028,0.0023)	(0.123)	(0.040)	(0.0025,0.0015)	(0.016)	(0.018)
	0.75	0.305	SM	(0.299,0.092)	0.715	0.328	(0.306,0.106)	0.675	0.268
			MSE	(0.0017,0.0017)	(0.030)	(0.015)	(0.0018,0.0016)	(0.056)	(0.020)
(0.5,0.2)	0.25	0.238	SM	(0.304,0.096)	0.199	0.277	(0.344,0.110)	0.255	0.198
			MSE	(0.0022,0.0021)	(0.011)	(0.009)	(0.0044,0.0020)	(0.024)	(0.014)
	0.5	0.192	SM	(0.279,0.095)	0.235	0.377	(0.318,0.107)	0.426	0.197
			MSE	(0.0022,0.0020)	(0.088)	(0.042)	(0.0037,0.0016)	(0.041)	(0.015)
	0.75	0.286	SM	(0.296,0.098)	0.806	0.271	(0.313,0.106)	0.647	0.251
			MSE	(0.0017,0.0017)	(0.031)	(0.016)	(0.0024,0.0015)	(0.064)	(0.021)

Table 5: Simulated means (SM) and Mean square errors (MSE) of naive and bias corrected GQL and GMM estimates of $\boldsymbol{\beta}$, σ^2_γ and ρ respectively, from 1000 simulations, for $\boldsymbol{\beta} = (0.3, 0.1)$ and selected values of measurement error variances $\boldsymbol{\sigma}^2_u = (\sigma^2_1, \sigma^2_2)'$ and correlations $\phi_1 = 0.25$, $\phi_2 = 0.5$.

				Naive Estimates			Bias Corrected Estimates		
$\boldsymbol{\sigma}^2_u$	σ^2_γ	ρ		$\hat{\boldsymbol{\beta}}'$	$\hat{\sigma}^2_\gamma$	$\hat{\rho}$	$\hat{\boldsymbol{\beta}}'$	$\hat{\sigma}^2_\gamma$	$\hat{\rho}$
(0.1,0.3)	0.25	0.285	SM	(0.279,0.086)	0.208	0.313	(0.281,0.086)	0.202	0.312
			MSE	(0.0027,0.0016)	(0.011)	(0.007)	(0.0023,0.0020)	(0.015)	(0.009)
	0.5	0.250	SM	(0.279,0.089)	0.387	0.330	(0.283,0.087)	0.411	0.303
			MSE	(0.0028,0.0020)	(0.031)	(0.018)	(0.0020, 0.0018)	(0.028)	(0.0140
	0.75	0.231	SM	(0.266,0.087)	0.572	0.317	(0.281,0.083)	0.627	0.286
			MSE	(0.0031,0.0016)	(0.061)	(0.022)	(0.0021,0.0020)	(0.048)	(0.018)
(0.3,0.3)	0.25	0.301	SM	(0.230,0.082)	0.303	0.252	(0.246,0.092)	0.214	0.312
			MSE	(0.0066,0.0021)	(0.012)	(0.011)	(0.0049,0.0011)	(0.0170	(0.010)
	0.5	0.278	SM	(0.246,0.078)	0.458	0.284	(0.250,0.095)	0.424	0.299
			MSE	(0.0044,0.0025)	(0.018)	(0.011)	(0.0042,0.0016)	(0.029)	(0.013)
	0.75	0.258	SM	(0.250,0.083)	0.656	0.309	(0.244,0.091)	0.633	0.281
			MSE	(0.0041,0.0021)	(0.040)	(0.017)	(0.0050,0.0015)	(0.049)	(0.017)
(0.5,0.2)	0.25	0.295	SM	(0.205,0.083)	0.304	0.254	(0.220,0.092)	0.210	0.299
			MSE	(0.0105,0.0018)	(0.011)	(0.010)	(0.0082,0.0023)	(0.017)	(0.010)
	0.5	0.262	SM	(0.209,0.083)	0.392	0.398	(0.228,0.101)	0.424	0.277
			MSE	(0.0098,0.0018)	(0.029)	(0.030)	(0.0068,0.0018)	(0.029)	(0.013)
	0.75	0.185	SM	(0.217,0.090)	0.612	0.244	(0.219,0.093)	0.626	0.224
			MSE	(0.0084,0.0015)	(0.049)	(0.018)	(0.0081,0.0017)	(0.047)	(0.016)

the absolute value of the bias in the bias corrected estimates for $\boldsymbol{\beta} = (0.3, 0.1)'$ and $\sigma_\gamma^2 = 0.25$, in Table 4 was (0.004,0.011). In Table 5, the absolute bias increases to (0.019,0.014). When σ_γ^2 increases to 0.5 the absolute bias in Table 4 was (0.001,0.009) whereas the absolute bias in Table 5 was (0.017,0.013). The mean squared errors in Tables 4 and 5 for the same parameters were however about the same in magnitude.

