

Empirical Likelihood Based Longitudinal Data Analysis

Tharshanna Nadarajah¹, Asokan Mulayath Variyath², J Concepción Loredo-Osti²

¹Department of Mathematics and Statistics, St. Francis Xavier University, Antigonish, Canada

²Department of Mathematics and Statistics, Memorial University of Newfoundland, St. John's, Canada

Email: tnadaraj@stfx.ca, varyath@mun.ca, jcloredoosti@mun.ca

How to cite this paper: Nadarajah, T., Variyath, A.M. and Loredo-Osti, J.C. (2020) Empirical Likelihood Based Longitudinal Data Analysis. *Open Journal of Statistics*, 10, 611-639.

<https://doi.org/10.4236/ojs.2020.104037>

Received: June 28, 2020

Accepted: July 10, 2020

Published: July 13, 2020

Copyright © 2020 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

In longitudinal data analysis, our primary interest is in the estimation of regression parameters for the marginal expectations of the longitudinal responses, and the longitudinal correlation parameters are of secondary interest. The joint likelihood function for longitudinal data is challenging, particularly due to correlated responses. Marginal models, such as generalized estimating equations (GEEs), have received much attention based on the assumption of the first two moments of the data and a working correlation structure. The confidence regions and hypothesis tests are constructed based on the asymptotic normality. This approach is sensitive to the misspecification of the variance function and the working correlation structure which may yield inefficient and inconsistent estimates leading to wrong conclusions. To overcome this problem, we propose an empirical likelihood (EL) procedure based on a set of estimating equations for the parameter of interest and discuss its characteristics and asymptotic properties. We also provide an algorithm based on EL principles for the estimation of the regression parameters and the construction of its confidence region. We have applied the proposed method in two case examples.

Keywords

Longitudinal Data, Generalized Estimating Equations, Empirical Likelihood, Adjusted Empirical Likelihood, Extended Empirical Likelihood

1. Introduction

Longitudinal studies are common in areas such as epidemiology, clinical trials, economics, agriculture, and survey sampling. In longitudinal studies, we are interested in the changes in the variables over time as a function of the covariates,

generally under the assumption that observations from different individuals are independent. For example, longitudinal studies are used to characterize growth and ageing, to assess the effect of risk factors on human health, and to evaluate the effectiveness of treatments. To obtain an unbiased, efficient, and reliable estimate, we must properly model the correlation between the repeated responses for each individual. However, the modelling of correlation, especially when the responses are discrete, is a challenging task even if the responses are collected over equispaced time points.

The approaches used for the analysis of longitudinal data can be classified as mixed effects models, transitional models, and marginal regression models. A potential disadvantage of mixed effects models is that they rely on parametric assumptions, which may lead to biased parameter estimates when a model is misspecified. Moreover, the estimation of the parameters is challenging when the random effects have a high dimension; it typically involves integrals that do not have an explicit form. Transitional models are more difficult to apply when there are missing data and the repeated measurements are not equally spaced in time. In addition, the interpretation of the regression parameters varies with the order of the serial correlation, and the regression parameter estimates are sensitive to the assumption of time dependence. Because of the aforementioned difficulties in modelling and performing inference, we focus on marginal models in this paper.

We start with brief review on existing methods for longitudinal data under the framework of generalized linear models (GLMs). The longitudinal observations consist of an outcome random variable y_{it} and a p -dimensional vector of covariates \mathbf{x}_{it} , observed for subjects $i = 1, \dots, k$ at a time point t , $t = 1, \dots, m_i$. For the i th subject, let $\mathbf{y}_i = (y_{i1}, \dots, y_{im_i})^T$ be the response vector, and let $\mathbf{X}_i = (\mathbf{x}_{i1}, \mathbf{x}_{i2}, \dots, \mathbf{x}_{it}, \dots, \mathbf{x}_{im_i})^T$ be the $m_i \times p$ matrix of covariates. Marginal models for longitudinal data can be extended to the GLM framework. The marginal density of y_{it} is assumed to follow an exponential family [1] of the form

$$f(y_{it}) = \exp\left[(y_{it}\theta_{it} - a(\theta_{it}))\phi + b(y_{it}, \phi)\right], \quad (1)$$

where $\theta_{it} = h(\eta_{it})$, h is a known injective function with $\eta_{it} = \mathbf{x}_{it}\boldsymbol{\beta}$, $\boldsymbol{\beta}$ is a $p \times 1$ vector of regression effects of \mathbf{x}_{it} on y_{it} , and $a(\cdot)$ and $b(\cdot)$ are functions that are assumed to be known. The mean and variance of y_{it} can be written

$$E(y_{it} | \mathbf{x}_{it}) = a'(\theta_{it}) = \mu_{it} \quad \text{and} \quad \text{Var}(y_{it}) = a''(\theta_{it}) = v(\mu_{it})\phi,$$

where ϕ is the unknown over-dispersion parameter and $v(\cdot)$ is a known variance function. For simplicity, we set the nuisance scale parameter ϕ to 1 for the rest of this paper.

The generalized estimating equation (GEE) approach [2] is a semiparametric method in which the estimating equations are derived without a full specification of the joint distribution of the observed data. This approach allows the user to specify any structure for the correlation matrix of the outcomes \mathbf{y}_i for esti-

mating the regression parameters. In this approach, [2] introduced a “working” correlation structure based on the GEE approach to obtain consistent and efficient estimators for the regression parameter β . The estimates of parameters are obtained by solving

$$g(\beta, \hat{\alpha}(\beta)) = \sum_{i=1}^k X_i^T A_i^{1/2} R_i^{-1}(\hat{\alpha}) A_i^{-1/2} (y_i - \mu_i) = 0, \quad (2)$$

where A_i is an $m_i \times m_i$ diagonal matrix with $\text{Var}(\mu_{it})$ as the t th diagonal element and $R_i(\hat{\alpha})$ is the $m_i \times m_i$ working correlation matrix of the m_i repeated measurements. For $j = 1, \dots, m_i$ and $j' = 1, \dots, m_i$, the (j, j') th element of R_i is the known, hypothesized, or estimated correlation. The working correlation may depend on an unknown $s \times 1$ correlation parameter vector α . The observation times and correlation matrix may differ from subject to subject, but the correlation matrix $R_i(\alpha)$ for the i th subject is fully specified by α . Some common working correlation structures are independent, autoregressive of order 1 (AR(1)), equally correlated (EQC), moving average of order 1 (MA(1)), or unstructured.

It has been demonstrated that in some situations the use of an arbitrary working correlation structure may lead to no solution for $\hat{\alpha}$, which may break down the entire GEE methodology (see [3]). In another study [4] showed that the GEE approach may yield an estimator of β , that, although consistent, is less efficient than that of the independence estimating equation approach under an arbitrary working correlation structure. To overcome this difficulty, [5] proposed using a stationary lag correlation structure instead of the working correlation matrix.

The estimate for β is obtained by solving the following estimating equations:

$$g(\beta, \rho) = \sum_{i=1}^k X_i^T A_i \Sigma_i^{-1}(\hat{\rho})(y_i - \mu_i) = 0, \quad (3)$$

where $\Sigma_i(\hat{\rho}) = A_i^{1/2} C_i^*(\rho) A_i^{1/2}$, with $C_i^*(\rho)$ the stationary lag correlation structure for the AR(1), MA(1), or EQC models. The stationary lag correlations can be estimated via the method of moments introduced by [6] and showed that the stationary lag correlation approach produces regression estimates that are consistent and more efficient than those obtained from the independence-assumption-based estimating equation approach (see [4]). Using simulation studies, [7] showed that there is a loss in the efficiency of the GEE estimators when the correlation structures are misspecified. Note, however, that the correlation structure is unknown in practice, and it is better to use a stationary lag-correlation structure which can accommodate the AR(1), EQC, and MA(1) structures. We, therefore, recommend defining a lag correlation structure for the longitudinal responses. When the number of estimations (r) is more than the number of parameters (p) (i.e. $r > p$), we have extra information about the parameter for improved efficiency, but it may not be possible to solve the estimating equations directly. To overcome this problem, [8] proposed an adaptive quadratic inference

function based on moment assumptions and the estimated variance; this does not involve direct estimation of the correlation parameter. Moreover, if the covariates are time-dependent the assumption $\lim_{k \rightarrow \infty} E[g(\beta_0, \hat{\alpha}(\beta_0))] = 0$ might not hold for an arbitrary working correlation structure, and so the GEE estimate of β is not necessarily consistent; see [9] [10] [11] [12] and [13]. The GEE estimator of β with the independent working correlation is always consistent, so [10] recommended using this correlation as a safe choice. However, ignoring the correlation of observations may lead to an inefficient estimate of the regression coefficients and an underestimate of the standard errors.

The GEE approach requires only the assumption of the existence of the first two marginal moments and a correlation structure. GEE estimators are consistent and asymptotically normal as long as the mean, variance, and correlation structure are correctly specified. Marginal models have satisfactory performance when the assumptions are satisfied. Misspecification can cause estimates based on marginal models to be inefficient and inconsistent, and inference in this situation can be completely inappropriate. Confidence regions and hypothesis tests are based on asymptotic normality, which may not hold since the finite-sample distribution may not be symmetric. These problems motivate us to investigate the applicability of empirical likelihood (EL), a nonparametric likelihood method, based on a set of GEEs for the parameter of interest.

The EL introduced by [14] has properties similar to the parametric likelihood. The EL combines the reliability of nonparametric methods with the flexibility and effectiveness of the likelihood approach. The EL has many nice properties parallel to those of parametric likelihood, including the ability to carry out hypothesis tests and construct confidence intervals without estimating the variance. The shape of EL confidence regions automatically reflects the emphasis of the observed data set. The EL method also offers advantages in parameter estimation and the formulation of goodness-of-fit tests. The EL has been successfully applied in areas such as linear models, GLMs, survey sampling, variable selection, survival analysis, and time series. We investigate the use of a nonparametric EL in subject-wise longitudinal data analysis, simultaneously estimating within-subject correlations using the method of moments by [6]. We used the adjusted EL to avoid computational issues. We explore the asymptotic properties of the proposed method and assess the method's performance based on a large number of simulations. Our approach provides consistent estimators and has comparable performance to marginal models when the model assumptions are correct. It is superior to marginal models when the variance function and correlation structure are even misspecified.

The remaining part of the paper is organized as follows. In Section 2, we develop the subject-wise EL via a set of GEEs of the parameter of interest and discuss its characteristics. We then introduce an adjusted EL (AEL) inference to longitudinal data. We discuss its characteristics and asymptotic properties in Section 3. In Section 4, we developed an algorithm based on EL principles for the

estimation of the regression parameters and the construction of its confidence region. In Section 5, the performance of the proposed method is assessed based on Monte Carlo simulations. The implementation of the proposed method in two real case examples is discussed in Section 6 and the conclusions are given in Section 7.

2. Empirical-Likelihood-Based Longitudinal Modelling

EL is a nonparametric-likelihood-based approach, introduced by [14], which is an alternative to parametric likelihood and bootstrap methods. This method enables us to fully employ the information available from the data for making asymptotically efficient inference about the population parameters. In this section, we introduce the EL-based longitudinal modelling.

In a seminal paper, [15] introduced the EL for linear models. EL confidence regions for regression coefficients in linear models were studied by [16]. The EL method can also be used to estimate the parameters defined by a set of estimating equations [17]. A comprehensive overview of the EL and its properties can be obtained from [18]. EL methods have attracted increasing attention over the last two decades, and the literature is extensive.

In longitudinal data analysis framework, [19] applied the EL approach using a subject-wise working independence model. This method ignores the within-subject correlation structure. Also note that [20] proposed a subject-wise EL by entering the longitudinal data and obtained asymptotic normality of the maximum EL estimator (MELE) of the regression coefficients. They did not consider the within-subject correlation structure. It is well known that the working-independence assumption may lead to a loss of efficiency in estimation when a within-subject correlation is present. To estimate the within-subject covariance matrices, [21] used the nonparametric sample covariance matrix obtained from the residuals of a GEE using the working-independence assumption. In this work, we show how to incorporate the within-subject correlation structure of the repeated measurements into the EL.

