TEST
https://doi.org/10.1007/s11749-019-00675-1

ORIGINAL PAPER

Analysis of correlated Birnbaum–Saunders data based on estimating equations

Aline B. Tsuyuguchi[1] · Gilberto A. Paula[2] · Michelli Barros[3]

Received: 9 October 2017 / Accepted: 14 July 2019

Abstract

Estimating equations for analyzing correlated Birnbaum–Saunders (BS) data are derived in this paper. A regression model is proposed for modeling the median of the life time until the failure, and a reweighted iterative process is developed for the joint estimation of the regression coefficients and the shape and correlation parameters. Diagnostic procedures, such as residual analysis and sensitivity studies based on case deletion and local influence, are given. Simulation studies are performed to assess the empirical distributions of the derived estimators and of a Pearson-type residual for correlated data. Finally, a longitudinal data set is analyzed by the procedures developed in the paper and extensions for the double case in which the median and the shape parameter are jointly modeled are discussed.

Keywords Skewness · Asymmetric data · Birnbaum–Saunders distribution · Correlated data · Diagnostic procedures · Estimating equations

Mathematics Subject Classification 60E05 · 62N99 · 62G10

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s11749-019-00675-1) contains supplementary material, which is available to authorized users.

✉ Gilberto A. Paula
giapaula@ime.usp.br

Aline B. Tsuyuguchi
aline@de.ufpe.br

Michelli Barros
michelli.karinne@ufcg.edu.br

1 Departamento de Estatística, Universidade Federal de Pernambuco, Recife, Brazil

2 Instituto de Matemática e Estatística, Universidade de São Paulo, São Paulo, Brazil

3 Departamento de Estatística, Universidade Federal de Campina Grande, Campina Grande, Brazil

Published online: 20 July 2019

1 Introduction

The BS distribution is a well-known failure time distribution originally developed for analyzing fatigue in materials (Birnbaum and Saunders 1969), particularly in areas where the failure time is related to some cumulative damage (see, for instance, Leiva 2016 and the references therein), and has been largely applied in the last decade for analyzing positive continuous data. The majority of the works developed under BS models assume independence among the experimental units, and few have been developed for analyzing correlated BS data. The correlated data include different data structures, such as multivariate observations, clustered data, repeated measurements, longitudinal data and spatially correlated data. Multivariate BS data have received special attention recently. For example, Kundu et al. (2013) derived the multivariate version of the BS distribution, Marchant et al. (2016) proposed a multivariate log-linear BS model, whereas Marchant et al. (2016) presented diagnostic procedures in multivariate generalized BS regression models. For repeated measurement and longitudinal BS data analysis, Villegas et al. (2011) proposed linear mixed models. A recent review on BS and related distributions is presented by Balakrishnan and Kundu (2019). However, to our knowledge, no study on correlated BS data based on estimating equations has been considered.

Under the approach of estimating equations, the distribution of the multivariate responses does not need to be specified. It is enough to know the first two moments of the marginal distributions and the correlation structure within each experimental unit. Under mild regularity conditions, consistent estimators may be derived for the regression coefficients as well as for their asymptotic variance–covariance matrix, even when this structure is misspecified. In addition, estimating equations may be preferred when the interest is on comparing marginal expected responses (see Alencar et al. 2012).

The aim of this paper is to propose an alternative approach for analyzing correlated BS data, such as clustered data, repeated-measures and longitudinal data, based on estimating equations. From the optimum class of estimating functions proposed by Crowder (1987) (see also, Godambe 1997; Artes and Jorgensen 2000), we derive an optimum class of estimating equations for analyzing correlated data in which the marginal distributions are assumed BS. A reweighted iterative process is developed for the parameter estimation, and the asymptotic and empirical properties of the derived estimators are discussed. Diagnostic procedures are proposed, and an application with a real data set is given. Extensions for the double case, in which the median is jointly modeled with the shape parameter, are also discussed.

The paper is organized as follows. In Sect. 2, a brief review on the BS distribution is given. A regression model for modeling the median of the life time until the failure of correlated BS data is proposed in Sect. 3, and an optimum class of estimating functions for the regression coefficients is derived. A joint iterative process for the parameter estimation is proposed, and some discussions on the asymptotic properties of the derived estimators are given. Simulation studies to assess the empirical distribution of the estimators are also performed. In Sect. 4, some diagnostic procedures, such as a Pearson-type residual for correlated data and some influence measures based on case deletion and local influence, are derived. The empirical distribution of the proposed residual is performed by a simulation study. An application with a real data for illustrating the methodology developed in the paper is presented in Sect. 5. Section 6 deals with the extension for the double case in which the median and the shape parameter are jointly modeled. Some conclusions are presented in the last section, whereas proofs and derivations are given in the supplementary material.

2 The BS distribution

Denote by t the life time until the failure which is assumed to follow the two-parameter BS distribution, denoted by $t \sim \mathrm{BS}(\alpha, \eta)$. The cumulative distribution function (CDF) of t is written as

$$F(t; \alpha, \eta) = \Phi\left[\frac{1}{\alpha}\left\{\left(\frac{t}{\eta}\right)^{\frac{1}{2}} - \left(\frac{\eta}{t}\right)^{\frac{1}{2}}\right\}\right], \quad t, \alpha, \eta > 0,$$

where $\Phi(\cdot)$ denotes the CDF of the standard normal distribution, η is the scale parameter, the median of t, and α is the shape parameter. In addition, one has that $\mathrm{E}(t) = \eta(\alpha^2 + 2)/2$ and $\mathrm{Var}(t) = \eta^2(5\alpha^4 + 4\alpha^2)/4$ (see, for instance, Leiva et al. 2007). One may write $t = (\eta/4)\{\alpha z + \sqrt{(\alpha z)^2 + 4}\}^2$, where $z \sim \mathrm{N}(0, 1)$, so that the qth quantile of t takes the form

$$\upsilon(q) = \frac{\eta}{4}\left\{\alpha z_q + \sqrt{(\alpha z_q)^2 + 4}\right\}^2,$$

where z_q denotes the qth quantile of $\mathrm{N}(0, 1)$. The above expression may be useful for estimating the quantiles from the estimates of η and α.