Furthermore, when compared to the bias corrected estimates in Table 4, the estimates obtained from the naive estimating equations appear to be unstable in the sense that it seems to perform well in some cases and then suddenly performs badly in other cases. This makes the naive estimates to be unreliable. For instance, when $\boldsymbol{\sigma}_u^2 = (0.3, 0.3)'$, $\sigma_\gamma^2 = 0.75$ and $\rho = 0.305$, the naive estimates $\hat{\boldsymbol{\beta}} = (0.299, 0.092)'$, $\hat{\sigma}_\gamma^2 = 0.715$ and $\hat{\rho} = 0.328$, respectively, appear to perform well. The bias corrected estimates, $\hat{\boldsymbol{\beta}} = (0.306, 0.106)'$, $\hat{\sigma}_\gamma^2 = 0.675$ and $\hat{\rho} = 0.268$, respectively, also performed well. However, for the same value of $\boldsymbol{\sigma}_u^2$, $\sigma_\gamma^2 = 0.5$ and $\rho = 0.292$, the naive estimates $\hat{\boldsymbol{\beta}} = (0.263, 0.077)'$, $\hat{\sigma}_\gamma^2 = 0.173$ and $\hat{\rho} = 0.479$, respectively, do not perform well. The estimates are also highly biased when compared to the bias corrected estimates in Table 4. The bias corrected estimates in this case, $\hat{\boldsymbol{\beta}} = (0.310, 0.109)'$, $\hat{\sigma}_\gamma^2 = 0.460$ and $\hat{\rho} = 0.269$, respectively, continue to perform well. Similarly, for the same value of $\boldsymbol{\sigma}_u^2$, with $\sigma_\gamma^2 = 0.25$ and $\rho = 0.200$, the naive estimates $\hat{\boldsymbol{\beta}} = (0.277, 0.083)'$, $\hat{\sigma}_\gamma^2 = 0.045$ and $\hat{\rho} = 0.356$, respectively, do not perform well and again highly biased when compared to the bias corrected estimates. The same unstable pattern can be seen in the naive estimates when $\boldsymbol{\sigma}_u^2 = (0.5, 0.2)'$. In all of these cases, the mean squared errors of both the naive and bias corrected estimates were about the same in magnitude. Also, the results in Tables 4 and 5 show that the bias corrected estimates will in general, perform better than the naive estimates as the bias in the covariates is increased by setting $\phi_1 = 0.25$ and $\phi_2 = 0.5$. For instance, when $\boldsymbol{\sigma}_u^2 = (0.1, 0.3)$, $\boldsymbol{\beta} = (0.3, 0.1)'$, $\sigma_\gamma^2 = 0.5$ and $\rho = 0.250$, the naive estimates were $\hat{\boldsymbol{\beta}} = (0.279, 0.089)'$, $\hat{\sigma}_\gamma^2 = 0.387$ and $\hat{\rho} = 0.330$, respectively; whereas, the bias corrected estimates were $\hat{\boldsymbol{\beta}} = (0.283, 0.087)'$, $\hat{\sigma}_\gamma^2 = 0.411$ and $\hat{\rho} = 0.303$. Also, when $\boldsymbol{\sigma}_u^2 = (0.5, 0.2)$, $\boldsymbol{\beta} = (0.3, 0.1)'$, $\sigma_\gamma^2 = 0.5$ and $\rho = 0.262$, the naive estimates were $\hat{\boldsymbol{\beta}} = (0.209, 0.083)'$, $\hat{\sigma}_\gamma^2 = 0.392$ and $\hat{\rho} = 0.398$, respectively; whereas, the bias corrected estimates were $\hat{\boldsymbol{\beta}} = (0.228, 0.101)'$, $\hat{\sigma}_\gamma^2 = 0.424$ and $\hat{\rho} = 0.277$.

Overall, the values of the estimates based on our proposed bias corrected methods, where close to their true values. In some cases, the bias of σ_γ^2 and ρ increased as the magnitude of the measurement error variances $\boldsymbol{\sigma}_u^2$ and the variance of the random effect σ_γ^2 increased. The results also show that the bias corrected estimates will be more reliable and stable when compared to the naive estimates as shown in Tables 4 and 5. In addition, the results appear to demonstrate that if the magnitude of the error in x_{it} is too large, then the observed values of the covariates will be dominated by the error and the estimates obtained based on these covariates will be highly biased and unreliable, in particular $\hat{\sigma}_\gamma^2$ and $\hat{\rho}$.