Following [15] and [17], we can extend the EL inference to longitudinal data based on a set of estimating functions $g(\boldsymbol{\beta}, \boldsymbol{\rho})$ given in (3). We incorporate the within-subject correlation structure of the repeated measurements into the EL using the well-known method of moments estimators by [6] for a given value of $\boldsymbol{\beta}$. The profile empirical log-likelihood function of $\boldsymbol{\beta}$ is defined by

$$\ell(\boldsymbol{\beta}) = \sup \left[\sum_{i=1}^k \log(p_i) : p_i \geq 0, i = 1, 2, \dots, k; \sum_{i=1}^k p_i = 1, \sum_{i=1}^k p_i g_i(\boldsymbol{\beta}, \boldsymbol{\rho}) = 0 \right].$$

The EL is maximized when

$$\hat{p}_i = \frac{1}{k \{1 + \hat{\boldsymbol{\lambda}}^T g_i(\boldsymbol{\beta}, \boldsymbol{\rho})\}}, i = 1, 2, \dots, k, \quad (4)$$

where the Lagrange multiplier $\hat{\boldsymbol{\lambda}} = \hat{\boldsymbol{\lambda}}(\boldsymbol{\beta})$ is the solution of

$$\sum_{i=1}^k \frac{g_i(\boldsymbol{\beta}, \boldsymbol{\rho})}{1 + \boldsymbol{\lambda}^T g_i(\boldsymbol{\beta}, \boldsymbol{\rho})} = 0. \quad (5)$$

This result leads to the profile empirical log-likelihood function

$$\ell(\boldsymbol{\beta}) = -k \log(k) - \sum_{i=1}^k \log(1 + \hat{\boldsymbol{\lambda}}^T(\boldsymbol{\beta}) g_i(\boldsymbol{\beta}, \boldsymbol{\rho}))$$

and the profile empirical log-likelihood ratio function

$$W_l(\boldsymbol{\beta}) = -\sum_{i=1}^k \log(k \hat{p}_i) = \sum_{i=1}^k \log[1 + \hat{\boldsymbol{\lambda}}^T(\boldsymbol{\beta}) g_i(\boldsymbol{\beta}, \boldsymbol{\rho})]. \quad (6)$$

Under some regularity conditions, we have $2W_l(\boldsymbol{\beta}_0) \xrightarrow{D} \chi_p^2$ as $k \rightarrow \infty$ if

$$E[g(\boldsymbol{\beta}_0, \hat{\boldsymbol{\rho}}(\boldsymbol{\beta}_0)) g^T(\boldsymbol{\beta}_0, \hat{\boldsymbol{\rho}}(\boldsymbol{\beta}_0))]$$

is full rank where $\boldsymbol{\beta}_0$ is the true parameter value. This conclusion is similar to that for the parametric likelihood ratio function. The vector $\boldsymbol{\beta}$ can be estimated by minimizing

$$W_l(\boldsymbol{\beta}) = \sum_{i=1}^k \log(1 + \hat{\boldsymbol{\lambda}}^T(\boldsymbol{\beta}) g(\boldsymbol{\beta}, \boldsymbol{\rho})) \quad (7)$$

with respect to $\boldsymbol{\beta}$. Note that the profile log-likelihood ratio function can be minimized with respect to $\boldsymbol{\beta}$ when $\boldsymbol{\rho}$ is known. In practice, $\boldsymbol{\rho}$ is unknown, but can be consistently estimated using the method of moments by [6].

The computation of the profile EL function is a key step in EL applications, and it involves constrained maximization. In some situations, the algorithm may fail because of poor initial values of the parameters. Moreover, the poor accuracy of EL confidence regions has been reported by several authors, including [17] [18] [22] [23] [24] and [25]. In the next subsection we will discuss how to address these problems in the context of longitudinal data.

Adjusted Empirical Likelihood

The computation of the profile EL ratio function $W_l(\boldsymbol{\beta})$ given in (7) is a key step in EL applications. The solution for $\boldsymbol{\lambda}$ must satisfy $\{1 + \hat{\boldsymbol{\lambda}}^T(\boldsymbol{\beta}) g_i(\boldsymbol{\beta}, \hat{\boldsymbol{\rho}}(\boldsymbol{\beta}))\} > 0$ for all $i = 1, \dots, k$. A necessary and sufficient condition for its existence is that the vector $\mathbf{0}$ is an interior point of the convex hull of $\{g_i(\boldsymbol{\beta}, \hat{\boldsymbol{\rho}}(\boldsymbol{\beta})), i = 1, \dots, k\}$. Under some moment conditions on $g(\boldsymbol{\beta}, \hat{\boldsymbol{\rho}}(\boldsymbol{\beta}))$ [18], the convex hull contains $\mathbf{0}$ as an interior point with probability 1 as $k \rightarrow \infty$. However, when $\boldsymbol{\beta}$ is not close to the true parameter value $\boldsymbol{\beta}_0$ or when k is small, it is possible that the solution of (5) does not exist. To avoid this problem, [25] introduced the adjusted EL (AEL). The AEL is obtained by adding a pseudo-observation to the data set. It overcomes the difficulties arising when the estimating equations for $\boldsymbol{\lambda}$ have no solution.

Let $g_i(\boldsymbol{\beta}) = g_i(\boldsymbol{\beta}, \hat{\boldsymbol{\rho}}(\boldsymbol{\beta}))$ and $\bar{g}_k(\boldsymbol{\beta}) = \frac{1}{k} \sum_{i=1}^k g_i(\boldsymbol{\beta})$ for any given $\boldsymbol{\beta}$. For some positive constant b_k , by the addition of an artificial observation

$$g_{k+1}(\boldsymbol{\beta}) = -\frac{b_k}{k} \sum_{i=1}^k g_i(\boldsymbol{\beta}) = -b_k \bar{g}_k(\boldsymbol{\beta})$$

with $b_k = \log(k)/2$. The adjusted profile empirical log-likelihood ratio function is

$$\begin{aligned} W_i^*(\boldsymbol{\beta}) &= \inf \left[-\sum_{i=1}^{k+1} \log[(k+1)p_i] : p_i \geq 0, i=1, 2, \dots, k+1; \sum_{i=1}^{k+1} p_i = 1, \sum_{i=1}^{k+1} p_i g_i(\boldsymbol{\beta}) = \mathbf{0} \right] \\ &= \sum_{i=1}^{k+1} \log \left[1 + \hat{\boldsymbol{\lambda}}^T(\boldsymbol{\beta}) g_i(\boldsymbol{\beta}) \right] \end{aligned}$$

with $\hat{\boldsymbol{\lambda}} = \hat{\boldsymbol{\lambda}}(\boldsymbol{\beta})$ being the solution of $\sum_{i=1}^{k+1} \frac{g_i(\boldsymbol{\beta})}{1 + \hat{\boldsymbol{\lambda}}^T g_i(\boldsymbol{\beta})} = \mathbf{0}$. Note that $\mathbf{0}$ always lies inside the convex hull of $\{g_i(\boldsymbol{\beta}, \hat{\rho}(\boldsymbol{\beta})), i=1, \dots, k+1\}$. The adjusted profile empirical log-likelihood ratio function is well defined after adding a pseudo value $g_{k+1}(\boldsymbol{\beta})$. For a wide range of b_k , following [25], we can show that the adjusted profile EL ratio function $W_i^*(\boldsymbol{\beta})$ has the same asymptotic properties as the unadjusted profile EL ratio function $W_i(\boldsymbol{\beta})$. We define the adjusted profile EL estimator of $\boldsymbol{\beta}$ to be the minimizer of

$$W_i^*(\boldsymbol{\beta}) = \sum_{i=1}^{k+1} \left[\log \left(1 + \hat{\boldsymbol{\lambda}}^T(\boldsymbol{\beta}) g_i(\boldsymbol{\beta}, \hat{\rho}(\boldsymbol{\beta})) \right) \right] \quad (8)$$

with respect to $\boldsymbol{\beta}$.

The adjustment is particularly useful because, even for some undesirable values of $\boldsymbol{\beta}$, the algorithm guarantees a solution. The confidence regions constructed via the AEL are found to have better coverage probabilities than those for the regular EL and the algorithm provides a promising solution for $\boldsymbol{\lambda}$, particularly when the sample size is small. The improved coverage probability is achieved without resorting to more complex procedures such as Bartlett correction or bootstrap calibration.

In the next section, following [17], we state and prove the results on the distributional properties of the adjusted profile EL estimates of $\hat{\boldsymbol{\beta}}$. We construct these theorems based on the GEE with lag correlation given in (3), since the GEE estimate of $\boldsymbol{\beta}$ under an arbitrary working-correlation structure is not necessarily consistent.

3. Main Results

In this section, we present the first-order asymptotic properties of $\hat{\boldsymbol{\beta}}$ and the adjusted profile empirical log-likelihood ratio statistics. We first introduce some notation and regularity conditions that are used in the theorems and lemma.

Regularity Conditions:

- A1. $E\{g(\boldsymbol{\beta}_0, \hat{\rho}(\boldsymbol{\beta}_0))\} = \mathbf{0}$, where $\boldsymbol{\beta}_0$ is the true value of $\boldsymbol{\beta}$,
 $g(\boldsymbol{\beta}, \hat{\rho}(\boldsymbol{\beta})) = \sum_{i=1}^k \mathbf{D}_i^T \boldsymbol{\Sigma}_i^{-1}(\hat{\rho})(\mathbf{y}_i - \boldsymbol{\mu}_i)$ be the estimating function for $\boldsymbol{\beta} \in \mathcal{R}^p$ (defined in (3)), $\mathbf{D}_i = \partial\{a_i'(\theta)\}/\partial\boldsymbol{\beta}$, $\boldsymbol{\Sigma}_i(\hat{\rho}) = \mathbf{A}_i^{1/2} \mathbf{C}_i^*(\hat{\rho}) \mathbf{A}_i^{1/2}$, and

$A_i = \text{diag} \{a_i''(\theta)\}$ for $i = 1, 2, \dots, k$. Let $\bar{g}_k(\boldsymbol{\beta}, \hat{\rho}(\boldsymbol{\beta})) = \frac{1}{k} \sum_{i=1}^k g_i(\boldsymbol{\beta}, \hat{\rho}(\boldsymbol{\beta}))$ and $g_{k+1}(\boldsymbol{\beta}, \hat{\rho}(\boldsymbol{\beta})) = -b_k \bar{g}_k(\boldsymbol{\beta}, \hat{\rho}(\boldsymbol{\beta}))$, where b_k is a positive constant.

A2. $\{a'(\theta)\}$ is three times continuously differentiable and $\{a''(\theta)\} > 0$ in Θ° , where Θ be the natural parameter space of the exponential family distributions presented in (1) and Θ° the interior of Θ . Also, $h(\eta)$ is three times continuously differentiable and $h'(\eta) > 0$.

A3. $E_{\boldsymbol{\beta}_0} \left\{ \frac{\partial g_k(\boldsymbol{\beta}, \rho)}{\partial \boldsymbol{\beta}} \right\}$ and

$V_k(\boldsymbol{\beta}_0, \hat{\rho}(\boldsymbol{\beta}_0)) = E_{\boldsymbol{\beta}_0} \{g_k(\boldsymbol{\beta}, \hat{\rho}(\boldsymbol{\beta}))g_k^T(\boldsymbol{\beta}, \hat{\rho}(\boldsymbol{\beta}))\}$ are positive definite.

A4. The rank of $E \left\{ \frac{\partial g_k(\boldsymbol{\beta}, \rho)}{\partial \boldsymbol{\beta}} \right\}$ is p in a neighbourhood of $\boldsymbol{\beta}_0$.

A5. There exist functions $G(\mathbf{y}, \mathbf{X})$ such that in a neighbourhood of $\boldsymbol{\beta}_0$.

$$\left| \frac{\partial g_k(\boldsymbol{\beta}, \rho)}{\partial \boldsymbol{\beta}} \right| < G(\mathbf{y}, \mathbf{X}), \|g_k(\mathbf{y}, \mathbf{X}, \boldsymbol{\beta}, \hat{\rho}(\boldsymbol{\beta}))\|^3 < G(\mathbf{y}, \mathbf{X})$$

with $E[G(\mathbf{y}, \mathbf{X})] < \infty$.

Theorem 3.1. Under regularity conditions A1-A5, suppose $(\mathbf{y}_i, \mathbf{X}_i), i = 1, 2, \dots, k$ is a set of independent and identically distributed random vectors. Let

$$2W_i^*(\boldsymbol{\beta}) = 2 \sum_{i=1}^{k+1} \log \left[1 + \hat{\boldsymbol{\lambda}}^T(\boldsymbol{\beta}) g_i(\boldsymbol{\beta}, \hat{\rho}(\boldsymbol{\beta})) \right] \tag{9}$$

be the adjusted profile empirical log-likelihood ratio function. Then, as $k \rightarrow \infty$, $\hat{\rho}(\boldsymbol{\beta})$ is a consistent estimator in the neighbourhood of $\boldsymbol{\beta}$; the correlation matrix of \mathbf{y}_i is $\mathbf{C}_i^*(\boldsymbol{\rho})$, defined in (3) and $W_i^*(\boldsymbol{\beta})$ attains its minimum value at some point $\hat{\boldsymbol{\beta}}$ in the interior of $\|\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0\| < k^{-1/3}$ in probability.