The BS distribution has been related to some continuous distributions. For example, $y = \log(t)$ is a symmetric distribution that is a particular case of the SN (sinh-normal) distribution (Rieck and Nedelman 1991) whose probability density function is given by

$$f(y; \alpha, \mu) = \frac{1}{\alpha\sqrt{2\pi}} \cosh\left(\frac{y-\mu}{2}\right) \exp\left[-2\left\{\frac{1}{\alpha}\sinh\left(\frac{y-\mu}{2}\right)\right\}^2\right],$$

where $y, \mu \in \mathbb{R}, \alpha > 0$, $\mathrm{E}(y) = \mu = \log(\eta)$ and $\mathrm{Var}(y) = 4g(\alpha)$ with $g(\alpha)$ being a function based on the modified Bessel function of third order. We will denote $y \sim \log\text{-BS}(\alpha, \eta)$. It is usual to derive the parameter estimates of the BS distribution from the log-BS distribution.

3 Median modeling

Let $\mathbf{t}_i = (t_{i1}, \ldots, t_{is_i})^\top$ be an $s_i \times 1$ vector containing positive responses for the ith experimental unit, for $i = 1, \ldots, n$. We will assume that $t_{ij} \sim \mathrm{BS}(\alpha, \eta_{ij})$ with regression structure $\log(\eta_{ij}) = \mathbf{x}_{ij}^\top \boldsymbol{\beta}$, where $\mathbf{x}_{ij} = (x_{ij1}, \ldots, x_{ijp})^\top$ contains values of explanatory variables and $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_p)^\top$ is the regression coefficient vector. Equivalently one has the following regression model: (i) $y_{ij}|\mathbf{x}_{ij} \sim$ log-BS(α, μ_{ij}) with (ii) $\mu_{ij} = \mathbf{x}_{ij}^\top \boldsymbol{\beta}$, for $i = 1, \ldots, n$ and $j = 1, \ldots, s_i$.

3.1 Extension of GEE

In order to incorporate the dependence among the responses within each experimental unit into the previous regression model, one may extend the GEE (Liang and Zeger 1986) by considering an optimum estimating function for $\boldsymbol{\beta}$ (see definition in Sect. S.1) as

$$\begin{aligned}\boldsymbol{\Psi}^*(\beta) &= \sum_{i=1}^n \mathrm{E}\left(\frac{\partial \mathbf{u}_i}{\partial \boldsymbol{\beta}^\top}\right)^\top \mathrm{Cov}(\mathbf{u}_i)^{-1}\mathbf{u}_i \\ &= \sum_{i=1}^n \mathbf{X}_i^\top \mathbf{N}_i \mathrm{Cov}(\mathbf{u}_i)^{-1}\mathbf{u}_i, \qquad (1)\end{aligned}$$

where $\mathbf{u}_i = (u_{i1}, \ldots, u_{is_i})^\top$, $u_{ij} = u_i(y_{ij}; \boldsymbol{\beta})$, $\mathbf{X}_i$ is an $s_i \times p$ matrix of rows $\mathbf{x}_{ij}^\top$, $\mathbf{N}_i = \mathrm{E}\left(\partial \mathbf{u}_i / \partial \boldsymbol{\mu}_i\right) = \mathrm{diag}\{\mathrm{E}(\dot{u}_{i1}), \ldots, \mathrm{E}(\dot{u}_{is_i})\}$ and $\mathrm{Cov}(\mathbf{u}_i)$ denotes the variance–covariance matrix of $\mathbf{u}_i$, for $i = 1, \ldots, n$ and $j = 1, \ldots, s_i$.

It is required unbiased estimating function, $\mathrm{E}\{\boldsymbol{\Psi}^*(\boldsymbol{\beta})\} = \mathbf{0}$, so we can consider $\mathrm{E}(u_{ij}) = 0, \forall ij$. Then, a candidate for u_{ij} is the score function of μ_{ij} given by

$$\begin{aligned}u_{ij} &= \partial \log\{f(y_{ij}; \alpha, \mu_{ij})\}/\partial \mu_{ij} \\ &= \frac{1}{\alpha^2}\sinh(y_{ij} - \mu_{ij}) - \frac{1}{2}\tanh\left(\frac{y_{ij} - \mu_{ij}}{2}\right),\end{aligned}$$

and from $\mathrm{E}(u_{ij}) = 0$ (see proof in Sect. S.2) it follows that $\mathrm{Var}(u_{ij}) = \mathrm{E}(u_{ij}^2) = -\mathrm{E}(\dot{u}_{ij})$ with

$$\mathrm{E}(\dot{u}_{ij}) = \frac{1}{4}\mathrm{E}\left\{\mathrm{sech}^2\left(\frac{y_{ij} - \mu_{ij}}{2}\right)\right\} - \frac{1}{2} - \frac{1}{\alpha^2}$$

(see proof in Sect. S.3), for $i = 1, \ldots, n$ and $j = 1, \ldots, s_i$.

We may express $\mathrm{Cov}(\mathbf{u}_i) = \boldsymbol{\Sigma}_{u_i}^{\frac{1}{2}} \mathbf{R}(\mathbf{u}_i) \boldsymbol{\Sigma}_{u_i}^{\frac{1}{2}}$, where $\boldsymbol{\Sigma}_{u_i} = \mathrm{diag}\{\mathrm{Var}(u_{i1}), \ldots, \mathrm{Var}(u_{is_i})\} = -\mathbf{N}_i$ and $\mathbf{R}(\mathbf{u}_i)$ denotes the correlation matrix among the elements of $\mathbf{u}_i$. Similarly to Liang and Zeger (1986) we will assume that the Pearson correlation matrix $\mathbf{R}(\mathbf{u}_i)$ has a parametric structure depending on the correlation vector $\boldsymbol{\rho} = (\rho_1, \ldots, \rho_q)^\top$, that does not depend on $\boldsymbol{\beta}$ neither α. Thus, the estimating function for $\boldsymbol{\beta}$ (given α and $\boldsymbol{\rho}$) assumes the form

$$\boldsymbol{\Psi}(\boldsymbol{\beta}) = \sum_{i=1}^n \mathbf{X}_i^\top \mathbf{W}_i \mathbf{N}_i^{-1} \mathbf{u}_i, \qquad (2)$$

where $\mathbf{W}_i = \mathbf{N}_i \mathbf{C}_i^{-1} \mathbf{N}_i$ and $\mathbf{C}_i = \boldsymbol{\Sigma}_{u_i}^{\frac{1}{2}} \mathbf{R}_i(\boldsymbol{\rho}) \boldsymbol{\Sigma}_{u_i}^{\frac{1}{2}}$ with $\mathbf{R}_i(\boldsymbol{\rho})$ being named working correlation matrix of the ith experimental unit, for $i = 1, \ldots, n$.