In order to assess the impact of having a large or small number of subjects K on the accuracy of the estimates, we used the modified estimating equations to compare the estimates when $K = 25$ and 300 with $\phi_1 = 0.25$, $\phi_2 = 0.5$. The results are shown in Table 6. The estimates were computed for $K = 300$, $T = 4$ due to the time it takes to complete 1000 simulations with $K = 1000$, $T = 4$. We found that when $K = 300$, $T = 4$, $\boldsymbol{\sigma}_u^2 = (0.1, 0.3)$, $\boldsymbol{\beta} = (0.3, 0.1)'$, $\sigma_\gamma^2 = 0.25$

Table 6: Simulated means (SM), Absolute value of bias (AVB) and standard errors (SSE) of naive and bias corrected GQL and GMM estimates of $\boldsymbol{\beta}$, σ_γ^2 and ρ respectively, from 1000 simulations, for $\boldsymbol{\beta} = (0.3, 0.1)$ and selected values of measurement error variances $\boldsymbol{\sigma}_u^2 = (\sigma_1^2, \sigma_2^2)'$ and correlations $\phi_1 = 0.25$, $\phi_2 = 0.5$.

				K = 25			**K = 300**		
$\boldsymbol{\sigma}_u^2$	σ_γ^2	ρ		$\hat{\boldsymbol{\beta}}'$	$\hat{\sigma}_\gamma^2$	$\hat{\rho}$	$\hat{\boldsymbol{\beta}}'$	$\hat{\sigma}_\gamma^2$	$\hat{\rho}$
(0.1,0.3)	0.25	0.190	SM	(0.270,0.068)	0.195	0.236	(0.284,0.092)	0.204	0.225
			AVB	(0.030,0.032)	(0.055)	(0.046)	(0.016,0.008)	(0.046)	(0.035)
			MSE	(0.0086,0.0154)	(0.022)	(0.019)	(0.0009,0.0007)	(0.006)	(0.002)
	0.75	0.192	SM	(0.283,0.099)	0.526	0.269	(0.281,0.093)	0.604	0.279
			AVB	(0.017,0.001)	(0.224)	(0.077)	(0.019,0.007)	(0.146)	(0.087)
			MSE	(0.0057,0.0070)	(0.122)	(0.027)	(0.0010,0.0005)	(0.032)	(0.015)
(0.3,0.3)	0.5	0.278	SM	(0.243,0.085)	0.408	0.295	(0.247,0.090)	0.442	0.288
			AVB	(0.057,0.015)	(0.092)	(0.017)	(0.053,0.010)	(0.058)	(0.010)
			MSE	(0.0119,0.0091)	(0.056)	(0.024)	(0.0034,0.0006)	(0.009)	(0.005)
	0.75	0.213	SM	(0.233,0.106)	0.511	0.316	(0.254,0.094)	0.634	0.245
			AVB	(0.067,0.006)	(0.239)	(0.103)	(0.046,0.006)	(0.116)	(0.032)
			MSE	(0.0113,0.0064)	(0.132)	(0.037)	(0.0026,0.0006)	(0.025)	(0.009)
(0.5,0.2)	0.25	0.191	SM	(0.226,0.079)	0.198	0.234	(0.213,0.091)	0.332	0.131
			AVB	(0.074,0.021.)	(0.052)	(0.043)	(0.087,0.009)	(0.082)	(0.060)
			MSE	(0.0169,0.0091)	(0.024)	(0.020)	(0.0080,0.0008)	(0.010)	(0.007)
	0.5	0.262	SM	(0.217,0.101)	0.389	0.294	(0.222,0.094)	0.470	0.266
			AVB	(0.083,0.001)	(0.127)	(0.032)	(0.078,0.006)	(0.030)	(0.004)
			MSE	(0.0117,0.0072)	(0.0572)	(0.0242)	(0.0066,0.0007)	(0.0074)	(0.0053)

and $\rho = 0.190$, the estimates were $\hat{\boldsymbol{\beta}} = (0.284, 0.092)'$, $\hat{\sigma}_\gamma^2 = 0.204$ and $\hat{\rho} = 0.225$ with mean square errors, $(0.0009, 0.0007)'$, 0.006 and 0.002, respectively. Also, with $\boldsymbol{\sigma}_u^2 = (0.3, 0.3)$, $\boldsymbol{\beta} = (0.3, 0.1)'$, $\sigma_\gamma^2 = 0.5$ and $\rho = 0.278$, the estimates were $\hat{\boldsymbol{\beta}} = (0.247, 0.090)'$, $\hat{\sigma}_\gamma^2 = 0.442$ and $\hat{\rho} = 0.288$ with mean square errors $(0.0034, 0.0006)'$, 0.009 and 0.005, respectively. Whereas, with $K = 25$, $T = 4$, $\boldsymbol{\sigma}_u^2 = (0.1, 0.3)$, $\boldsymbol{\beta} = (0.3, 0.1)'$, $\sigma_\gamma^2 = 0.25$ and $\rho = 0.190$ we obtained, $\hat{\boldsymbol{\beta}} = (0.270, 0.068)'$, $\hat{\sigma}_\gamma^2 = 0.195$ and $\hat{\rho} = 0.236$ with mean square errors $(0.0086, 0.0154)'$, 0.022 and 0.019, respectively. When $K = 25$, $T = 4$, $\boldsymbol{\sigma}_u^2 = (0.3, 0.3)$, $\boldsymbol{\beta} = (0.3, 0.1)'$, $\sigma_\gamma^2 = 0.5$ and $\rho = 0.278$, we found that $\hat{\boldsymbol{\beta}} = (0.243, 0.085)'$, $\hat{\sigma}_\gamma^2 = 0.408$ and $\hat{\rho} = 0.295$ with mean square errors $(0.0119, 0.0091)'$, 0.056 and 0.024, respectively. These results show clearly that the estimates for $K = 300$ are more efficient than the estimates obtained with $K = 25$, in the sense that the mean square errors were much smaller in magnitude. In addition, though the magnitude of the errors in the covariates was large, the estimates when $K = 300$ were closer to the true values than when $K = 25$. This indicates that the modified estimates performs better as K increases.