This result corresponds to Lemma 1 in [17] which is about the consistency of maximum empirical likelihood estimates (MELE) for independent and identically distributed data. By following [17], under the regularity conditions A1-A5, we can obtain a subject-wise MELE, as $k \rightarrow \infty$, with probability tending to 1 the equation $W_i^*(\boldsymbol{\beta})$ has a solution within the open ball $\|\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0\| < k^{-1/3}$. It is noted that the proof is similar to the proof of Lemma 1 in [19] and the details are omitted here.

Theorem 3.2. In addition to the regularity conditions A1-A5, suppose that $\frac{\partial^2 g(\boldsymbol{\beta}, \rho)}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}^T}$ is bounded by some integrable function $G(\mathbf{y}, \mathbf{X})$ in the neighbourhood. Then, there exists a sequence of adjusted profile EL estimates $\hat{\boldsymbol{\beta}}$ of $\boldsymbol{\beta}$ such that

$$\sqrt{k}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0) \xrightarrow{D} N(\mathbf{0}, \Delta),$$

where

$$\Delta = \left[E_{\boldsymbol{\beta}_0} \left\{ \frac{\partial g(\boldsymbol{\beta}, \hat{\rho}(\boldsymbol{\beta}))}{\partial \boldsymbol{\beta}} \right\}^T \left\{ E_{\boldsymbol{\beta}_0} \{g(\boldsymbol{\beta}, \hat{\rho}(\boldsymbol{\beta}))g^T(\boldsymbol{\beta}, \hat{\rho}(\boldsymbol{\beta}))\}^{-1} \right\} E_{\boldsymbol{\beta}_0} \left\{ \frac{\partial g(\boldsymbol{\beta}, \hat{\rho}(\boldsymbol{\beta}))}{\partial \boldsymbol{\beta}} \right\} \right]^{-1}$$

It is noted that the proof of Theorem 3.2 is similar to the proof of Theorem 1 in [17]. The details are thus omitted here.

Theorem 3.3. *Under regularity conditions A1-A5, the adjusted profile empirical log-likelihood ratio function $2W_l^*(\beta_0)$, where β_0 is the true value of β , is asymptotically chi-squared distributed with degrees of freedom p .*

The proof of Theorem 3.3 can be achieved by using similar arguments as those used in the proof of Theorem 2 in [17]. The details are thus omitted here.

4. Algorithm

To implement our method, we need an efficient algorithm. We minimize the profile EL ratio function $W_l(\beta)$ with respect to β using a Newton-Raphson algorithm. At each Newton-Raphson iteration, we compute the Lagrange multiplier for updated values of β and $\hat{\rho}(\beta)$. We used the modified Newton-Raphson algorithm proposed by [26] for computing the Lagrange multiplier for a given value of the parameter. We implemented this method, which is numerically stable. The algorithm given in Sections 4.1, 4.2, and 4.3 can easily be extended to the AEL by the addition of a pseudo-value $g_{k+1}(\beta) = -b_k \bar{g}_k(\beta, \hat{\rho}(\beta))$, where b_k is a positive constant.

4.1. Computation of Lagrange Multiplier

The Lagrange multiplier λ is estimated by solving the equation

$$\sum_{i=1}^k \frac{g_i(\beta, \hat{\rho}(\beta))}{1 + \lambda^T g_i(\beta, \hat{\rho}(\beta))} = 0$$

for a given set of vectors $g_i(\beta, \hat{\rho}(\beta))$, $i = 1, 2, \dots, k$. Note that the above equation is the derivative of R with respect to λ for a given β , where

$$R = \sum_{i=1}^k \log \{1 + \lambda^T g_i(\beta, \hat{\rho}(\beta))\}. \quad (10)$$

In the EL problem, the solution must satisfy

$$1 + \lambda^T g_i(\beta, \hat{\rho}(\beta)) > 0, \quad i = 1, 2, \dots, k.$$

The modified Newton-Raphson algorithm for estimating λ for a given value of β and $\hat{\rho}(\beta)$ is as follows:

1. Set $\lambda^c = 0$, $c = 0$, $\gamma^c = 1$, $\epsilon = 1e^{-08}$, $\rho = \rho^0$, and $\beta = \beta^0$.
2. Let R^λ and $R^{\lambda\lambda}$ be the first and second partial derivatives of R (given in (10)) with respect to λ :

$$R^\lambda = \sum_{i=1}^k \left[\frac{g_i(\beta, \hat{\rho}(\beta))}{\{1 + \lambda^T g_i(\beta, \hat{\rho}(\beta))\}} \right],$$

$$R^{\lambda\lambda} = - \sum_{i=1}^k \left[\frac{g_i(\beta, \hat{\rho}(\beta)) g_i^T(\beta, \hat{\rho}(\beta))}{\{1 + \lambda^T g_i(\beta, \hat{\rho}(\beta))\}^2} \right].$$

Compute R^λ and $R^{\lambda\lambda}$ for $\lambda = \lambda^c$ and let $\Delta(\lambda^c) = -[R^{\lambda\lambda}]^{-1} R^\lambda$.

- If $\|\Delta(\lambda^c)\| < \epsilon$ stop the algorithm and report λ^c ; otherwise continue.
3. Calculate $\delta^c = \gamma^c \Delta(\lambda^c)$. If $1 + (\lambda^c - \delta^c) g_i(\beta, \hat{\rho}(\beta)) \leq 0$ for some i , set $\gamma^c = \frac{\gamma^c}{2}$ and go to Step 2.
 4. Set $\lambda^{c+1} = \lambda^c - \delta^c$, $c = c + 1$, and $\gamma^{c+1} = (c + 1)^{-\frac{1}{2}}$ and go to Step 2. Step 2 will guarantee that $p_i > 0$ and the optimization is carried out in the right direction.

4.2. Algorithm for Optimizing Profile Empirical Likelihood Ratio Function

Let $\hat{\lambda}(\beta)$ be the estimated value of λ for a given β . We minimize the profile EL ratio function defined in (7) over β . The Newton-Raphson algorithm is as follows:

1. Set $\beta = \beta^0$, $h = 0$, and $\epsilon = 1e^{-08}$.
2. Let $\hat{\lambda} = \lambda(\beta)$ and $\hat{\rho}(\beta)$ be the estimated values of λ and ρ .
3. Compute the new estimate of β via

$$\beta^{(h+1)} = \beta^{(h)} - \{W_i^{\beta\beta}(\beta^{(h)})\}^{-1} \{W_i^\beta(\beta^{(h)})\} \tag{11}$$

where $W_i(\beta)$ is the profile empirical log-likelihood ratio function defined in (7), with

$$W_i^\beta = \frac{\partial W_i(\beta)}{\partial \beta}, \quad W_i^{\beta\beta} = \frac{\partial^2 W_i(\beta)}{\partial \beta \partial \beta^T}.$$

Note that to compute W_i^β and $W_i^{\beta\beta}$, we need to estimate the Lagrange multiplier $\hat{\lambda}(\beta)$ as in Section 4.1. In practice, ρ is unknown, and the correlations can be consistently estimated by [24] using the method of moments.

4. If $\min|\beta^{(h+1)} - \beta^{(h)}| < \epsilon$ stop the algorithm and report $\beta^{(h+1)}$; otherwise set $h = h + 1$ and go to Step 3.

The simplified expressions for W_i^β and $W_i^{\beta\beta}$ are as follows. Let R^β , $R^{\beta\beta}$, and $R^{\beta\lambda}$ be the first and second partial derivatives of (10) with respect to β and λ

$$R^\beta = \sum_{i=1}^k \left[\frac{g'_i(\beta, \hat{\rho}(\beta)) \lambda}{\{1 + \lambda^T g_i(\beta, \hat{\rho}(\beta))\}} \right],$$

$$R^{\beta\beta} = \sum_{i=1}^k \left\{ \left[\frac{g''_i(\beta, \hat{\rho}(\beta)) \lambda^T}{\{1 + \lambda^T g_i(\beta)\}} \right] - \left[\frac{g'_i(\beta, \hat{\rho}(\beta)) \lambda \lambda^T [g'_i(\beta, \hat{\rho}(\beta))]^T}{\{1 + \lambda^T g_i(\beta, \hat{\rho}(\beta))\}^2} \right] \right\},$$

and

$$R^{\beta\lambda} = \sum_{i=1}^k \left[\frac{\{1 + \lambda^T g_i(\beta)\} g'_i(\beta, \hat{\rho}(\beta)) - g'_i(\beta, \hat{\rho}(\beta)) \lambda [g_i(\beta)]^T}{\{1 + \lambda^T g_i(\beta)\}^2} \right].$$

The first derivative of $W_i(\beta)$ with respect to β is

$$W_i^\beta = \sum_{i=1}^k \left[\frac{\left[\frac{\partial \lambda(\beta)}{\partial \beta} \right]^T g_i(\beta, \hat{\rho}(\beta)) + g_i'(\beta, \hat{\rho}(\beta)) \lambda(\beta)}{\{1 + \lambda^T(\beta) g_i(\beta, \hat{\rho}(\beta))\}} \right]$$

$$= \left[\frac{\partial \lambda(\beta)}{\partial \beta} \right]^T R^\lambda + R^\beta.$$

Note that for $\lambda = \hat{\lambda}(\beta)$, $R^\lambda = 0$. Therefore,

$$W_i^\beta = R^\beta. \tag{12}$$

Similarly, the second derivative of $W_i(\beta)$ with respect to β is

$$W_i^{\beta\beta} = \sum_{i=1}^k \left[\frac{\{1 + \lambda^T(\beta) g_i(\beta, \hat{\rho}(\beta))\} \left\{ \left[\frac{\partial^2 \lambda(\beta)}{\partial \beta \partial \beta^T} \right] g_i(\beta, \hat{\rho}(\beta)) + 2g_i'(\beta) \left[\frac{\partial \lambda(\beta)}{\partial \beta} \right]^T + g_i''(\beta, \hat{\rho}(\beta)) \lambda(\beta) \right\}}{\{1 + \lambda^T(\beta) g_i(\beta, \hat{\rho}(\beta))\}^2} \right]$$

$$- \sum_{i=1}^k \left[\frac{\left\{ \left[\frac{\partial \lambda(\beta)}{\partial \beta} \right]^T g_i(\beta, \hat{\rho}(\beta)) + g_i'(\beta, \hat{\rho}(\beta)) \lambda(\beta) \right\} \left\{ \left[\frac{\partial \lambda(\beta)}{\partial \beta} \right]^T g_i(\beta, \hat{\rho}(\beta)) + g_i'(\beta, \hat{\rho}(\beta)) \lambda(\beta) \right\}^T}{\{1 + \lambda^T(\beta) g_i(\beta, \hat{\rho}(\beta))\}^2} \right]$$

$$= \left[\frac{\partial \lambda(\beta)}{\partial \beta} \right]^T R^{\lambda\lambda} \left[\frac{\partial \lambda(\beta)}{\partial \beta} \right] + 2 \left[\frac{\partial \lambda(\beta)}{\partial \beta} \right]^T R^{\beta\lambda} + R^{\beta\beta}.$$

Following [18], a local quadratic approximation to R leads to

$$\left[\frac{\partial \lambda(\beta)}{\partial \beta} \right] = (R^{\lambda\lambda})^{-1} R^{\beta\lambda},$$

so

$$W_i^{\beta\beta} = R^{\beta\beta} - R^{\beta\lambda} (R^{\lambda\lambda})^{-1} R^{\lambda\beta}. \tag{13}$$

4.3. Construction of Confidence Interval

We use the bisection method to construct the lower and upper confidence limits based on the profile EL ratio for β . Let $\hat{\beta} = (\hat{\beta}_1, \hat{\beta}_2)^T$ be the estimated value of β from Section 4.2, where $\hat{\beta}_1$ is a scalar and $\hat{\beta}_2$ is the $1 \times p - 1$ vector of parameters and we are interested to construct confidence interval for β_1 .

1. Compute a reasonable lower confidence limit $\beta_{1,L}$ for β_1 . Set $L_1 = \hat{\beta}_1$, $L_2 = \hat{\beta}_1 - a \times SE(\hat{\beta}_1)$, and $\epsilon = 1e^{-05}$, where $SE(\hat{\beta}_1)$ is the standard error of $\hat{\beta}_1$ using any existing method. We can choose a such that

$W_i(L_2, \hat{\beta}_2) > [\chi_{1,1-\alpha}^2] / 2 > W_i(L_1, \hat{\beta}_2)$, where $\chi_{1,1-\alpha}^2$ is the $(1-\alpha)$ th quantile from a χ^2 distribution with one degree of freedom.