The model defined by $t_{ij} \sim \mathrm{BS}(\alpha, \eta_{ij})$ such that $\log(\eta_{ij}) = \mathbf{x}_{ij}^\top \boldsymbol{\beta}$ with $\mathbf{R}_i(\boldsymbol{\rho})$ being the working correlation matrix, for $i = 1, \ldots, n$ and $j = 1, \ldots, s_i$, will be named BS-GEE model.

3.2 Iterative process

For obtaining the estimate $\hat{\boldsymbol{\beta}}$, we will apply the Newton scoring method, that is parallel to the Fisher scoring method (see, for instance, Jorgensen et al. 1996). Thus, to solve $\boldsymbol{\Psi}(\hat{\boldsymbol{\beta}}) = \mathbf{0}$ we will apply the Newton algorithm with $\boldsymbol{\Psi}'(\boldsymbol{\beta})$ being replaced by its expectation $\mathrm{E}\{\boldsymbol{\Psi}'(\boldsymbol{\beta})\}$. The iterative process will be alternated with consistent estimates for α and $\boldsymbol{\rho}$. Then, by fixing $(\alpha, \boldsymbol{\rho}^\top)^\top$, and using the expression for $\mathrm{E}\{\boldsymbol{\Psi}'(\boldsymbol{\beta})\}$ from Sect. S.4 we obtain the following iterative process:

$$\begin{aligned}\boldsymbol{\beta}^{(m+1)} &= \boldsymbol{\beta}^{(m)} - [\mathrm{E}\{\boldsymbol{\Psi}'(\boldsymbol{\beta}^{(m)})\}]^{-1} \boldsymbol{\Psi}(\boldsymbol{\beta}^{(m)}) \\ &= \boldsymbol{\beta}^{(m)} - \left\{\sum_{i=1}^{n} \mathbf{X}_i^\top \mathbf{W}_i^{(m)} \mathbf{X}_i\right\}^{-1} \left\{\sum_{i=1}^{n} \mathbf{X}_i^\top \mathbf{W}_i^{(m)} (\mathbf{N}_i^{(m)})^{-1} \mathbf{u}_i^{(m)}\right\},\end{aligned} \tag{3}$$

for $m = 0, 1, 2\ldots$, where $\mathrm{E}\{\boldsymbol{\Psi}'(\boldsymbol{\beta})\} = \mathrm{E}\{\partial \boldsymbol{\Psi}(\boldsymbol{\beta})/\partial \boldsymbol{\beta}^\top\} = \sum_{i=1}^{n} \mathbf{X}_i^\top \mathbf{W}_i \mathbf{X}_i$. The iterative process (3) may be re-expressed in the reweighted form

$$\boldsymbol{\beta}^{(m+1)} = \left\{\sum_{i=1}^{n} \mathbf{X}_i^\top \mathbf{W}_i^{(m)} \mathbf{X}_i\right\}^{-1} \left\{\sum_{i=1}^{n} \mathbf{X}_i^\top \mathbf{W}_i^{(m)} \mathbf{z}_i^{(m)}\right\}, \tag{4}$$

for $m = 0, 1, 2\ldots$, where $\mathbf{z}_i = \mathbf{X}_i \boldsymbol{\beta} - \mathbf{N}_i^{-1} \mathbf{u}_i$ is a modified dependent variable, for $i = 1, \ldots, n$. We will consider for starting value $\boldsymbol{\beta}^{(0)}$ and for α estimate the respective maximum likelihood estimates from the marginal model under independent observations. For instance, from Leiva et al. (2007), the α estimate (for $\boldsymbol{\beta}$ fixed) may be expressed as

$$\hat{\alpha} = \sqrt{\frac{\sum_{i=1}^{n} \sum_{j=1}^{s_i} \left\{4 \sinh^2\left(\frac{y_{ij} - \mu_{ij}}{2}\right)\right\}}{\sum_{i=1}^{n} s_i}}. \tag{5}$$

This estimate may be, alternatively, derived from the estimating equations $\boldsymbol{\Gamma}(\hat{\boldsymbol{\theta}}) = \mathbf{0}$ given in Sect. 6 for the joint modeling of median and shape parameters. Moment estimators will be considered for estimating $\boldsymbol{\rho}$. Below we describe such estimators, by fixing $\boldsymbol{\beta}$ and α, for some usual structures.

Independent

In this case, one has $\mathbf{R}_i = \mathbf{I}_{s_i}$, where $\mathbf{I}_{s_i}$ denotes the identity matrix of order s_i.