5 Application To Real Data

In this section, we use longitudinal count responses on the number of patents awarded to 168 firms in the United States from 1974 to 1979 along with associated covariate information on the type of firm, log of the book value of capital in 1972 and research and development (R & D) expenditures from 1971 to 1979, to demonstrate how the methods we have proposed in this paper can be applied to a real data. The patent count data has also been analyzed by Haussman, Hall and Griliches (1984), and Sutradhar (2011). In the data, the code used to identify firms that were considered to be non-scientific was 0 while scientific firms were coded as 1. Aside from the covariates, the patent counts may also be influenced by some unobservable random effects γ_i such as internal or bureaucratic processes that may lead to delays in registration of patents in a given year. A close examination of the data shows that measurement error could be an issue in the data as some firms may have reported a lower or higher R & D expenditures. Furthermore, since the repeated patent counts of a firm are likely to be longitudinally correlated and the R & D expenditures are changing with time,

Following the notations in this paper, for this data we have $K = 168$, $T = 6$ with the year 1974 corresponding to $t = 1$. As far as covariates are concerned, we considered $p = 6$ covariates, $\boldsymbol{x}_{it} = (x_{it1}, x_{it2}, x_{it3}, x_{it4}, x_{it5}, x_{it6})'$, where the first four covariates x_{it1} to x_{it4} are the R & D expenditures at time points $t, t-1, t-2$ and $t-3$, respectively. The estimated parameters associated with the R & D expenditures will provide information to the analyst on the effect of the expenditures in the current year (lag 0) and in the past 3 years (lags 1, 2 and 3) on the number of patents awarded to firms. The fifth (x_{it5}) and sixth covariates (x_{it6}) are type of firm and log book value of capital, respectively. For instance, for Firm 1 corresponding to $i = 1$, the $T \times p$ (6×6) matrix of covariates

$\boldsymbol{X}_1$ and vector of observed patent counts $\boldsymbol{y}_1$, are given by

$$\boldsymbol{X}_1 = \begin{pmatrix} -0.685 & -0.151 & 0.084 & -0.216 & 1 & 1.975 \\ -1.485 & -0.685 & -0.151 & 0.084 & 1 & 1.975 \\ -1.195 & -1.485 & -0.685 & -0.151 & 1 & 1.975 \\ -0.610 & -1.195 & -1.485 & -0.685 & 1 & 1.975 \\ -0.581 & -0.610 & -1.195 & -1.485 & 1 & 1.975 \\ -0.609 & -0.581 & -0.610 & -1.195 & 1 & 1.975 \end{pmatrix} \text{ and } \boldsymbol{y}_1 = \begin{pmatrix} 2 \\ 3 \\ 2 \\ 1 \\ 1 \\ 1 \end{pmatrix},$$