2. Compute the profile empirical log-likelihood ratio values $W_1 = 2W_i(L_1, \hat{\beta}_2)$ and $W_2 = 2W_i(L_2, \hat{\beta}_2)$.

3. Minimize the profile EL ratio function defined in (7) over β_2 for a given $L_{new} = (L_1 + L_2)/2$. Let $\hat{\beta}_{2_{new}}$ be the new estimate of β_2 and $W_{new} = 2W_1(L_{new}, \hat{\beta}_{2_{new}})$.
 4. If $W_{new} < \chi^2_{1,1-\alpha}$, set $L_1 = L_{new}$ and $W_1 = W_{new}$; else set $L_2 = L_{new}$ and $W_2 = W_{new}$.
 5. If $|W_1 - W_2| < \epsilon$ stop the algorithm and report $\beta_{1,L}$; otherwise go to Step 3.
- We can use this approach to construct the upper confidence limit by setting $U_1 = \hat{\beta}_1$ and $U_2 = \hat{\beta}_1 + a \times SE(\hat{\beta}_1)$.

5. Performance Analysis

In this section, we conduct simulation studies to investigate the performance of our EL-based approach. We compute the coverage probabilities based on the ordinary EL, AEL and compare them with those of the GEE approach, which is based on a normal approximation. We also compute the coverage probabilities based on the extended empirical likelihood (EEL) by [27] and [28], which expands the EL domain geometrically and improve the coverage probabilities. We use different working correlations for the comparison. We generate count and continuous responses with different correlation structures and compare the methods under different working correlation structures.

5.1. Correlation Models for Stationary Count Data

We consider the stationary correlation models for count data discussed by [29] and [30]. The three models used to generate the data are

(i) Poisson Autoregressive Order 1 (AR(1)) Model

Let $y_{it} \sim \text{Poi}(\tilde{\mu}_i)$, where $\tilde{\mu}_i = \exp(\tilde{x}_i\beta)$. The repeated responses follow the AR lag 1 dynamic model given by

$$y_{it} = \rho * y_{i,t-1} + d_{it}, t = 2, \dots, m_i. \tag{14}$$

Given $y_{i,t-1}$, $\rho * y_{i,t-1}$ is the binomial thinning operation. That is,

$$\rho * y_{i,t-1} = \sum_{j=1}^{y_{i,t-1}} b_j(\rho) = z_{i,t-1},$$

where the $b_j(\rho)$ are independent and identically distributed Bernoulli(ρ) random variables. We assume that $d_{it} \sim \text{Poi}(\tilde{\mu}_i(1-\rho))$ and it is independent of $z_{i,t-1}$. Let $\tilde{x}_i = (\tilde{x}_{i1}, \dots, \tilde{x}_{ip})$ be the time-independent covariate for the i th individual.

(ii) Poisson Moving Average Order 1 (MA(1)) Model

The repeated responses follow the MA lag 1 dynamic model given by

$$y_{it} = \rho * d_{i,t-1} + d_{it}, t = 2, \dots, m_i, \tag{15}$$

where $\rho * d_{i,t-1} = \sum_{j=1}^{d_{i,t-1}} b_j(\rho)$ is a binomial thinning operation and

$d_{it} \sim \text{Poi}\left[\frac{\tilde{\mu}_i}{1+\rho}\right]$, $t = 0, \dots, m_i$, with $\tilde{\mu}_i = \exp(\tilde{x}_i\beta)$. Here $t = 0$ is the initial time.

(iii) Poisson Equally Correlated Model

Let $y_{i0} \sim \text{Poi}(\tilde{\mu}_i)$ and $d_{it} \sim \text{Poi}[\tilde{\mu}_i(1-\rho)]$ for all $t=1, \dots, m_i$. The repeated responses follow the dynamic equicorrelation model given by

$$y_{it} = \rho * y_{i0} + d_{it}, \text{ for } t=1, \dots, m_i. \quad (16)$$

We simulated 1000 data sets from each of these models follow the AR(1), EQC, or MA(1) structure, and used EL-based methods to estimate the parameters using different working correlation such as AR(1), EQC, and MA(1) as well as lag correlation. In each simulation we use the parameters

$\beta = (\beta_1, \beta_2)^T = (0.3, 0.2)^T$ and $\rho = 0.5$. We consider $k=100$ subjects and $m=4$ time points. For the i th subject, we generate the covariates $\tilde{x}_i = (\tilde{x}_{i1}, \tilde{x}_{i2})$ from a normal distribution with mean 0 and standard deviation 1. For the analysis, we consider the working correlation to be either a true correlation or a lag correlation. We did not consider other possible values for ρ since the working correlation structure may lead to no solution for $\hat{\alpha}$ in some situations.

Table 1 gives the average estimated values of the regression coefficients with the corresponding simulated standard errors in parentheses for the independent, AR(1), EQC, and MA(1) models. We also give the coverage probabilities for β_1 and β_2 for the 0.95 and 0.99 confidence levels with the average width of the CI in parentheses. The results in **Table 1** shows that the estimates $\hat{\beta}_1$ and $\hat{\beta}_2$ are close to the true values, width, and the coverage probabilities of the intervals based on the EL, EEL, and AEL are similar to those of the GEE. For instance, in the AR(1)/AR(1) case (true model/working correlation structure) the coverage probabilities of $\hat{\beta}_1$ based on the GEE, EL, EEL, and AEL are 0.947, 0.928, 0.937, and 0.937 respectively for the nominal level of 0.95. For $\hat{\beta}_2$, these probabilities are 0.954, 0.934, 0.940, and 0.942 for the same nominal level. Note that the intervals based on the EL have a slight undercoverage compared with those based on GEE. The EEL and AEL give substantially better coverage probabilities. Moreover, EEL and AEL are consistently more accurate than the EL. The results for lag correlations have similar patterns.

5.2. Misspecified Working Correlation Structure

In the simulation studies discussed in Section 5.1 we considered the correlation structure used to generate the data as the working correlation in the GEE-based modelling. However, in practice, we do not know the correlation structure of the data. As discussed before, if the working correlation is misspecified, we may lose the efficiency of the parameter estimates [3] [4].

We conducted a simulation study to assess the loss of efficiency. We generated repeated counts with the AR(1) correlation structure given in Section 5.1(i) with $\rho = 0.49$ and 0.70 and $m = 5$ time points. We used three working correlation structures: EQC, MA(1), and lag correlation. **Table 2** gives the results for the GEE, EL, EEL, and AEL.

Table 2 shows that the EL, EEL, and AEL are superior to the GEE when the correlation structure is misspecified. Note that, in this EL-based approach, we

could construct CIs without estimating the variance of the parameter of interest. For example, in the AR(1)/EQC case the coverage probabilities of $\hat{\beta}_1$ based on the GEE, EL, EEL, and AEL are 0.917, 0.928, 0.934, and 0.935 respectively for the nominal 0.95 level. For $\hat{\beta}_2$, these probabilities are 0.916, 0.929, 0.937, and 0.937 for the same nominal level. In this situation, the GEE with stationary lag correlation performs better than the GEE with a misspecified working correlation. However, the EL, EEL, and AEL perform as well as the former method, despite being nonparametric methods based on a data-driven likelihood ratio function. We did not consider all possible cases, for instance, a true EQC or MA(1) correlation model, since under different working correlation structures the correlation parameter $\hat{\alpha}$ does not exist (see [5]).

5.3. Over-Dispersed Stationary Count Data

In this section, we consider the performance of our approach when the variance function is misspecified, in the context of stationary count data. We generate over-dispersed stationary count data y_{it} using $\tilde{\mu}_i = u_i \exp(\tilde{x}_i \beta)$ for the three models discussed in Section 5.1, where u_i is a random sample such that $E(u_i) = 1$ and $\text{Var}(u_i) = \omega$. Marginally, we have $E(y_{it}) = \tilde{\mu}_i$ and $\text{Var}(y_{it}) = \tilde{\mu}_i(1 + \tilde{\mu}_i\omega)$. The distribution of u is chosen to be gamma with shape parameter ω and scale parameter $1/\omega$, where ω is the over-dispersion parameter. We choose over-dispersion parameter $\omega = 1/4$. However, the GEE, EL, EEL, and AEL CIs are constructed under the assumption that there is no over-dispersion.

Table 3 gives the average estimated values of the regression coefficients, the corresponding simulated standard errors in parentheses, the coverage probabilities for β_1 and β_2 for the 0.95 and 0.99 confidence levels, and the average width of the CI in parentheses for the independent, AR(1), EQC, and MA(1) models. **Table 3** shows that when there is over-dispersion, the EL, EEL, and AEL outperform the GEE. In the AR(1)/AR(1) case the coverage probabilities of $\hat{\beta}_1$ based on the GEE, EL, EEL, and AEL are 0.876, 0.916, 0.926, and 0.931 respectively for the nominal 0.95 level. For $\hat{\beta}_2$, these probabilities are 0.891, 0.920, 0.929, and 0.931 for the same nominal level. This indicates that the EL, EEL, and AEL are fairly robust to model misspecification. Note that the construction of the CI based on the EL, EEL, and AEL does not require the estimation of the scale parameter.

5.4. Correlation Models for Continuous Data

In this section, we investigate the performance of our EL approach on a class of stationary and nonstationary correlation models for longitudinal continuous data. The random errors $(\epsilon_1, \epsilon_2, \epsilon_3, \epsilon_4)^T$ are generated from the multivariate normal distribution with marginal mean 0, marginal variance 1, and an auto-correlation coefficient $\rho = 0.5$. In this performance analysis, we consider three correlation models: exchangeable, AR(1), and MA(1).

(i) AR(1) Structure

For $t = 1, \dots, m_i$, for the i th individual

$$y_{it} = \mathbf{x}_{it}\boldsymbol{\beta} + \epsilon_{it}, \quad (17)$$

and we assume that

$$\epsilon_{it} = \rho\epsilon_{it} + a_{it},$$

with $|\rho| < 1$ and $a_{it} \sim N(0,1)$.

(ii) MA(1) Structure The ϵ_{it} in (17) follow the model

$$\epsilon_{it} = \rho a_{i,t-1} + a_{it}$$

where ρ is a suitable scale parameter that does not necessarily satisfy $|\rho| < 1$, and $a_{it} \sim N(0,1)$.

(iii) Equicorrelation (EQC) Structure The ϵ_{it} in (17) follow the model

$$\epsilon_{it} = \rho a_{i0} + a_{it},$$

where a_{i0} is an error value at the initial time, and ρ is a suitable correlation parameter. We assume that

$$a_{it} \sim N(0,1) \text{ and } a_{i0} \sim N(0,1),$$

and a_{it} and a_{i0} are independent for all t .

We simulated 1000 data sets from the above models under stationary and non-stationary covariates, using the parameters $\boldsymbol{\beta} = (\beta_1, \beta_2)^\top = (0.4, 0.5)^\top$, $\rho = 0.5$, and $m = 4$. For the i th subject, we generate the covariates $\tilde{\mathbf{x}}_i = (\tilde{x}_{i1}, \tilde{x}_{i2})$ from a normal distribution with mean 0 and standard deviation 1. **Table 4** gives the mean estimated values of the regression coefficients, the corresponding simulated standard errors in parentheses, the simulated coverage probabilities for β_1 and β_2 for the 0.95 and 0.99 confidence levels, and the average width of the CI in parentheses for the independent, AR(1), EQC, and MA(1) models with stationary covariates. **Table 5** gives the results for nonstationary covariates.

The coverage probabilities of the intervals based on the EL, EEL, and AEL are similar to those of the GEE. For instance, in the MA(1)/MA(1) case in **Table 4** the coverage probabilities of $\hat{\beta}_1$ based on the GEE, EL, EEL, and AEL are 0.955, 0.945, 0.955, and 0.954 respectively for the nominal 0.95 level. For $\hat{\beta}_2$, these probabilities are 0.958, 0.944, 0.948, and 0.951 for the same nominal level. Note that the intervals based on the EL have a slight undercoverage compared with those for the GEE. Also, the EEL and AEL are consistently more accurate than the EL. The lag-correlation-based coverage probabilities have similar patterns.

5.5. Correlation Models for Misspecified Continuous Data

In this section, we compare the performances of the methods when the correlation model for continuous data is misspecified. The stationary and nonstationary correlation models for longitudinal continuous data are generated from (17) for the parameter set in Section 5.4, and the correlated random errors $(\epsilon_1, \epsilon_2, \epsilon_3, \epsilon_4)^\top$ are generated from the $\chi^2(1)-1$ distribution instead of the normal distribution for the three correlation models:

Table 1. Coverage probabilities of regression estimates for count data with stationary covariates for the independent, AR(1), EQC, and MA(1) models.