Unstructured

Here the correlation matrix $\mathbf{R}_i$ is unstructured and one has $s_i(s_i - 1)/2$ parameters to be estimated for each group. Denoting $\mathbf{R}_i = \{\rho_{ijj'}\}$, the (j, j')th element of $\mathbf{R}_i$ may be estimated by

$$\hat{\rho}_{jj'} = \frac{1}{n}\sum_{i=1}^{n} \frac{u_{ij}}{\sqrt{\mathrm{E}(u_{ij}^2)}} \frac{u_{ij'}}{\sqrt{\mathrm{E}(u_{ij'}^2)}}.$$

Exchangeable

In this case $\mathbf{R}_i = \mathbf{R}_i(\rho)$, where the (j, j')th element of $\mathbf{R}_i$ becomes given by $\mathrm{R}_{ijj'} = 1$, for $j = j'$, and $\mathrm{R}_{ijj'} = \rho$, for $j \neq j'$. A consistent estimator for ρ may expressed as

$$\hat{\rho} = \frac{1}{n}\sum_{i=1}^{n} \frac{1}{s_i(s_i - 1)} \sum_{j=1}^{s_i} \sum_{j'=1, j'\neq j}^{s_i} \frac{u_{ij}}{\sqrt{\mathrm{E}(u_{ij}^2)}} \frac{u_{ij'}}{\sqrt{\mathrm{E}(u_{ij'}^2)}}.$$

First-order autoregressive

Here we assume $\mathbf{R}_i = \mathbf{R}_i(\rho)$, where the (j, j')th element of $\mathbf{R}_i$ becomes given by $\mathrm{R}_{ijj'} = 1$, for $j = j'$, and $\mathrm{R}_{ijj'} = \rho^{|j-j'|}$, for $j \neq j'$. A consistent estimator for ρ may be expressed as

$$\hat{\rho} = \frac{1}{n}\sum_{i=1}^{n} \frac{1}{(s_i - 1)} \sum_{j=1}^{s_i - 1} \frac{u_{ij}}{\sqrt{\mathrm{E}(u_{ij}^2)}} \frac{u_{i(j+1)}}{\sqrt{\mathrm{E}(u_{i(j+1)}^2)}}.$$

Thus, we may propose the following iterative process to obtain the parameter estimates:

1. Give starting values $\alpha^{(0)}$, $\boldsymbol{\beta}^{(0)}$ and $\boldsymbol{\rho}^{(0)}$; for example, one may take $\alpha^{(0)}$ and $\boldsymbol{\beta}^{(0)}$ as the respective maximum likelihood estimates under the independent case and $\boldsymbol{\rho}^{(0)} = \mathbf{0}$;
2. Update $\boldsymbol{\beta}$ from (4);
3. Update α from (5) and $\boldsymbol{\rho}$ from some correlation structure;
4. Repeat (1)–(3) above until the convergence.

The quantity $\mathrm{E}\left\{\mathrm{sech}^2\left(\frac{y_{ij}-\mu_{ij}}{2}\right)\right\}$, that appears in the expressions of $\mathrm{E}(\dot{u}_{ij})$, does not have closed-form solution, so it should be evaluated by integral approximation methods.

3.3 Inference

From Artes (1997) one has that $\hat{\boldsymbol{\beta}}$, obtained from the iterative process (4), is such as

$$\sqrt{n}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) \xrightarrow{D} \mathrm{N}_p(\mathbf{0}, \mathbf{J}_{\Psi}^{-1}(\boldsymbol{\beta})),$$

where $\mathbf{J}_{\Psi}(\boldsymbol{\beta}) = \lim_{n\to\infty} n^{-1}\mathbf{J}_{n\Psi}(\boldsymbol{\beta})$, with

$$\mathbf{J}_{n\Psi}(\boldsymbol{\beta}) = \left\{\sum_{i=1}^{n} \mathbf{S}_i(\boldsymbol{\beta})\right\}\left\{\sum_{i=1}^{n} \mathbf{V}_i(\boldsymbol{\beta})\right\}^{-1}\left\{\sum_{i=1}^{n} \mathbf{S}_i(\boldsymbol{\beta})\right\}$$

and $\sum_{i=1}^{n} \mathbf{S}_i(\boldsymbol{\beta})$ and $\sum_{i=1}^{n} \mathbf{V}_i(\boldsymbol{\beta})$ being, respectively, the sensitivity and variability matrices of $\boldsymbol{\Psi}(\boldsymbol{\beta})$ (see Sect. S.4). Therefore, a consistent estimator of the variance–covariance matrix of $\hat{\boldsymbol{\beta}}$ may be expressed as

$$\{\widehat{\mathbf{J}}_{\Psi}(\boldsymbol{\beta})\}^{-1} = \left\{\sum_{i=1}^{n} \mathbf{X}_i^{\top}\widehat{\mathbf{W}}_i\mathbf{X}_i\right\}^{-1}\left\{\sum_{i=1}^{n} \mathbf{X}_i^{\top}\widehat{\mathbf{N}}_i\widehat{\mathbf{C}}_i^{-1}\hat{\mathbf{u}}_i\hat{\mathbf{u}}_i^{\top}\widehat{\mathbf{C}}_i^{-1}\widehat{\mathbf{N}}_i\mathbf{X}_i\right\}\left\{\sum_{i=1}^{n} \mathbf{X}_i^{\top}\widehat{\mathbf{W}}_i\mathbf{X}_i\right\}^{-1},$$

where $\widehat{\mathbf{W}}_i = \widehat{\mathbf{N}}_i^{\top}\widehat{\mathbf{C}}_i^{-1}\widehat{\mathbf{N}}_i$ and $\widehat{\mathbf{C}}_i = \widehat{\boldsymbol{\Sigma}}_{u_i}^{\frac{1}{2}}\mathbf{R}_i(\hat{\rho})\widehat{\boldsymbol{\Sigma}}_{u_i}^{\frac{1}{2}}$ with $\widehat{\mathbf{N}}_i = \mathrm{diag}\{\widehat{\mathrm{E}}(\dot{u}_i), \ldots, \widehat{\mathrm{E}}(\dot{u}_{is_i})\}$, for $i = 1, \ldots, n$. If the regressors are dropped, the estimate of the asymptotic variance for the location parameter yields

$$\widehat{\mathrm{Var}}(\hat{\mu}) = \frac{\sum_{i=1}^{n} \mathbf{1}_{s_i}^{\top}\widehat{\mathbf{N}}_i\widehat{\mathbf{C}}_i^{-1}\hat{\mathbf{u}}_i\hat{\mathbf{u}}_i^{\top}\widehat{\mathbf{C}}_i^{-1}\widehat{\mathbf{N}}_i\mathbf{1}_{s_i}}{(\sum_{i=1}^{n} \mathbf{1}_{s_i}^{\top}\widehat{\mathbf{W}}_i\mathbf{1}_{s_i})^2},$$

where $\mathbf{1}_{s_i}$ denotes an $s_i \times 1$ vector of ones.