respectively. It is clear that since the type of firm and log book value of capital are not changing with time, the error variances $\hat{\sigma}_5^2 = \hat{\sigma}_6^2 = 0$ and correlation parameters for these covariates $\hat{\phi}_5 = \hat{\phi}_6 = 0$. Now, using (3.4) in §3.1 to estimate the error variances, σ_u^2 and magnitude of the correlation parameters $\phi_u, u = 1, 2, \ldots, 6$, we find that $\hat{\sigma}_1^2 = 0.0855$, $\hat{\sigma}_2^2 = 0.0807$, $\hat{\sigma}_3^2 = 0.0921$, $\hat{\sigma}_4^2 = 0.0996$, $\hat{\sigma}_5^2 = \hat{\sigma}_6^2 = 0$, and $\hat{\phi}_u = 0.2$, $u = 1, \ldots, 4$, $\hat{\phi}_u = 0$, $u = 5, 6$. These results show that the magnitude of the measurement errors in the patent count data may not be large enough to mask the effect of the covariates. Next, using the iterative procedures for the modified GQL and the modified GMM approaches of §3.1-3.3 we obtained estimates of the covariate effects, variance and correlation index parameters as $\hat{\boldsymbol{\beta}} = (0.4394, -0.1058, -0.0033, 0.0006, 0.2672, 0.2104)$, $\hat{\sigma}_\gamma^2 = 0.3598$, and $\hat{\rho} = 0.0671$, respectively. The estimated parameters show that firm type, log book value of capital in 1972 and R & D expenditure at time t (lag 0) had a large positive effect on patent counts with scientific firms having more patent count awards than non-scientific firms. The effect of the most recent (lag 0) R & D expenditures on patent counts appear to be large (0.4394), while the effect of the expenditures in the previous year (lag 1) and two years earlier (lag 2) were moderate and negative (-0.1058 and -0.0033, respectively). The estimated parameters also show that the effect of the lag 3 R & D expenditures is positive but appear to be very minimal.

In order to assess the performance of the model in fitting the data, we compared the yearly mean, $\bar{y}_t = \sum_{i=1}^{K} y_{it}/K$ and standard deviation $S_t = \sqrt{\sum_{i=1}^{K}(y_{it} - \bar{y}_t)^2/(K-1)}$, $t = 1, \ldots, 6$ of the patent counts with the overall yearly mean and standard deviations of the fitted values computed based on (2.8) and (2.10). The results, in Table 7, show that the fitted values are quite close to the true yearly means and standard deviations of the patent counts. For instance, in 1974, 1976, and 1978, the true means were 2.952, 2.369 and 2.399, respectively with corresponding fitted values of 2.314, 2.317 and 2.427.

6 Concluding Remarks

Previous research on the effect of measurement error in covariates on model parameters have focused largely on Gaussian data. The problem for count and binary data has, however, not been adequately addressed in the literature. In particular, to our knowledge, no research is currently available in the literature when the observed longitudinal counts are influenced by unobservable latent variables and covariates that are measured with error. In this paper we have outlined some results to bridge that

Table 7: Yearly mean and standard deviation (SD) of patent counts and yearly fitted values based on GQL and GMM estimates of $\boldsymbol{\beta}$, σ_γ^2 and ρ.

	Patent Count					
	1974	1975	1976	1977	1978	1979
Mean	2.952	2.435	2.369	2.244	2.399	2.161
SD	3.508	2.952	2.704	2.851	3.094	2.854
	Fitted Values					
	1974	1975	1976	1977	1978	1979
Mean	2.314	2.238	2.317	2.359	2.427	2.441
SD	2.131	2.078	2.132	2.161	2.207	2.217

gap. We proposed a dynamic conditionally Poisson mixed model for the data and developed estimating equations and iterative methods for computing unbiased and consistent estimates of the model parameters, namely, the effect of the covariates, the variance of the latent variable and the correlation index parameter. We also developed an R code for implementing the techniques and used it in our simulation studies and application to a real data. The results of our simulation studies showed that the proposed computational methods performed well when the magnitude of the measurement error is not so large as to dominate or mask the effect of the true covariates. In some longitudinal studies, the number of observations for each subject may be unequal. Also, the observations may be measured at unequally spaced intervals. We note that the approach proposed in this paper can be extended to these situations.

Acknowledgement

The author is grateful to the editors and referees whose comments have led to improvements in the quality of this paper. This research is partially supported by the Natural Sciences and Engineering Research Council of Canada.

Appendix

Proof of Theorem 1.

(a) It is clear that the integrand in the expression $\int [h_1^*(F_\epsilon) - h_2^*(F_\epsilon)] dF_\epsilon(\boldsymbol{y})$, is a continuous function of ϵ. Therefore, using Leibniz rule (see Flanders (1973)) and (3.13), it can be shown

$$\left\{ \int \left[\frac{\partial h_1^*(F_\epsilon)}{\partial \boldsymbol{\beta}(F_\epsilon)} - \frac{\partial h_2^*(F_\epsilon)}{\partial \boldsymbol{\beta}(F_\epsilon)} \right] dF \cdot \frac{\partial \boldsymbol{\beta}(F_\epsilon)}{\partial \epsilon} \right\} \Bigg|_{\epsilon=0} + \int \left[h_1^*(F) - h_2^*(F) \right] dG = \mathbf{0}. \qquad \text{(A.1)}$$