True model/ Working correlation	Estimate	Statistic	Coverage Probabilities							
			95% level				99% level			
			GEE	EL	EEL	AEL	GEE	EL	EEL	AEL
IND/IND $\rho = 0.50$	$\hat{\beta}_1$	0.2975 (0.046)	0.953 (0.184)	0.929 (0.175)	0.936 (0.178)	0.937 (0.179)	0.986 (0.242)	0.985 (0.232)	0.987 (0.240)	0.987 (0.238)
	$\hat{\beta}_2$	0.1978 (0.048)	0.944 (0.188)	0.927 (0.179)	0.934 (0.183)	0.936 (0.184)	0.990 (0.248)	0.980 (0.238)	0.984 (0.246)	0.984 (0.244)
AR(1)/AR(1) $\rho = 0.50$	$\hat{\beta}_1$	0.3006 (0.065)	0.947 (0.258)	0.928 (0.246)	0.937 (0.251)	0.937 (0.252)	0.985 (0.339)	0.980 (0.326)	0.981 (0.338)	0.981 (0.335)
	$\hat{\beta}_2$	0.1978 (0.067)	0.954 (0.264)	0.934 (0.251)	0.940 (0.256)	0.942 (0.257)	0.990 (0.347)	0.980 (0.332)	0.984 (0.343)	0.983 (0.340)
AR(1)/Lag $\rho = 0.50$	$\hat{\beta}_1$	0.3006 (0.066)	0.946 (0.256)	0.930 (0.246)	0.936 (0.250)	0.939 (0.252)	0.985 (0.337)	0.978 (0.325)	0.981 (0.337)	0.981 (0.334)
	$\hat{\beta}_2$	0.1978 (0.067)	0.952 (0.263)	0.931 (0.250)	0.938 (0.255)	0.940 (0.256)	0.990 (0.345)	0.980 (0.331)	0.985 (0.342)	0.985 (0.339)
EQC/EQC $\rho = 0.50$	$\hat{\beta}_1$	0.2984 (0.074)	0.955 (0.288)	0.942 (0.274)	0.946 (0.280)	0.948 (0.281)	0.988 (0.379)	0.986 (0.364)	0.988 (0.376)	0.987 (0.373)
	$\hat{\beta}_2$	0.1936 (0.074)	0.950 (0.295)	0.939 (0.282)	0.941 (0.288)	0.941 (0.289)	0.990 (0.387)	0.986 (0.372)	0.986 (0.386)	0.986 (0.383)
EQC/Lag $\rho = 0.50$	$\hat{\beta}_1$	0.2985 (0.074)	0.954 (0.288)	0.940 (0.274)	0.948 (0.280)	0.948 (0.281)	0.987 (0.379)	0.984 (0.363)	0.987 (0.376)	0.986 (0.373)
	$\hat{\beta}_2$	0.1933 (0.075)	0.952 (0.294)	0.937 (0.281)	0.940 (0.287)	0.943 (0.288)	0.990 (0.387)	0.986 (0.372)	0.986 (0.385)	0.986 (0.382)
MA(1)/MA(1) $\rho = 0.50$	$\hat{\beta}_1$	0.2989 (0.058)	0.943 (0.222)	0.926 (0.211)	0.929 (0.215)	0.931 (0.216)	0.989 (0.291)	0.982 (0.280)	0.985 (0.289)	0.983 (0.287)
	$\hat{\beta}_2$	0.2022 (0.056)	0.952 (0.227)	0.932 (0.216)	0.935 (0.220)	0.936 (0.221)	0.994 (0.298)	0.984 (0.285)	0.991 (0.296)	0.989 (0.293)
MA(1)/Lag $\rho = 0.50$	$\hat{\beta}_1$	0.2990 (0.058)	0.944 (0.220)	0.928 (0.210)	0.929 (0.215)	0.929 (0.216)	0.986 (0.290)	0.980 (0.279)	0.985 (0.289)	0.985 (0.287)
	$\hat{\beta}_2$	0.2024 (0.056)	0.946 (0.225)	0.931 (0.215)	0.934 (0.220)	0.937 (0.221)	0.992 (0.296)	0.983 (0.285)	0.989 (0.295)	0.988 (0.293)

Table 2. Coverage probabilities of regression estimates for count data with stationary covariates when the working correlation is misspecified for an AR(1) model.

True model/ Working correlation	Estimate	Statistic	Coverage Probabilities							
			95% level				99% level			
			GEE	EL	EEL	AEL	GEE	EL	EEL	AEL
AR(1)/AR(1) $\rho = 0.70$	$\hat{\beta}_1$	0.3001 (0.070)	0.952 (0.279)	0.938 (0.265)	0.944 (0.270)	0.946 (0.272)	0.987 (0.367)	0.983 (0.351)	0.984 (0.364)	0.984 (0.361)
	$\hat{\beta}_2$	0.2009 (0.073)	0.950 (0.286)	0.928 (0.270)	0.938 (0.275)	0.941 (0.276)	0.988 (0.375)	0.986 (0.356)	0.989 (0.369)	0.988 (0.366)
AR(1)/EQC $\rho = 0.70$	$\hat{\beta}_1$	0.2997 (0.073)	0.911 (0.247)	0.932 (0.273)	0.935 (0.278)	0.936 (0.280)	0.973 (0.325)	0.975 (0.361)	0.978 (0.375)	0.978 (0.371)
	$\hat{\beta}_2$	0.1956 (0.076)	0.902 (0.252)	0.934 (0.278)	0.936 (0.284)	0.936 (0.286)	0.963 (0.331)	0.977 (0.368)	0.982 (0.381)	0.980 (0.378)
AR(1)/Lag $\rho = 0.70$	$\hat{\beta}_1$	0.3002 (0.070)	0.952 (0.278)	0.932 (0.264)	0.940 (0.268)	0.940 (0.270)	0.986 (0.366)	0.982 (0.349)	0.982 (0.361)	0.982 (0.359)
	$\hat{\beta}_2$	0.2007 (0.073)	0.950 (0.284)	0.926 (0.268)	0.938 (0.273)	0.940 (0.275)	0.988 (0.373)	0.985 (0.354)	0.988 (0.367)	0.987 (0.364)
AR(1)/AR(1) $\rho = 0.49$	$\hat{\beta}_1$	0.2989 (0.062)	0.938 (0.237)	0.918 (0.225)	0.922 (0.229)	0.923 (0.231)	0.988 (0.311)	0.977 (0.298)	0.982 (0.309)	0.978 (0.306)
	$\hat{\beta}_2$	0.1956 (0.062)	0.940 (0.243)	0.920 (0.231)	0.928 (0.235)	0.928 (0.236)	0.993 (0.0.319)	0.985 (0.305)	0.988 (0.316)	0.987 (0.313)
AR(1)/EQC $\rho = 0.49$	$\hat{\beta}_1$	0.2992 (0.061)	0.899 (0.207)	0.931 (0.231)	0.936 (0.236)	0.936 (0.237)	0.968 (0.272)	0.978 (0.306)	0.985 (0.317)	0.984 (0.314)
	$\hat{\beta}_2$	0.1987 (0.062)	0.908 (0.212)	0.945 (0.236)	0.948 (0.241)	0.950 (0.242)	0.980 (0.278)	0.987 (0.313)	0.992 (0.324)	0.990 (0.321)
AR(1)/MA(1) $\rho = 0.49$	$\hat{\beta}_1$	0.2991 (0.061)	0.897 (0.205)	0.931 (0.228)	0.934 (0.232)	0.936 (0.233)	0.968 (0.270)	0.979 (0.302)	0.985 (0.313)	0.984 (0.310)
	$\hat{\beta}_2$	0.1985 (0.061)	0.905 (0.210)	0.936 (0.233)	0.944 (0.238)	0.944 (0.239)	0.981 (0.276)	0.990 (0.309)	0.993 (0.319)	0.991 (0.317)
AR(1)/Lag $\rho = 0.49$	$\hat{\beta}_1$	0.3006 (0.066)	0.946 (0.256)	0.930 (0.246)	0.936 (0.250)	0.939 (0.252)	0.985 (0.337)	0.978 (0.325)	0.981 (0.337)	0.981 (0.334)
	$\hat{\beta}_2$	0.1978 (0.067)	0.952 (0.263)	0.931 (0.250)	0.938 (0.255)	0.940 (0.256)	0.990 (0.345)	0.980 (0.331)	0.985 (0.342)	0.985 (0.339)

Table 3. Coverage probabilities of regression estimates for over-dispersion of count data with stationary covariates for the independent, AR(1), EQC, and MA(1) models.

True model/ Working correlation	Estimate	Statistic	Coverage Probabilities							
			95% level				99% level			
			GEE	EL	EEL	AEL	GEE	EL	EEL	AEL
IND/IND $\rho = 0.50$	$\hat{\beta}_1$	0.2978 (0.074)	0.808 (0.191)	0.937 (0.273)	0.941 (0.278)	0.941 (0.280)	0.902 (0.252)	0.985 (0.361)	0.989 (0.374)	0.988 (0.370)
	$\hat{\beta}_2$	0.1980 (0.071)	0.813 (0.188)	0.934 (0.265)	0.937 (0.270)	0.937 (0.271)	0.919 (0.247)	0.979 (0.349)	0.985 (0.361)	0.983 (0.358)
AR(1)/AR(1) $\rho = 0.50$	$\hat{\beta}_1$	0.2974 (0.080)	0.876 (0.276)	0.916 (0.314)	0.926 (0.317)	0.931 (0.319)	0.971 (0.363)	0.978 (0.410)	0.982 (0.425)	0.982 (0.421)
	$\hat{\beta}_2$	0.2016 (0.086)	0.898 (0.282)	0.924 (0.316)	0.929 (0.323)	0.932 (0.324)	0.973 (0.370)	0.983 (0.417)	0.987 (0.432)	0.986 (0.428)
AR(1)/Lag $\rho = 0.50$	$\hat{\beta}_1$	0.2916 (0.085)	0.899 (0.282)	0.928 (0.310)	0.930 (0.316)	0.931 (0.318)	0.973 (0.371)	0.978 (0.408)	0.983 (0.423)	0.983 (0.419)
	$\hat{\beta}_2$	0.1978 (0.088)	0.952 (0.288)	0.931 (0.315)	0.938 (0.321)	0.940 (0.323)	0.990 (0.378)	0.980 (0.415)	0.985 (0.430)	0.985 (0.426)
EQC/EQC $\rho = 0.50$	$\hat{\beta}_1$	0.2960 (0.091)	0.903 (0.305)	0.917 (0.336)	0.922 (0.342)	0.926 (0.344)	0.967 (0.401)	0.975 (0.443)	0.980 (0.459)	0.977 (0.455)
	$\hat{\beta}_2$	0.2000 (0.094)	0.892 (0.311)	0.905 (0.339)	0.912 (0.346)	0.913 (0.347)	0.963 (0.409)	0.970 (0.447)	0.979 (0.463)	0.977 (0.458)
EQC/Lag $\rho = 0.50$	$\hat{\beta}_1$	0.2957 (0.091)	0.901 (0.305)	0.917 (0.335)	0.923 (0.341)	0.924 (0.343)	0.967 (0.400)	0.974 (0.441)	0.978 (0.457)	0.976 (0.453)
	$\hat{\beta}_2$	0.2001 (0.095)	0.894 (0.311)	0.905 (0.338)	0.912 (0.344)	0.912 (0.346)	0.961 (0.409)	0.974 (0.445)	0.978 (0.461)	0.977 (0.457)
MA(1)/MA(1) $\rho = 0.50$	$\hat{\beta}_1$	0.2980 (0.080)	0.859 (0.234)	0.928 (0.289)	0.933 (0.295)	0.935 (0.297)	0.949 (0.307)	0.980 (0.382)	0.983 (0.395)	0.983 (0.392)
	$\hat{\beta}_2$	0.1987 (0.080)	0.861 (0.239)	0.915 (0.289)	0.917 (0.295)	0.917 (0.296)	0.938 (0.314)	0.975 (0.381)	0.982 (0.395)	0.981 (0.391)
MA(1)/Lag $\rho = 0.50$	$\hat{\beta}_1$	0.2975 (0.080)	0.897 (0.257)	0.922 (0.288)	0.935 (0.294)	0.936 (0.296)	0.967 (0.337)	0.979 (0.380)	0.982 (0.394)	0.981 (0.391)
	$\hat{\beta}_2$	0.1986 (0.080)	0.891 (0.262)	0.915 (0.288)	0.920 (0.294)	0.920 (0.295)	0.962 (0.344)	0.973 (0.380)	0.984 (0.393)	0.981 (0.390)

Table 4. Coverage probabilities of regression estimates for continuous data with stationary covariates for the independent, AR(1), EQC, and MA(1) models.