For assessing the hypothesis testing $\mathrm{H}_0 : \mathbf{A}\boldsymbol{\beta} = \mathbf{0}$ against $\mathrm{H}_1 : \mathbf{A}\boldsymbol{\beta} \neq \mathbf{0}$, where $\mathbf{A}$ is a $r \times p$ matrix of row rank r $(r \leq p)$, one may apply a Wald-type test whose respective statistic is given by $\xi_W = (\mathbf{A}\hat{\boldsymbol{\beta}})^{\top}\{\widehat{\mathbf{J}}_{\Psi}(\boldsymbol{\beta})\}\mathbf{A}\hat{\boldsymbol{\beta}}$. For large sample and under usual regularity conditions, it follows that $\xi_W \sim \chi_r^2$, where χ_r^2 denotes the chi-squared distribution with r degrees of freedom.

3.4 Simulation study

In this section, we describe a simulation study to assess the large sample behavior of the estimators from the iterative process developed in Sect. 3.2, for analyzing regression models under correlated BS data and based on estimating equations.

First, we generate $\mathbf{t}_i = (t_{i1}, \ldots, t_{is})^{\top}$ from a multivariate BS distribution, as described in Kundu et al. (2013), with location parameter vector $\boldsymbol{\eta}_i = (\eta_{i1}, \ldots, \eta_{is})^{\top}$, shape parameter α and correlation matrix $\boldsymbol{\Sigma}$, for $i = 1, \ldots, n$. Then, we calculate $y_{ij} = \log(t_{ij})$ and $\mu_{ij} = \log(\eta_{ij})$, where

$$\mu_{ij} = \beta_0 + \beta_1 x_{ij},$$

with x_{ij}'s being fixed values generated from a uniform distribution in the range [0, 1], for $i = 1, \ldots, n$ and $j = 1, \ldots, s$. The $\boldsymbol{\Sigma}$ matrix is structured as either first-order autoregressive correlation (AR(1)) or exchangeable correlation, with ρ denoting the correlation coefficient. The values assigned for the parameters are $\beta_0 = 4$, $\beta_1 = -2$, $\alpha = 0.5$ and $\rho = 0.3, 0.6$ and 0.9, whereas $n = 10, 30, 50$ and 80 and $s = 3, 5$ and 10. The bias (in absolute value) and the mean squared error (MSE) were calculated, and for each scenario, 5000 replicates were considered. In Tables S1–S8, we describe the results of the simulation study, in which the responses were generated from BS multivariate distributions and fitted under BS-GEE models. Tables S1–S4 are concerning, respectively, with simulated responses under AR(1) and exchangeable structures and fitted under the same correlation structures, whereas in Table S5–S8 one has the results when the responses are simulated under AR(1) (exchangeable) correlation structure and fitted under exchangeable (AR(1)) correlation structure.

We may notice from Tables S1–S4 that MSE and bias of $\hat{\beta}_0$, $\hat{\beta}_1$, $\hat{\alpha}$ and $\hat{\rho}$ decrease as s increases as well as the sample size n increases with strong indication of consistency for the four estimators and for all the scenarios considered. From Tables S5–S8 that describe the results considering misspecification of the correlation structure, one may also notice that MSE and bias of $\hat{\beta}_0$, $\hat{\beta}_1$, $\hat{\alpha}$ and $\hat{\rho}$ decrease as s increases and the sample size n increases. There is indication of consistency for the estimators $\hat{\beta}_0$, $\hat{\beta}_1$ and $\hat{\alpha}$ for all the scenarios, whereas the estimator of ρ does not seem to be consistent. This last result is expected since we are using the moment estimator for the correlation coefficient under the generated multivariate distribution.

Table S9 presents the results from the simulation study for $n = 100$ and $s = 5$ in which the efficiency of the robust ($\mathrm{EF_R}$) and naive ($\mathrm{EF_N}$) estimators of β_0 and β_1 are evaluated. The naive estimator of $\hat{\boldsymbol{\beta}}$ is given by $\mathrm{Var_N}(\hat{\boldsymbol{\beta}}) = \{\sum_{i=1}^{n} \mathbf{X}_i^\top \widehat{\mathbf{W}}_i \mathbf{X}_i\}^{-1}$. When data are generated and fitted under the same correlation structure, in general the efficiency becomes very close to one for both estimators. However, under misspecification the robust estimator loses efficiency which agrees with results described by Liang et al. (1992) and Fitzmaurice et al. (1993).

4 Diagnostic methods

The aim of diagnostic procedures is to assess the existence of departures from the model assumptions, as well as to detect discrepant observations, particularly influential and outlying ones. Such observations may exercise great influence on the parameter estimates and in some situations may cause inferential changes. There is a vast literature on diagnostic procedures in generalized estimating equations (see, for instance, Venezuela et al. 2007, 2011; Hardin and Hilbe 2012 and the references therein). However, in this work, we will discuss some extensions of the procedures developed by Venezuela et al. (2007) to the models (i)–(ii) proposed in Sect. 3.

4.1 Residual analysis

Since the full likelihood function is not known for the proposed model in Sect. 3, quantile and deviance component residuals are not possible to be evaluated, unless for the marginal models. So, we will derive a Pearson-type residual for the proposed model.

At the convergence, the iterative process (4) may be expressed as

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^\top \widehat{\mathbf{W}} \mathbf{X})^{-1} \mathbf{X}^\top \widehat{\mathbf{W}} \widehat{\mathbf{z}},$$

where $\mathbf{X} = (\mathbf{X}_1^\top, \ldots, \mathbf{X}_n^\top)^\top$ is the $\sum s_i \times p$ model matrix, $\widehat{\mathbf{W}} = \mathrm{diag}\{\widehat{\mathbf{W}}_1, \ldots, \widehat{\mathbf{W}}_n\}$ is the $\sum s_i \times \sum s_i$ weight matrix, whereas $\mathbf{z} = (\mathbf{z}_1^\top, \ldots, \mathbf{z}_n^\top)^\top$ with $\mathbf{z}_i = \mathbf{X}_i \boldsymbol{\beta} - \mathbf{N}_i^{-1} \mathbf{u}_i$ being a pseudo-response for the ith group, for $i = 1, \ldots, n$. Therefore, $\hat{\boldsymbol{\beta}}$ may be interpreted as the least squares solution of the linear regression of $\widehat{\mathbf{z}}$ onto the columns of the matrix $\mathbf{X}$ with weight matrix $\widehat{\mathbf{W}}$.