The influence function in (3.14) is obtained at $G = \delta_{\boldsymbol{y}}$. In (3.14),

$$\begin{aligned}
\frac{\partial h_1^*(F)}{\partial \boldsymbol{\beta}} &= \left.\frac{\partial h_1^*(F_\epsilon)}{\partial \boldsymbol{\beta}(F_\epsilon)}\right|_{\epsilon=0} \\
&= \frac{\partial \boldsymbol{M}_{1\phi}(F)}{\partial \boldsymbol{\beta}}[\boldsymbol{X}' - \boldsymbol{B}_{1\phi}\{\boldsymbol{\beta}(F)\otimes \mathbf{1}_T'\}]\sum_{k=1}^{T}\boldsymbol{J}_k\boldsymbol{U}_x^{1/2}(F)\boldsymbol{Q}_z\boldsymbol{U}_x^{-1/2}(F)\boldsymbol{D}_{z,k}^{-1/2}(F)\boldsymbol{y} \\
&\quad - \boldsymbol{M}_{1\phi}(F)\boldsymbol{B}_{1\phi}\frac{\partial\{\boldsymbol{\beta}(F)\otimes \mathbf{1}_T'\}}{\partial \boldsymbol{\beta}}\sum_{k=1}^{T}\boldsymbol{J}_k\boldsymbol{U}_x^{1/2}(F)\boldsymbol{Q}_z\boldsymbol{U}_x^{-1/2}(F)\boldsymbol{D}_{z,k}^{-1/2}(F)\boldsymbol{y} \\
&\quad + \boldsymbol{M}_{1\phi}(F)[\boldsymbol{X}' - \boldsymbol{B}_{1\phi}\{\boldsymbol{\beta}(F)\otimes \mathbf{1}_T'\}]\sum_{k=1}^{T}\boldsymbol{J}_k\left\{\frac{\partial \boldsymbol{U}_x^{1/2}(F)}{\partial \boldsymbol{\beta}}\boldsymbol{Q}_z\boldsymbol{U}_x^{-1/2}(F)\boldsymbol{D}_{z,k}^{-1/2}(F)\right. \\
&\quad + \boldsymbol{U}_x^{1/2}(F)\boldsymbol{Q}_z\frac{\partial \boldsymbol{U}_x^{-1/2}(F)}{\partial \boldsymbol{\beta}}\boldsymbol{D}_{z,k}^{-1/2}(F) \\
&\quad \left. + \boldsymbol{U}_x^{1/2}(F)\boldsymbol{Q}_z\boldsymbol{U}_x^{-1/2}(F)\frac{\partial \boldsymbol{D}_{z,k}^{-1/2}(F)}{\partial \boldsymbol{\beta}}\right\}\boldsymbol{y}, \qquad \text{(A.2)}
\end{aligned}$$

where

$$\frac{\partial \boldsymbol{M}_{1\phi}(F)}{\partial \boldsymbol{\beta}} : \;T\times T\times p;\quad \frac{\partial\{\boldsymbol{\beta}(F)\otimes \mathbf{1}_T'\}}{\partial \boldsymbol{\beta}} : \; p\times T\times p;$$
$$\frac{\partial \boldsymbol{U}_x^{1/2}(F)}{\partial \boldsymbol{\beta}} : \; T\times T\times p;\ \text{ and }\ \frac{\partial \boldsymbol{D}_{z,k}^{-1/2}(F)}{\partial \boldsymbol{\beta}} : \; T\times T\times p,$$

are 3-dimensional arrays (tensors). The components of these tensors are defined in the following way.

For $r = 1, \ldots, T$, $s = 1, \ldots, T$ and $l = 1, \ldots, p$, the (r, s, l)th element of $\frac{\partial \boldsymbol{M}_{1\phi}(F)}{\partial \boldsymbol{\beta}}$ is defined as $M_{r,s,l}^{(1)} = \frac{\partial \boldsymbol{M}_{1\phi(r,s)}(F)}{\partial \beta_l}$, with

$$M_{r,r}^{(1)} = \frac{\partial \boldsymbol{M}_{1\phi(r,r)}(F)}{\partial \boldsymbol{\beta}} = -m_1\boldsymbol{B}_{1\phi}\boldsymbol{\beta} : \; p\times 1,$$

when $r = s$ and $M_{r,s}^{(1)} = \mathbf{0} : \; p\times 1$, when $r \neq s$. Similarly, For $r = 1, \ldots, p$, $s = 1, \ldots, T$ and $l = 1, \ldots, p$, we have

$$\beta_{r,s}^{(1)} = \frac{\partial\{\boldsymbol{\beta}(F)\otimes \mathbf{1}_T'\}_{(r,r)}}{\partial \boldsymbol{\beta}} = \boldsymbol{j} = (0, 0, \ldots, 0, 1, 0, \ldots, 0)',$$

where 1 is the rth element of the $p\times 1$ vector $\boldsymbol{j}$ and zeros elsewhere. Also, for $r = 1, \ldots, T$, $s = 1, \ldots, T$ and $l = 1, \ldots, p$, we have