True model/ Working correlation	Estimate	Statistic	Coverage Probabilities							
			95% level				99% level			
			GEE	EL	EEL	AEL	GEE	EL	EEL	AEL
IND/IND $\rho = 0.50$	$\hat{\beta}_1$	0.4010 (0.050)	0.955 (0.197)	0.947 (0.191)	0.950 (0.195)	0.951 (0.196)	0.989 (0.259)	0.988 (0.254)	0.988 (0.262)	0.988 (0.260)
	$\hat{\beta}_2$	0.5015 (0.051)	0.953 (0.197)	0.936 (0.191)	0.942 (0.195)	0.942 (0.196)	0.993 (0.259)	0.985 (0.253)	0.990 (0.262)	0.989 (0.260)
AR(1)/AR(1) $\rho = 0.50$	$\hat{\beta}_1$	0.4019 (0.071)	0.949 (0.278)	0.944 (0.270)	0.950 (0.276)	0.952 (0.277)	0.990 (0.365)	0.989 (0.358)	0.990 (0.371)	0.990 (0.368)
	$\hat{\beta}_2$	0.5027 (0.068)	0.959 (0.278)	0.948 (0.270)	0.951 (0.275)	0.952 (0.277)	0.994 (0.365)	0.993 (0.358)	0.993 (0.371)	0.993 (0.367)
AR(1)/Lag $\rho = 0.50$	$\hat{\beta}_1$	0.4018 (0.071)	0.946 (0.277)	0.944 (0.270)	0.949 (0.275)	0.950 (0.277)	0.990 (0.364)	0.988 (0.357)	0.989 (0.370)	0.989 (0.367)
	$\hat{\beta}_2$	0.5026 (0.077)	0.959 (0.310)	0.948 (0.302)	0.950 (0.308)	0.953 (0.310)	0.994 (0.408)	0.992 (0.401)	0.994 (0.415)	0.994 (0.411)
EQC/EQC $\rho = 0.50$	$\hat{\beta}_1$	0.4016 (0.079)	0.949 (0.310)	0.942 (0.302)	0.949 (0.308)	0.951 (0.310)	0.993 (0.408)	0.986 (0.400)	0.988 (0.415)	0.987 (0.411)
	$\hat{\beta}_2$	0.5023 (0.091)	0.963 (0.307)	0.947 (0.332)	0.953 (0.338)	0.954 (0.340)	0.995 (0.404)	0.989 (0.437)	0.994 (0.453)	0.992 (0.448)
EQC/Lag $\rho = 0.50$	$\hat{\beta}_1$	0.4015 (0.079)	0.948 (0.310)	0.945 (0.301)	0.948 (0.307)	0.949 (0.309)	0.992 (0.407)	0.986 (0.399)	0.988 (0.414)	0.987 (0.410)
	$\hat{\beta}_2$	0.5024 (0.077)	0.961 (0.310)	0.947 (0.301)	0.953 (0.308)	0.954 (0.309)	0.996 (0.407)	0.989 (0.400)	0.994 (0.414)	0.994 (0.410)
MA(1)/MA(1) $\rho = 0.50$	$\hat{\beta}_1$	0.3981 (0.065)	0.955 (0.254)	0.945 (0.247)	0.955 (0.252)	0.954 (0.253)	0.992 (0.334)	0.986 (0.327)	0.991 (0.338)	0.991 (0.335)
	$\hat{\beta}_2$	0.4974 (0.062)	0.958 (0.254)	0.944 (0.246)	0.948 (0.251)	0.951 (0.252)	0.995 (0.334)	0.989 (0.327)	0.992 (0.338)	0.990 (0.335)
MA(1)/Lag $\rho = 0.50$	$\hat{\beta}_1$	0.3981 (0.065)	0.956 (0.252)	0.942 (0.246)	0.951 (0.251)	0.953 (0.253)	0.994 (0.332)	0.986 (0.326)	0.989 (0.338)	0.987 (0.335)
	$\hat{\beta}_2$	0.4974 (0.062)	0.957 (0.252)	0.943 (0.246)	0.948 (0.251)	0.950 (0.252)	0.994 (0.332)	0.989 (0.326)	0.991 (0.338)	0.991 (0.335)

- AR(1): $\epsilon_{it} = \rho\epsilon_{i,t-1} + a_{it}, t = 1, 2, 3, 4,$
- EQC: $\epsilon_{it} = \rho a_{i,0} + a_{it}, t = 1, 2, 3, 4,$
- MA(1): $\epsilon_{it} = \rho a_{i,t-1} + a_{it}, t = 1, 2, 3, 4.$

However, the confidence regions for the GEE are constructed under the normality assumption.

Table 6 gives the mean estimated values of the coefficients and the corresponding simulated standard errors in parentheses. It also includes the coverage probability for β_1 and β_2 for the 0.95 and 0.99 confidence levels and the average width of the CI in parentheses for samples of sizes $k = 50$ and $k = 100$ for the independent, AR(1), EQC, and MA(1) models with stationary covariates. **Table 7** gives the results for nonstationary covariates.

When the model is misspecified, the EL, EEL, and AEL outperform the GEE. For example, in the AR(1)/Lag case in **Table 6** the coverage probabilities of $\hat{\beta}_1$ based on the GEE, EL, EEL, and AEL are 0.790, 0.918, 0.931, and 0.932 respectively for the nominal 0.95 level. For $\hat{\beta}_2$, these probabilities are 0.801, 0.924, 0.937, and 0.937 for the same nominal level. Note that we do not need to estimate a scale parameter in the construction of the CI in the EL setup, and also in the EL we did not model the over-dispersion. **Table 7** shows that when the covariates are time-dependent the GEE has substantial undercoverage compared with the results for time-independent covariates, as discussed by [9].

6. Applications

In this section, we illustrate the applicability of our proposed method to two real-world examples.

6.1. Health Care Utilization Study

We consider longitudinal health care utilization data [5] that was collected by Eastern Health, St. John's, Newfoundland, Canada. These longitudinal count data contain complete records for $k = 144$ individuals for the $m = 4$ years from 1985 to 1988. The response of interest was the number of visits to a physician by each individual during a given year. Information on four covariates, namely, gender, number of chronic conditions, education level, and age, was recorded for each individual. Background information allows us to assume that the response variable, marginally, follows the Poisson distribution, and the repeated counts over the four years will be longitudinally correlated. Since the data indicate over-dispersion, we consider a negative binomial model with two variance functions

$$\text{var}(y) = \mu + \alpha\mu$$

and

$$\text{var}(y) = \mu + \alpha\mu^2.$$

Thus, the variance function is different from that of the Poisson model, $\text{var}(y) = \mu$. To confirm the over-dispersion, we test $H_0 : \alpha = 0$ against

Table 5. Coverage probabilities of regression estimates for continuous data with nonstationary covariates for the independent, AR(1), EQC, and MA(1) models.

True model/ Working correlation	Estimate	Statistic	Coverage Probabilities							
			95% level				99% level			
			GEE	EL	EEL	AEL	GEE	EL	EEL	AEL
IND/IND $\rho = 0.50$	$\hat{\beta}_1$	0.4003 (0.051)	0.947 (0.196)	0.940 (0.194)	0.945 (0.198)	0.945 (0.199)	0.986 (0.258)	0.986 (0.257)	0.988 (0.266)	0.988 (0.263)
	$\hat{\beta}_2$	0.4996 (0.051)	0.948 (0.196)	0.946 (0.195)	0.949 (0.199)	0.949 (0.199)	0.985 (0.258)	0.986 (0.258)	0.988 (0.267)	0.988 (0.264)
AR(1)/AR(1) $\rho = 0.50$	$\hat{\beta}_1$	0.3998 (0.043)	0.939 (0.160)	0.926 (0.158)	0.936 (0.161)	0.937 (0.162)	0.989 (0.211)	0.987 (0.209)	0.989 (0.216)	0.989 (0.214)
	$\hat{\beta}_2$	0.4980 (0.040)	0.959 (0.160)	0.949 (0.159)	0.954 (0.162)	0.955 (0.163)	0.992 (0.211)	0.992 (0.210)	0.993 (0.218)	0.993 (0.215)
AR(1)/Lag $\rho = 0.50$	$\hat{\beta}_1$	0.3998 (0.043)	0.935 (0.159)	0.923 (0.157)	0.932 (0.160)	0.934 (0.161)	0.988 (0.209)	0.986 (0.208)	0.989 (0.216)	0.988 (0.214)
	$\hat{\beta}_2$	0.4979 (0.040)	0.958 (0.159)	0.946 (0.158)	0.955 (0.161)	0.955 (0.162)	0.991 (0.209)	0.993 (0.209)	0.994 (0.217)	0.993 (0.215)
EQC/EQC $\rho = 0.50$	$\hat{\beta}_1$	0.4018 (0.040)	0.954 (0.155)	0.947 (0.153)	0.955 (0.156)	0.956 (0.157)	0.992 (0.204)	0.988 (0.203)	0.989 (0.210)	0.989 (0.208)
	$\hat{\beta}_2$	0.5001 (0.040)	0.953 (0.155)	0.947 (0.153)	0.952 (0.156)	0.954 (0.157)	0.989 (0.204)	0.989 (0.203)	0.991 (0.210)	0.990 (0.208)
EQC/Lag $\rho = 0.50$	$\hat{\beta}_1$	0.4019 (0.040)	0.947 (0.154)	0.945 (0.153)	0.950 (0.156)	0.951 (0.156)	0.988 (0.202)	0.985 (0.202)	0.990 (0.209)	0.990 (0.207)
	$\hat{\beta}_2$	0.5011 (0.040)	0.946 (0.154)	0.938 (0.153)	0.944 (0.156)	0.945 (0.157)	0.988 (0.202)	0.987 (0.202)	0.989 (0.209)	0.989 (0.207)
MA(1)/MA(1) $\rho = 0.50$	$\hat{\beta}_1$	0.4000 (0.043)	0.942 (0.165)	0.936 (0.166)	0.946 (0.169)	0.947 (0.170)	0.992 (0.217)	0.993 (0.220)	0.995 (0.227)	0.993 (0.225)
	$\hat{\beta}_2$	0.5027 (0.044)	0.938 (0.165)	0.931 (0.165)	0.939 (0.168)	0.939 (0.169)	0.987 (0.217)	0.979 (0.218)	0.986 (0.226)	0.983 (0.223)
MA(1)/Lag $\rho = 0.50$	$\hat{\beta}_1$	0.4004 (0.038)	0.926 (0.138)	0.926 (0.136)	0.931 (0.139)	0.932 (0.140)	0.982 (0.181)	0.976 (0.180)	0.983 (0.187)	0.981 (0.185)
	$\hat{\beta}_2$	0.5001 (0.035)	0.943 (0.138)	0.950 (0.137)	0.957 (0.140)	0.957 (0.141)	0.992 (0.181)	0.988 (0.182)	0.990 (0.188)	0.990 (0.187)

Table 6. Coverage probabilities of regression estimates for misspecified data with stationary covariates for the independent, AR(1), EQC, and MA(1) models ($k = 50$).