Thus, the orthogonal projection matrix of $\widehat{\mathbf{z}}$ onto $\mathbf{X}$ with weight matrix $\widehat{\mathbf{W}}$ becomes given by

$$\widehat{\mathbf{H}} = \mathrm{diag}\{\widehat{\mathbf{H}}_1, \ldots, \widehat{\mathbf{H}}_n\},$$

where $\widehat{\mathbf{H}}_i = \widehat{\mathbf{W}}_i^{\frac{1}{2}} \mathbf{X}_i (\mathbf{X}^\top \widehat{\mathbf{W}} \mathbf{X})^{-1} \mathbf{X}_i^\top \widehat{\mathbf{W}}_i^{\frac{1}{2}}$, with $\widehat{\mathbf{W}}_i^{\frac{1}{2}}$ being an $s_i \times s_i$ square root matrix derived from the eigenvalue decomposition of $\widehat{\mathbf{W}}_i$, for $i = 1, \ldots, n$. The principal diagonal elements of the matrix $\widehat{\mathbf{H}}_i$, named $\widehat{h}_{ijj}$ for $j = 1, \ldots, s_i$, may be considered as individual leverage measures. Also, we may construct group leverage measures as $\hat{h}_i = s_i^{-1} \sum_{j=1}^{s_i} \hat{h}_{ijj}$, for $i = 1, \ldots, n$. Index plots of $\hat{h}_{ijj}$ and $\hat{h}_i$ may be useful to reveal observations (or group of observations) with high influence on the fitted values, for $i = 1, \ldots, n$ and $j = 1, \ldots, s_i$.

Ordinary residuals for assessing the least squares solution may be defined as $r_{ij} = \mathbf{e}_{ij}^\top \widehat{\mathbf{W}}_i^{\frac{1}{2}} (\widehat{\mathbf{z}}_i - \mathbf{X}_i \hat{\boldsymbol{\beta}})$, where $\mathbf{e}_{ij}$ is an $s_i \times 1$ vector of zeros with 1 at the jth position. Similarly to Venezuela et al. (2007), we may define a standardized version, named standardized Pearson-type residual, as

$$t_{r_{ij}} = \frac{\mathbf{e}_{ij}^\top \widehat{\mathbf{W}}_i^{\frac{1}{2}} (\widehat{\mathbf{z}}_i - \mathbf{X}_i \hat{\boldsymbol{\beta}})}{\sqrt{1 - \widehat{h}_{ijj}}},$$

for $i = 1, \ldots, n$ and $j = 1, \ldots, s_i$. The index plot of $t_{r_{ij}}$ and the normal probability plot of $t_{r_{ij}}$ with a generated confidence band may reveal outlying observations as well as the adequacy of the proposed model. This confidence band, named envelope, was proposed by Atkinson (1981) for the normal case and has been extended for non-normal error models by various authors even for correlated data such as GEE (see, for instance, Venezuela et al. 2007). A recent review on this methodology may be found in Moral et al. (2017).

4.1.1 Simulation study

In this section, we describe a small simulation study with the same parameter settings assumed in Sect. 3.3 for assessing the empirical distribution of the standardized Pearson-type residual $t_{r_{ij}}$, for $i = 1, \ldots, n$ and $j = 1, \ldots, s$. The responses were generated from BS multivariate distributions and fitted under BS-GEE models. Tables S10 and S11 present the results when the data are generated and fitted under the same correlation structure, whereas in Tables S12 and S13 one has the misspecification of the correlation structure. Summary statistics for the standardized Pearson-type residual are evaluated as the average of $R = 5000$ replicates.

As we may see from Tables S10–S13, in general, the empirical distribution of $t_{r_{ij}}$ presents a mean very close to zero, a standard deviation very close to 1, a negligible skewness and kurtosis near 3. The Kolmogorov–Smirnov (KS) statistics is not significant in all the scenarios, confirming the very good agreement between the empirical distribution of the standardized Pearson-type residual and the standard normal distribution.

Even though one has for all the scenarios studied a very good approximation of the standardized Pearson-type residual to the normal distribution, we recommend the normal probability plot for the residual $t_{r_{ij}}$ added by the envelope. This graph may be useful to identify possible outlying observations as well as departures from the assumption of marginal BS distribution for the responses. Other graphs, such as of $t_{r_{ij}}$ against the fitted values, may be performed to assess departures from the assumption of variance homogeneity. Estimating equations for the heterogeneity case are derived in Sect. 6.

4.2 Cook distance

Cook distance (Cook 1977) is the most popular influence measure developed originally for linear models for assessing the effect of dropping individual observations on the regression estimates. Pregibon (1981) extended the procedure for generalized linear models by introducing the one-step approximation for obtaining one approximated Cook distance, whereas Preisser and Qaqish (1996) (see also Venezuela et al. 2007) applied this procedure in generalized estimating equations. However, this methodology may not be appropriate for time series models with autoregressive errors, as pointed out by Kim and Huggins (1998) and Paula et al. (2009). These authors argue the application of local influence approach.