$$U_{r,r}^{(1)} = \frac{\partial \boldsymbol{U}_{x(r,r)}^{1/2}(F)}{\partial \boldsymbol{\beta}} = \mu_{x,r}^{1/2}\boldsymbol{x}_r/2 : \; p\times 1,$$

when $r = s$ and $U^{(1)}_{r,s} = \mathbf{0} :\ p \times 1$, when $r \neq s$. Next, we have that

$$U^{(2)}_{r,r} = \frac{\partial \boldsymbol{U}^{-1/2}_{x(r,r)}(F)}{\partial \boldsymbol{\beta}} = -\mu^{-1/2}_{x,r}\boldsymbol{x}_{ir}/2 :\ p \times 1,$$

when $r = s$ and $U^{(2)}_{r,s} = \mathbf{0} :\ p \times 1$, when $r \neq s$. Finally, we have that for $r = 1, \ldots, T$, $s = 1, \ldots, T$ and $l = 1, \ldots, p$,

$$D^{(1)}_{r,r} = \frac{\partial \boldsymbol{D}^{-1/2}_{z,k(r,r)}(F)}{\partial \boldsymbol{\beta}} = \frac{\partial \hat{d}_{z,kr}}{\partial \boldsymbol{\beta}} :\ p \times 1,$$

when $r = s$ and $D^{(2)}_{r,s} = \mathbf{0} :\ p \times 1$, when $r \neq s$, where

$$\frac{\partial \hat{d}_{z,kr}}{\partial \boldsymbol{\beta}} = \mu_{x,r}[\boldsymbol{x}_r - \Lambda\boldsymbol{\beta}]\exp(\boldsymbol{\beta}'\Lambda\boldsymbol{\beta})[\exp(\sigma^2_\gamma - 1)] + \ \mu_{x,k}[\boldsymbol{x}_k - \Lambda\boldsymbol{\beta}]\exp(\boldsymbol{\beta}'\Lambda\boldsymbol{\beta})[\exp(\sigma^2_\gamma - 1)]$$
$$+ \ \mu_{x,r}\mu_{x,k}[\boldsymbol{x}_r + \boldsymbol{x}_k - 2(\Lambda + \Lambda_\phi)\boldsymbol{\beta}]\exp(\boldsymbol{\beta}'(\Lambda + \Lambda_\phi)\boldsymbol{\beta})[\exp(\sigma^2_\gamma - 1)].$$

The partial derivative $\frac{\partial h^*_2(F)}{\partial \boldsymbol{\beta}}$ is similarly defined.

(b) We begin by taking the first order Taylor series expansion of $\boldsymbol{\beta}(F_\epsilon)$ with respect to ϵ. In particular, at $\epsilon = 1$ we obtain

$$\boldsymbol{\beta}(G) - \boldsymbol{\beta}(F) \approx \left.\frac{\partial \beta(F_\epsilon)}{\partial \epsilon}\right|_{\epsilon=0}. \tag{A.3}$$

Using (A.1), we note that at $G = F_k$ the expression (A.3) becomes

$$\boldsymbol{\beta}(G) - \boldsymbol{\beta}(F) \approx \frac{1}{K}\sum_{i=1}^{K} IF(\boldsymbol{y}_i, \boldsymbol{\beta}, F)]. \tag{A.4}$$

The consistency of $\boldsymbol{\beta}(F_K)$ then follows from applying Theorem 6.2 of Billingsley (1986).

(c) The asymptotic normality of $\boldsymbol{\beta}(F_K)$ follows from rewriting (A.4) and applying Lindeberg's central limit theorem (Billingsley, 1986, Thm 27.2, Pg 369).

References

Billingsley, P. (1986), Probabilty and measure, New York: John Wiley and Sons Inc.

Buonaccorsi, J. P. (2010), Measurement error: models, methods and applications, London: Chapman and Hall.

Buonaccorsi, J. P., Demidenko, E. and Tosteson, T. (2000), "Estimation in longitudinal random effects models with measurement error," *Statistica Sinica*, 10, 885-903.

Bun, M. J. G. and Carree, M. A. (2005), "Bias-corrected estimation in dynamic panel data models," *Journal of Business and Economic Statistics*, 23, 200-210.

Carroll, R. J., Lin, X. and Wang, N. (1997), "Generalized linear mixed measurement error models," In Modelling Longitudinal and Spatially Correlated Data, New York: Springer.

Cheng, C. and Ness, J. W. V. (1999), Statistical Regression with measurement error, Kendall's Library of Statistics: Arnold, London.

Cook, J. R. and Stefanski, L. A. (1994), "Simulation-extrapolation estimation in parametric measurement error models," *Journal of the American Statistical Association*, 89, 1314-1328.

Devlin, S. J., Gnanadesikan, R. and Kettenring, J. R. (1975), "Robust estimation and outlier detection with correlation coefficients," *Biometrika*, 62, 531-545.