True model/ Working correlation	Estimate	Statistic	Coverage Probabilities							
			95% level				99% level			
			GEE	EL	EEL	AEL	GEE	EL	EEL	AEL
IND/IND $\rho = 0.50$	$\hat{\beta}_1$	0.3971 (0.103)	0.836 (0.281)	0.906 (0.357)	0.909 (0.367)	0.913 (0.369)	0.930 (0.369)	0.963 (0.479)	0.975 (0.512)	0.971 (0.501)
	$\hat{\beta}_2$	0.4988 (0.101)	0.838 (0.281)	0.904 (0.350)	0.926 (0.370)	0.926 (0.372)	0.928 (0.369)	0.966 (0.473)	0.978 (0.513)	0.976 (0.502)
AR(1)/AR(1) $\rho = 0.50$	$\hat{\beta}_1$	0.3954 (0.157)	0.805 (0.397)	0.923 (0.587)	0.934 (0.610)	0.934 (0.613)	0.910 (0.522)	0.971 (0.783)	0.973 (0.827)	0.977 (0.814)
	$\hat{\beta}_2$	0.4986 (0.154)	0.807 (0.397)	0.935 (0.593)	0.942 (0.610)	0.944 (0.613)	0.913 (0.522)	0.975 (0.784)	0.983 (0.827)	0.980 (0.815)
AR(1)/Lag $\rho = 0.50$	$\hat{\beta}_1$	0.3949 (0.156)	0.790 (0.391)	0.918 (0.581)	0.931 (0.601)	0.932 (0.604)	0.902 (0.513)	0.971 (0.772)	0.980 (0.816)	0.977 (0.803)
	$\hat{\beta}_2$	0.4983 (0.152)	0.801 (0.391)	0.924 (0.581)	0.937 (0.602)	0.937 (0.604)	0.911 (0.513)	0.977 (0.714)	0.986 (0.818)	0.983 (0.805)
EQC/EQC $\rho = 0.50$	$\hat{\beta}_1$	0.3970 (0.143)	0.885 (0.444)	0.927 (0.570)	0.940 (0.590)	0.943 (0.593)	0.953 (0.584)	0.981 (0.759)	0.989 (0.803)	0.987 (0.790)
	$\hat{\beta}_2$	0.4935 (0.140)	0.893 (0.444)	0.943 (0.565)	0.955 (0.585)	0.955 (0.588)	0.960 (0.584)	0.984 (0.753)	0.991 (0.798)	0.990 (0.785)
EQC/Lag $\rho = 0.50$	$\hat{\beta}_1$	0.3973 (0.143)	0.811 (0.371)	0.928 (0.568)	0.935 (0.588)	0.941 (0.591)	0.902 (0.487)	0.981 (0.756)	0.989 (0.800)	0.987 (0.787)
	$\hat{\beta}_2$	0.4940 (0.140)	0.812 (0.371)	0.941 (0.562)	0.953 (0.583)	0.956 (0.586)	0.928 (0.487)	0.986 (0.750)	0.990 (0.795)	0.989 (0.782)
MA(1)/MA(1) $\rho = 0.50$	$\hat{\beta}_1$	0.3996 (0.144)	0.802 (0.363)	0.935 (0.563)	0.941 (0.584)	0.942 (0.587)	0.897 (0.477)	0.976 (0.750)	0.984 (0.796)	0.981 (0.782)
	$\hat{\beta}_2$	0.4974 (0.139)	0.808 (0.363)	0.949 (0.560)	0.955 (0.580)	0.955 (0.583)	0.917 (0.477)	0.985 (0.747)	0.991 (0.792)	0.989 (0.779)
MA(1)/Lag $\rho = 0.50$	$\hat{\beta}_1$	0.3997 (0.144)	0.802 (0.366)	0.932 (0.561)	0.943 (0.582)	0.943 (0.585)	0.906 (0.481)	0.977 (0.748)	0.983 (0.794)	0.982 (0.780)
	$\hat{\beta}_2$	0.4980 (0.138)	0.808 (0.366)	0.948 (0.558)	0.960 (0.578)	0.960 (0.581)	0.921 (0.481)	0.985 (0.745)	0.991 (0.791)	0.989 (0.777)

Table 7. Coverage probabilities of regression estimates for misspecified data with nonstationary covariates for the independent, AR(1), EQC, and MA(1) models ($k = 100$).

True model/ Working correlation	Estimate	Statistic	Coverage Probabilities							
			95% level				99% level			
			GEE	EL	EEL	AEL	GEE	EL	EEL	AEL
IND/IND $\rho = 0.50$	$\hat{\beta}_1$	0.3992 (0.098)	0.843 (0.279)	0.930 (0.377)	0.943 (0.392)	0.943 (0.393)	0.945 (0.366)	0.983 (0.491)	0.990 (0.544)	0.989 (0.530)
	$\hat{\beta}_2$	0.4936 (0.098)	0.846 (0.279)	0.926 (0.377)	0.934 (0.392)	0.935 (0.394)	0.936 (0.366)	0.979 (0.489)	0.987 (0.544)	0.985 (0.531)
AR(1)/AR(1) $\rho = 0.50$	$\hat{\beta}_1$	0.3949 (0.094)	0.776 (0.221)	0.925 (0.354)	0.930 (0.368)	0.932 (0.369)	0.890 (0.291)	0.972 (0.477)	0.983 (0.512)	0.979 (0.500)
	$\hat{\beta}_2$	0.4982 (0.092)	0.777 (0.221)	0.933 (0.353)	0.942 (0.368)	0.945 (0.369)	0.887 (0.291)	0.984 (0.476)	0.988 (0.511)	0.986 (0.499)
AR(1)/Lag $\rho = 0.50$	$\hat{\beta}_1$	0.3990 (0.094)	0.750 (0.214)	0.919 (0.355)	0.931 (0.369)	0.933 (0.371)	0.870 (0.281)	0.981 (0.477)	0.988 (0.513)	0.985 (0.500)
	$\hat{\beta}_2$	0.4952 (0.096)	0.762 (0.214)	0.917 (0.355)	0.927 (0.369)	0.928 (0.371)	0.864 (0.281)	0.972 (0.478)	0.982 (0.514)	0.979 (0.501)
EQC/EQC $\rho = 0.50$	$\hat{\beta}_1$	0.3980 (0.107)	0.772 (0.256)	0.928 (0.408)	0.938 (0.424)	0.938 (0.426)	0.896 (0.336)	0.979 (0.548)	0.986 (0.588)	0.985 (0.574)
	$\hat{\beta}_2$	0.5018 (0.106)	0.778 (0.256)	0.923 (0.404)	0.934 (0.421)	0.936 (0.422)	0.890 (0.336)	0.980 (0.543)	0.988 (0.583)	0.984 (0.569)
EQC/Lag $\rho = 0.50$	$\hat{\beta}_1$	0.3981 (0.109)	0.751 (0.253)	0.921 (0.405)	0.932 (0.421)	0.934 (0.423)	0.888 (0.332)	0.979 (0.544)	0.983 (0.585)	0.983 (0.570)
	$\hat{\beta}_2$	0.5026 (0.107)	0.761 (0.253)	0.918 (0.402)	0.937 (0.419)	0.938 (0.420)	0.880 (0.332)	0.978 (0.541)	0.986 (0.581)	0.983 (0.566)
MA(1)/MA(1) $\rho = 0.50$	$\hat{\beta}_1$	0.3961 (0.100)	0.709 (0.213)	0.920 (0.369)	0.930 (0.385)	0.931 (0.386)	0.827 (0.274)	0.980 (0.497)	0.989 (0.534)	0.987 (0.521)
	$\hat{\beta}_2$	0.5009 (0.100)	0.728 (0.213)	0.926 (0.369)	0.941 (0.385)	0.941 (0.386)	0.834 (0.274)	0.983 (0.497)	0.991 (0.534)	0.989 (0.522)
MA(1)/Lag $\rho = 0.50$	$\hat{\beta}_1$	0.3972 (0.096)	0.749 (0.222)	0.913 (0.356)	0.929 (0.370)	0.930 (0.372)	0.866 (0.292)	0.981 (0.478)	0.989 (0.514)	0.987 (0.502)
	$\hat{\beta}_2$	0.5002 (0.096)	0.762 (0.222)	0.929 (0.356)	0.938 (0.370)	0.939 (0.372)	0.872 (0.292)	0.981 (0.478)	0.991 (0.515)	0.989 (0.502)

$H_a : \alpha > 0$ using the likelihood ratio test. The result confirms the presence of over-dispersion in both variance function models.

Our analysis used the GEE with a working correlation matrix (AR(1), EQC, MA(1), or lag correlation) and our EL approach. **Table 8** gives the regression parameter estimates and 95% CIs. The gender covariate was coded as 1 for male and 0 for female. Under the AR(1) structure, the estimate of its regression coefficient is $\hat{\beta}_1 = -0.1929$, suggesting that females make more visits to physicians. The GEE CI indicates that this variable is significant, but the EL CI does not. The estimated values $\hat{\beta}_2 = 0.1668$ and $\hat{\beta}_4 = 0.0308$ suggest that individuals with chronic diseases and older individuals pay more visits to physicians, as expected. The corresponding CIs show that both variables are significant. The education covariate was coded as 1 for less than high school and 0 for higher education. The value $\hat{\beta}_3 = -0.4738$ indicates that educated individuals pay more visits to physicians, showing that they are more concerned about their health or can afford it. The corresponding CIs show that this variable is significant. **Table 8** shows that different working correlations lead to slightly different parameter estimates, but the overall conclusion remains the same. Since the data indicate over-dispersion, the GEE-based approach may be inefficient, as shown

Table 8. Regression estimates for health care utilization count data.

Covariates	Estimate	95% Confidence Interval	
		GEE	EL
Working Correlation: AR(1)			
Gender effect ($\hat{\beta}_1$)	-0.1929	(-0.313, -0.073)	(-0.421, 0.020)
Chronic effect ($\hat{\beta}_2$)	0.1668	(0.177, 0.216)	(0.094, 0.241)
Education effect ($\hat{\beta}_3$)	-0.4738	(-0.624, -0.324)	(-0.768, -0.180)
Age effect ($\hat{\beta}_4$)	0.0308	(0.029, 0.033)	(0.028, 0.033)
Working Correlation: EQC			
Gender effect ($\hat{\beta}_1$)	-0.1772	(-0.306, -0.048)	(-0.407, 0.034)
Chronic effect ($\hat{\beta}_2$)	0.1681	(0.115, 0.222)	(0.095, 0.237)
Education effect ($\hat{\beta}_3$)	-0.4354	(-0.597, -0.274)	(-0.726, -0.146)
Age effect ($\hat{\beta}_4$)	0.0302	(0.028, 0.033)	(0.027, 0.033)
Working Correlation: MA(1)			
Gender effect ($\hat{\beta}_1$)	-0.1922	(-0.299, -0.086)	(-0.421, 0.021)
Chronic effect ($\hat{\beta}_2$)	0.1669	(0.123, 0.211)	(0.094, 0.241)
Education effect ($\hat{\beta}_3$)	-0.4720	(-0.605, -0.339)	(-0.766, -0.179)
Age effect ($\hat{\beta}_4$)	0.0308	(0.029, 0.033)	(0.028, 0.033)
Working Correlation: Lag			
Gender effect ($\hat{\beta}_1$)	-0.1819	(-0.311, -0.053)	(-0.411, 0.029)
Chronic effect ($\hat{\beta}_2$)	0.1677	(0.114, 0.221)	(0.095, 0.238)
Education effect ($\hat{\beta}_3$)	0.4469	(-0.608, -0.286)	(-0.738, -0.156)
Age effect ($\hat{\beta}_4$)	0.0304	(0.028, 0.033)	(0.027, 0.033)

in our performance analysis. We conclude that the EL approach is more appropriate for this data set, and the significant variables identified by this approach are more reliable.

6.2. Longitudinal CD4 Cell Counts of HIV Seroconverters

This data set contains 2376 observations of the CD4 cell counts of $k = 369$ men infected with the HIV virus [31]. The goal of our analysis is to estimate the average evolution over time of the CD4 counts by considering the effects of AGE, SMOKE (smoking status measured by packs of cigarettes per day), DRUG (yes = 1; no = 0), SEXP (number of sex partners), DEPRESSION (measured by the CESD scale) and YEAR (time since seroconversion). To examine whether there are any interaction effects between the covariates, we included all the two-factor interactions in our model.

This data set has the subject-specific evolution over time of the CD4 cell counts with and without drug use. The cell counts are right-skewed, so the analysis was conducted on square-root transformed CD4 cell counts whose distribution is more nearly Gaussian. Tables 9-11 summarize the analysis for the AR(1), EQC, and lag working correlations. The GEE indicates that SMOKE, DRUG, SEXP, AGE \times SEXP, SMOKE \times DRUG, SMOKE \times SEXP, and DRUG \times SEXP are significant. Under EQC, AGE \times SMOKE and AGE \times DRUG are also significant. The EL selects SMOKE, DRUG, SEXP, and DRUG \times SEX. Under EQC and lag AGE \times SEXP is also significant. The GEE approach is sensitive to the choice of correlation structure. In this real data set, the true correlation structure is unknown, so the lag correlation approach is appropriate since it can accommodate all three correlation structures. The Shapiro-Wilk test shows that the square-root transformed CD4 cell counts are not normally distributed. The GEE-based method is, therefore, not appropriate. We, therefore, conclude that the EL is a better choice.

Table 9. Estimated coefficients for CD4 data set using AR(1) working correlation.