Then, similarly to Preisser and Qaqish (1996), we obtain from the iterative process (4) the following one-step approximation:

$$\boldsymbol{\beta}_{(ij)}^{(1)} = \hat{\boldsymbol{\beta}} - \frac{r_{ij}}{(1 - \hat{h}_{ijj})} (\mathbf{X}^\top \widehat{\mathbf{W}} \mathbf{X})^{-1} \mathbf{X}_i^\top \widehat{\mathbf{W}}_i^{\frac{1}{2}} \mathbf{e}_{ij},$$

where $\boldsymbol{\beta}_{(ij)}^{(1)}$ denotes the one-step approximation for the estimate $\hat{\boldsymbol{\beta}}_{(ij)}$, after removing the jth measure of the ith experimental unit. The quality of the one-step approximation above depends on the accuracy of the estimators used in the expression. Then, the Cook distance is approximated by

$$\begin{aligned}\mathrm{C}_{ij} &= \frac{1}{p}(\boldsymbol{\beta}^{(1)}_{(ij)} - \hat{\boldsymbol{\beta}})^{\top}(\mathbf{X}^{\top}\widehat{\mathbf{W}}\mathbf{X})(\boldsymbol{\beta}^{(1)}_{(ij)} - \hat{\boldsymbol{\beta}}) \\ &= \frac{1}{p} t^2_{r_{ij}} \frac{\hat{h}_{ijj}}{(1-\hat{h}_{ijj})}.\end{aligned}$$

Index plot of C_{ij} may reveal observations with large influence on $\hat{\boldsymbol{\beta}}$. The one-step approximation when the ith experimental unit is dropped is derived in Sect. S.6.

4.3 Local influence

Cook (1986) introduced the local influence method as a general way for assessing the sensitivity of the parameter estimates, under small perturbations in the model or data. Cadigan and Farrell (2002) extended the local influence method by replacing the likelihood displacement by a more general influence measure, whereas Venezuela et al. (2011) applied the methodology in generalized estimating equations. Basically, the likelihood equations are replaced by the estimating equations

$$\boldsymbol{\Psi}(\hat{\boldsymbol{\beta}}) = \left\{ \frac{\partial \mathcal{F}(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}} \right\}\Bigg|_{\boldsymbol{\beta}=\hat{\boldsymbol{\beta}}} = \mathbf{0},$$

where $\mathcal{F}(\boldsymbol{\beta})$ is named fit function, that is assumed twice differentiable in $\boldsymbol{\beta}$ with unique interior parameter estimate. Thus, an appropriate influence measure that generalizes the likelihood displacement is defined as $\mathrm{FD}_{\omega} = 2\{\mathcal{F}(\hat{\boldsymbol{\beta}}) - \mathcal{F}(\hat{\boldsymbol{\beta}}_{\omega})\}$, where $\boldsymbol{\omega} = (\omega_1, \ldots, \omega_q)^{\top}$ denotes a perturbation vector and $\hat{\boldsymbol{\beta}}_{\omega}$ is the solution of the perturbed estimating equations $\boldsymbol{\Psi}(\hat{\boldsymbol{\beta}}_{\omega}|\boldsymbol{\omega}) = \mathbf{0}$. There is a no-perturbation vector $\boldsymbol{\omega}_0$ such that $\boldsymbol{\Psi}(\boldsymbol{\beta}_{\omega}|\boldsymbol{\omega}_0) = \boldsymbol{\Psi}(\boldsymbol{\beta})$. The normal conformal curvature (proposed by Poon and Poon 1999) in the unitary direction $\mathbf{d}$ is given by $\mathrm{B}_d(\boldsymbol{\beta}) = |\mathbf{d}^{\top}\boldsymbol{\Delta}^{\top}\{\ddot{\mathcal{F}}(\boldsymbol{\beta})\}^{-1}\boldsymbol{\Delta}\mathbf{d}|/\sqrt{tr(\boldsymbol{\Delta}^{\top}\{\ddot{\mathcal{F}}(\boldsymbol{\beta})\}^{-1}\boldsymbol{\Delta})^2}$, where $0 \leq \mathrm{B}_d(\boldsymbol{\beta}) \leq 1$ with $\boldsymbol{\Delta} = \partial\boldsymbol{\Psi}(\boldsymbol{\beta}|\boldsymbol{\omega})/\partial\boldsymbol{\omega}^{\top}$ evaluated at $\boldsymbol{\beta} = \hat{\boldsymbol{\beta}}$ and $\boldsymbol{\omega} = \boldsymbol{\omega}_0$.

Usual graphs to assess the sensitivity of the parameter estimates are the index plot of $\mathbf{d}_{max}$ (the eigenvector corresponding to the largest eigenvalue of the matrix $\mathbf{B} = \boldsymbol{\Delta}^{\top}\{\ddot{\mathcal{F}}(\boldsymbol{\beta})\}^{-1}\boldsymbol{\Delta}$) and the index plot of $\mathrm{B}_i = \mathrm{B}_{d_i}(\boldsymbol{\beta})$ with $\mathbf{d}_i$ denoting a vector of zeros with 1 at the ith position. For the calculation, we will replace $\ddot{\mathcal{F}}(\boldsymbol{\beta})$ by the sensitivity matrix $\mathrm{E}\{\ddot{\mathcal{F}}(\boldsymbol{\beta})\}$. In the sequel, we will derive the elements of the matrix Δ for some perturbation schemes.

4.3.1 Case-weight perturbation

The estimating equation for $\boldsymbol{\beta}$, given in (1), may be re-expressed as

$$\boldsymbol{\Psi}(\boldsymbol{\beta}) = \mathbf{X}^{\top}\mathbf{W}\mathbf{N}^{-1}\mathbf{u},$$

where $\mathbf{X} = (\mathbf{X}_1^\top, \ldots, \mathbf{X}_n^\top)^\top$, $\mathbf{W} = \text{blockdiag}\{\mathbf{W}_1, \ldots, \mathbf{W}_n\}$, $\mathbf{N} = \text{blockdiag}\{\mathbf{N}_1, \ldots, \mathbf{N}_n\}$ and $\mathbf{u} = (\mathbf{u}_1^\top, \ldots, \mathbf{u}_n^\top)^\top$. Then, the case-weight perturbation scheme leads to the estimating equation

$$\boldsymbol{\Psi}(\boldsymbol{\beta}|\boldsymbol{\omega}) = \mathbf{X}^\top \mathbf{W}\mathbf{N}^{-1}\text{diag}(\boldsymbol{\omega})\mathbf{u},$$

where $\boldsymbol{\omega} = (\boldsymbol{\omega}_1^\top, \ldots, \boldsymbol{\omega}_n^\top)^\top$ with $\boldsymbol{\omega}_i = (\omega_{i1}, \ldots, \omega_{is_i})^\top$, $i = 1, \ldots, n$, and the no-perturbation vector $\boldsymbol{\omega}_0$ is formed by 1's. We obtain

$$\boldsymbol{\Delta} = \frac{\partial \boldsymbol{\Psi}(\boldsymbol{\beta}|\boldsymbol{\omega})}{\partial \boldsymbol{\omega}^\top} = \mathbf{X}^\top \mathbf{W}\mathbf{N}^{-1}\text{diag}(\mathbf{u})$$

evaluated at $\boldsymbol{\beta} = \hat{\boldsymbol{\beta}}$ and $\boldsymbol{\omega} = \boldsymbol{\omega}_0$.