Fan, Z., Sutradhar, B. and Rao, R. P. (2012), "Bias corrected generalized method of moments and generalized quasi-likelihood inferences in linear models for panel data with measurement error," *Sankhya B*, 74, 126-148.

Fuller, W. A. (1996), Introduction to statistical time series. New York: John Wiley and Sons Inc.

Gleser, L. J. (1990), "Improvement of the naïve estimation in nonlinear errors-in-variables regression," In Statistical Analysis of Measurement Error Models and Application (P. J. Brown, W. A. Fuller, eds.), *Cotemporary Mathematics*, 112, 99-114.

Hampel, F. R. (1974), "The influence curve and its role in robust estimation," *Journal of the American Statistical Association*, 69, 383-393.

Hausman, J. A., Hall, B. H. and Griliches, Z. (1984), "Econometric models for count data with an application to the patent-R and D relationship," *Econometrica*, 52, 908-938.

Jowaheer, V., Sutradhar, B. C. and Fan, Z. (2013), "Inferences in binary regression models for independent data with measurement errors in covariates," In Synergies of soft computing and statistics for intelligent data analysis, Springer Berlin Heidelberg.

Kanter, M. (1975), "Autoregression for discrete processes mod 2," *Journal of Applied Probability*, 12, 371-375.

McCullagh, P. and Nelder, J. A. (1989), Generalized linear models, London: Chapman and Hall.

McKenzie, E. (1988), "Some ARMA models for dependent sequences of Poisson counts," *Advances in Applied Probability*, 20, 822-835.

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

Oyet, A. J. and Sutradhar, B. C. (2013), "Longitudinal modelling of infectious disease," *Sankhya B*, 75, 319-342.

Qaqish, B. F. (2003), "A family of multivariate binary distributions for simulating correlated binary variables with specified marginal means and correlation," *Biometrika*, 90, 455-463.

Rabe-Hesketh, S., Pickles, A. and Skrondal, A. (2003), "Correcting for covariate measurement error in logistic regression using nonparametric maximum likelihood estimation," *Statistical Modelling*, 3, 215-232.

Rosner, B. Willett, W. C. and Spiegelmann, D. (1989), "Correction of logistic regression relative risk estimates and confidence intervals for systematic within-person measurement error," *Statistics in Medicine*, 8(9), 1051-1069.

Schafer, D. W. (1993), "Likelihood analysis for probit regression with measurement error," *Biometrika*, 80, 899-904.

Schmid, C. H., Segal, M. R. and Rosner, B. (1994), "Incorporating measurement error in the estimation of autoregressive models for longitudinal data," *Journal of Statistical Planning and Inference*, 42, 1-18.

Schneeweiss, H. and Augustin, T. (2006), "Some recent advances in measurement error models and methods," *Allgemeines Statistisches Archiv*, 90, 183-197.

Selvaratnam, S., Oyet, A., Yi, Y. and Gadag, V. (2017), "Estimation of a generalized linear mixed model for response-adaptive designs in multi-center clinical trials," *The Canadian Journal of Statistics*, 45, 310-325.

Stefanski, L. A. (1985), "Effect of measurement error on parameter estimation," *Biometrika*, 72, 583-592.

Staudenmayer, J. and Buonaccorsi, J.P. (2005), "Measurement error in linear autoregressive models," *Journal of the American Statistical Association*, 100, 841-852.

Sutradhar, B. C. (2011), Dynamic Mixed Models for Familial Longitudinal Data, New York: Springer.

Sutradhar, B.C. and Rao, J.N.K. (1996), "Estimation of regression parameters in generalized linear models for cluster correlated data with measurement error," *Canadian Journal of Statistics*, 24, 177-192.

Sutradhar, B. C. and Rao, R. P. (2016), "Inferences in longitudinal count data models with measurement error in time dependent covariates," *Sankhya B*, 78, 39-65.

Verbeke, G. and Molenberghs, G. (2000), Linear Mixed Models for Longitudinal Data, Springer: New York.

Wang, N., Carroll, R. J. and Liang, K. Y. (1996), "Quasilikelihood estimation in measurement error models with correlated replicates," *Biometrics*, 52, 401-411.

Wansbeek, T. (2001), "GMM estimation in panel data models with measurement error," *Journal of econometrics*, 104, 259-268.

Wasserman, L (2006), All of Nonparametric Statistics, Springer: New York.

Xiao, Z., Shao, J., Xu, R. and Palta, M. (2007), "Efficiency of GMM estimation in panel data models with measurement error," *Sankhya: Indian Journal of Statistics*, 69, 101-118.

Zhang, C. and Oyet, A. J. (2014), "Second order longitudinal dynamic models with covariates: estimation and forecasting," *Metrika*, 77, 837-859.

Received: December 12, 2023

Accepted: April 3, 2024