Variable	Method	
	GEE	EL
INTERCEPT	25.37 (25.25, 25.49)	25.37 (24.97, 25.77)
AGE	-0.001 (-0.014, 0.012)	-0.001 (-0.060, 0.060)
SMOKE	0.938 (0.864, 1.012)	0.938 (0.669, 1.211)
DRUG	0.716 (0.597, 0.834)	0.716 (0.316, 1.115)
SEXP	0.390 (0.365, 0.414)	0.390 (0.306, 0.471)
AGE \times SMOKE	-0.001 (-0.007, 0.004)	-0.001 (-0.032, 0.031)
AGE \times DRUG	0.001 (-0.013, 0.013)	0.001 (-0.069, 0.071)
AGE \times SEXP	0.008 (0.006, 0.009)	0.008 (-0.003, 0.016)
SMOKE \times DRUG	-0.242 (-0.315, -0.169)	-0.242 (-0.500, 0.023)
SMOKE \times SEXP	0.041 (0.033, 0.049)	0.041 (-0.005, 0.089)
DRUG \times SEXP	-0.270 (-0.295, -0.245)	-0.270 (-0.358, -0.183)

Table 10. Estimated coefficients for CD4 data set using EQC working correlation.

Variable	Method	
	GEE	EL
INTERCEPT	25.10 (24.96, 25.25)	25.10 (24.70, 25.50)
AGE	-0.023 (-0.037, 0.008)	-0.023 (-0.085, 0.043)
SMOKE	1.241 (1.161, 1.322)	1.241 (0.972, 1.515)
DRUG	1.132 (1.006, 1.257)	1.132 (0.732, 1.532)
SEXP	0.545 (0.521, 0.569)	0.545 (0.459, 0.633)
AGE × SMOKE	-0.011 (-0.016, -0.006)	-0.011 (-0.044, 0.023)
AGE × DRUG	0.036 (0.022, 0.050)	0.036 (-0.031, 0.106)
AGE × SEXP	0.017 (0.015, 0.018)	0.017 (0.007, 0.027)
SMOKE × DRUG	-0.398 (-0.477, -0.319)	-0.398 (-0.650, 0.131)
SMOKE × SEXP	0.038 (0.030, 0.045)	0.038 (-0.010, 0.091)
DRUG × SEXP	-0.184 (-0.209, -0.159)	-0.184 (-0.274, -0.091)

Table 11. Estimated coefficients for CD4 data set using lag working correlation.

Variable	Method	
	GEE	EL
INTERCEPT	25.35 (25.22, 25.46)	25.35 (25.34, 25.36)
AGE	-0.001 (-0.015, 0.012)	-0.001 (-0.061, 0.059)
SMOKE	0.942 (0.867, 1.016)	0.942 (0.673, 1.215)
DRUG	0.727 (0.608, 0.845)	0.727 (0.327, 1.127)
SEXP	0.389 (0.364, 0.414)	0.389 (0.305, 0.470)
AGE × SMOKE	-0.002 (-0.007, 0.003)	-0.002 (-0.032, 0.030)
AGE × DRUG	0.002 (-0.011, 0.015)	0.002 (-0.067, 0.072)
AGE × SEXP	0.007 (0.005, 0.009)	0.007 (-0.003, 0.016)
SMOKE × DRUG	-0.233 (-0.306, -0.160)	-0.233 (-0.490, 0.032)
SMOKE × SEXP	0.042 (0.035, 0.050)	0.042 (-0.004, 0.090)
DRUG × SEXP	-0.268 (-0.356, -0.182)	-0.268 (-0.356, -0.182)

7. Conclusions

Longitudinal data modelling using the GEE approach assumes a working correlation model for the within-subject correlation of the responses. When the working correlation is incorrectly specified, the GEE based estimates are not necessarily consistent and may lose efficiency. Any misspecification can cause estimates based on marginal models to be inefficient and misleading conclusions. Also, the construction of a confidence region and hypothesis testing are based on asymptotic normality, which may not hold since the finite-sample distribution may not be symmetric.

Taking these issues into account, we have proposed an EL-based longitudinal modelling based on a data-driven likelihood ratio approach sharing many of the

properties of the parametric likelihood. We do not need to specify the complete parametric distribution to perform the inference. We can, therefore, use likelihood methods without assuming that the data come from a known family of distributions. We defined the subject-wise profile EL based on a set of GEEs. The estimation and confidence region construction using the EL approach are proposed, which has advantages over other methods such as those based on normal approximations. We introduced the adjusted EL to avoid any computational issues, which improve the coverage probabilities. A major advantage of EI is that involves no prior assumptions about the shape of an EL-based confidence region, which is data-driven. The construction of the confidence region based on the EL method does not involve any variance estimation.

The proposed approach yields more efficient estimators than the conventional GEE approach and achieves the same asymptotic properties as [17]. Our performance analysis showed that our method for longitudinal count and continuous responses is comparable to the GEE when the model assumptions are satisfied. For instance, when the working correlation is correctly specified, the coverage probabilities of the intervals based on the EL, EEL, and AEL are similar to those of the GEE. CIs based on the regular EL have slight undercoverage compared with those of the GEE; the coverage probabilities are substantially improved with the EEL and AEL. Moreover, these methods are consistently more accurate than the regular EL. When the working correlation is misspecified, the coverage probabilities of the intervals based on the EL, EEL, and AEL are shown to be equally efficient to the GEE estimator with stationary lag correlation structure. Also, the results show that when the working correlation is misspecified, the GEE estimator with stationary lag correlation structure, EL, EEL, and AEL outperforms the GEE with an incorrect working correlation structure. When the model is misspecified such as marginal variance, our method outperforms the GEE. This result shows that EL methods are robust to misspecification. Moreover, the EL-based CI has a data-driven shape, whereas the GEE-based CI is always symmetric due to normal approximation.

Acknowledgements

The authors' research was supported by grants from Natural Sciences & Engineering Research Council of Canada and Canadian Institute of Health Research.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

- [1] McCullagh, P. and Nelder, J. (1989) *Generalized Linear Models*. 2nd Edition, CRC Press, London. <https://doi.org/10.1007/978-1-4899-3242-6>
- [2] Liang, K.Y. and Zeger, S.L. (1986) *Longitudinal Data Analysis Using Generalized*

- Linear Models. *Biometrika*, **73**, 13-22. <https://doi.org/10.1093/biomet/73.1.13>
- [3] Crowder, M.J. (1995) On Use of a Working Correlation Matrix in Using Generalized Linear Models for Repeated Measures. *Biometrika*, **82**, 407-410. <https://doi.org/10.1093/biomet/82.2.407>
- [4] Sutradhar, B.C. and Das, K. (1999) On the Efficiency of Regression Estimators in Generalized Linear Models for Longitudinal Data. *Biometrika*, **86**, 459-465. <https://doi.org/10.1093/biomet/86.2.459>
- [5] Sutradhar, B.C. (2003) An Overview on Regression Models for Discrete Longitudinal Responses. *Statistical Science*, **18**, 377-393. <https://doi.org/10.1214/ss/1076102426>
- [6] Sutradhar, B.C. and Kovacevic, M. (2000) Analysing Ordinal Longitudinal Survey Data: Generalized Estimating Equations Approach. *Biometrika*, **87**, 837-848. <https://doi.org/10.1093/biomet/87.4.837>
- [7] Nadarajah, T., Variyath, A.M. and Loredó-Osti, J.C. (2016) Penalized Generalized Quasi-Likelihood Based Variable Selection for Longitudinal Data. In: *ISS-2015 Proceedings Volume on Advances in Parametric and Semiparametric Analysis of Multivariate, Time Series, Spatial-Temporal, and Familial-Longitudinal Data*, Springer Lecture Notes in Statistics, Springer, Berlin, 233-250. https://doi.org/10.1007/978-3-319-31260-6_8
- [8] Qu, A., Lindsay, B.G. and Li, B. (2000) Improving Estimating Equations Using Quadratic Inference Functions. *Biometrika*, **87**, 823-836. <https://doi.org/10.1093/biomet/87.4.823>
- [9] Hu, F.-C. (1993) A Statistical Methodology for Analyzing the Causal Health Effect of a Time Dependent Exposure from Longitudinal Data. Dissertation, Harvard School of Public Health, Department of Biostatistics, Boston.
- [10] Pepe, M.S. and Anderson, G.L. (1994) A Cautionary Note on Inference for Marginal Regression Models with Longitudinal Data and General Correlated Response Data. *Communications in Statistics, Part B Simulation and Computation*, **23**, 939-951. <https://doi.org/10.1080/03610919408813210>
- [11] Emond, M.J., Ritz, J. and Oakes, D. (1997) Bias in GEE Estimates from Misspecified Models for Longitudinal Data. *Communications in Statistics*, **26**, 15-32. <https://doi.org/10.1080/03610929708831899>
- [12] Pan, W., Louis, T.A. and Connett, J.E. (2000) A Note on Marginal Linear Regression with Correlated Response Data. *American Statistician*, **54**, 191-195. <https://doi.org/10.1080/00031305.2000.10474544>
- [13] Diggle, P.J., Heagerty, P.J., Liang, K.-Y. and Zeger, S.L. (2002) *The Analysis of Longitudinal Data*. 2nd Edition, Oxford Statistical Science, Oxford University Press, Oxford.
- [14] Owen, A.B. (1988) Empirical Likelihood Ratio Confidence Interval for a Single Functional. *Biometrika*, **75**, 237-249. <https://doi.org/10.1093/biomet/75.2.237>
- [15] Owen, A.B. (1991) Empirical Likelihood for Linear Models. *The Annals of Statistics*, **19**, 1725-1747. <https://doi.org/10.1214/aos/1176348368>
- [16] Chen, S.X. (1994) Empirical Likelihood Confidence Intervals for Linear Regression Coefficients. *Journal of Multivariate Analysis*, **49**, 24-40. <https://doi.org/10.1006/jmva.1994.1011>
- [17] Qin, J. and Lawless, J. (1994) Empirical Likelihood and General Estimating Equations. *Annals of Statistics*, **22**, 300-325. <https://doi.org/10.1214/aos/1176325370>
- [18] Owen, A.B. (2001) *Empirical Likelihood*. 2nd Edition, CRC Press, London.

- <https://doi.org/10.1201/9781420036152>
- [19] You, J., Chen, G. and Zhou, Y. (2006) Block Empirical Likelihood for Longitudinal Partially Linear Regression Models. *Canadian Journal of Statistics*, **34**, 79-96. <https://doi.org/10.1002/cjs.5550340107>
- [20] Xue, L.G. and Zhu, L.X. (2007) Empirical Likelihood Semiparametric Regression Analysis for Longitudinal Data. *Biometrika*, **94**, 921-937. <https://doi.org/10.1093/biomet/asm066>
- [21] Wang, S., Qian, L. and Carroll, J.R. (2010) Generalized Empirical Likelihood Methods for Analyzing Longitudinal Data. *Biometrics*, **97**, 79-93. <https://doi.org/10.1093/biomet/asp073>
- [22] Hall, P. and La Scala, B. (1990) Methodology and Algorithm of Empirical Likelihood. *International Statistical Review*, **58**, 109-127. <https://doi.org/10.2307/1403462>
- [23] Corcoran, S.A., Davison, A.C. and Spady, R.H. (1995) Reliable Inference from Empirical Likelihood. Technical Report, Nuffield College, University of Oxford, Oxford.
- [24] Tsao, M. (2004) Bounds on Coverage Probabilities of the Empirical Likelihood Ratio Confidence Regions. *The Annals of Statistics*, **32**, 1215-1221. <https://doi.org/10.1214/009053604000000337>
- [25] Chen, J., Variyath, A.M. and Abraham, B. (2008) Adjusted Empirical Likelihood and Its Properties. *Journal of Computational Graphics and Statistics*, **17**, 426-443. <https://doi.org/10.1198/106186008X321068>
- [26] Chen, J., Sitter, R.R. and Wu, C. (2002) Using Empirical Likelihood Methods to Obtain Range Restricted Weights in Regression Estimators for Surveys. *Biometrika*, **89**, 230-237. <https://doi.org/10.1093/biomet/89.1.230>
- [27] Tsao, M. (2013) Extending the Empirical Likelihood by Domain Expansion. *Canadian Journal of Statistics*, **41**, 257-274. <https://doi.org/10.1002/cjs.11175>
- [28] Tsao, M. and Wu, F. (2013) Empirical Likelihood on the Full Parameter Space. *Revised for the Annals of Statistics*, **41**, 2176-2196. <https://doi.org/10.1214/13-AOS1143>
- [29] McKenzie, E. (1988) Some ARMA Models for Dependent Sequences of Poisson Counts. *Advances in Applied Probability*, **20**, 822-835. <https://doi.org/10.1017/S0001867800018395>
- [30] Sutradhar, B.C. (2011) Dynamic Mixed Models for Familial Longitudinal Data. Springer Series in Statistics, Springer, Berlin. <https://doi.org/10.1007/978-1-4419-8342-8>
- [31] Zeger, L.S. and Diggle, J.P. (1994) Semiparametric Models for Longitudinal Data with Application to CD4 Cell Numbers in HIV Seroconverters. *Biometrics*, **50**, 689-699. <https://doi.org/10.2307/2532783>