4.3.2 Response perturbation

We will consider the response perturbation

$$y_{\omega_{ij}} = y_{ij} + \omega_{ij} s_{y_{ij}},$$

where $\omega_{ij} \in \mathbb{R}$, $\boldsymbol{\omega}_0$ is a vector of zeros and $s_{y_{ij}}$ denotes the estimated standard deviation of y_{ij}. The perturbed estimating equation takes the form

$$\boldsymbol{\Psi}(\boldsymbol{\beta}|\boldsymbol{\omega}) = \mathbf{X}^\top \mathbf{W}\mathbf{N}^{-1}\mathbf{u}_\omega,$$

where $\boldsymbol{\omega} = (\boldsymbol{\omega}_1^\top, \ldots, \boldsymbol{\omega}_n^\top)^\top$ with $\boldsymbol{\omega}_i = (\omega_{i1}, \ldots, \omega_{is_i})^\top$, $i = 1, \ldots, n$, and the no-perturbation vector $\boldsymbol{\omega}_0$ is formed by 0's, $\mathbf{u}_\omega = (\mathbf{u}_{\omega_1}^\top, \ldots, \mathbf{u}_{\omega_n}^\top)^\top$ with $\mathbf{u}_{\omega_i} = (u_{\omega_{i1}}, \ldots, u_{\omega_{is_i}})^\top$ and

$$u_{\omega_{ij}} = \frac{1}{\alpha^2}\sinh(y_{\omega_{ij}} - \mu_{ij}) - \frac{1}{2}\tanh\left(\frac{y_{\omega_{ij}} - \mu_{ij}}{2}\right),$$

for $i = 1, \ldots, n$ and $j = 1, \ldots, s_i$. Therefore

$$\boldsymbol{\Delta} = \frac{\partial \boldsymbol{\Psi}(\boldsymbol{\beta}|\boldsymbol{\omega})}{\partial \boldsymbol{\omega}^\top} = \mathbf{X}^\top \mathbf{W}\mathbf{N}^{-1}\text{diag}\left(\frac{\partial \mathbf{u}_\omega}{\partial \boldsymbol{\omega}^\top}\right)$$

evaluated at $\boldsymbol{\beta} = \hat{\boldsymbol{\beta}}$ and $\boldsymbol{\omega} = \boldsymbol{\omega}_0$, where the diagonal elements of the matrix $\{\partial \mathbf{u}_\omega / \partial \boldsymbol{\omega}^\top\}$ are given by

$$\frac{\partial u_{\omega_{ij}}}{\partial \omega_{ij}} = \left[\frac{1}{\alpha^2}\cosh(y_{\omega_{ij}} - \mu_{ij}) - \frac{1}{4}\text{sech}^2\left\{\frac{y_{\omega_{ij}} - \mu_{ij}}{2}\right\}\right] s_{y_{ij}},$$

for $i = 1, \ldots, n$ and $j = 1, \ldots, s_i$.

4.3.3 Single-covariate perturbation

We will consider additive perturbation scheme for the kth column of $\mathbf{X}$ (assumed continuous), in which each component is given by

$$x_{k\omega_{ij}} = x_{ijk} + \omega_{ij} s_{x_k},$$

where $\omega_{ij} \in \mathbb{R}$, $\boldsymbol{\omega}_0$ is a vector of zeros and s_{x_k} denotes the estimated standard deviation of x_k. One has

$$\mu_{\omega_{ij}} = \beta_1 x_{ij1} + \beta_2 x_{ij2} + \cdots + \beta_k (x_{ijk} + \omega_{ij} s_{x_k}) + \cdots + \beta_p x_{pij}$$

and the estimating equation (1) may be re-expressed as

$$\boldsymbol{\Psi}(\boldsymbol{\beta}) = \mathbf{X}^\top \mathbf{N} \mathbf{C}^{-1} \mathbf{u},$$

where $\mathbf{C} = \boldsymbol{\Sigma}_u^{\frac{1}{2}} \mathbf{R}(\boldsymbol{\rho}) \boldsymbol{\Sigma}_u^{\frac{1}{2}}$ with $\mathbf{R}(\boldsymbol{\rho}) = \text{blockdiag}\{\mathbf{R}_1(\boldsymbol{\rho}), \ldots, \mathbf{R}_n(\boldsymbol{\rho})\}$ and $\boldsymbol{\Sigma}_u^{\frac{1}{2}} = -\mathbf{N}^{\frac{1}{2}}$. Then, the perturbed estimating equation takes the form

$$\boldsymbol{\Psi}(\boldsymbol{\beta}|\boldsymbol{\omega}) = \mathbf{X}_\omega^\top \mathbf{N}_\omega \mathbf{C}_\omega^{-1} \mathbf{u}_\omega,$$

with $\mathbf{C}_\omega = \mathbf{N}_\omega^{\frac{1}{2}} \mathbf{R}(\boldsymbol{\rho}) \mathbf{N}_\omega^{\frac{1}{2}}$, and consequently we obtain

$$\boldsymbol{\Delta} = \frac{\partial \boldsymbol{\Psi}(\boldsymbol{\beta}|\boldsymbol{\omega})}{\partial \boldsymbol{\omega}^\top} = \mathbf{X}_\omega^\top \mathbf{N}_\omega \mathbf{C}_\omega^{-1} \frac{\partial \mathbf{u}_\omega}{\partial \boldsymbol{\omega}^\top} + \mathbf{A} \mathbf{N}_\omega \mathbf{C}_\omega^{-1} \mathbf{u}_\omega,$$

evaluated at $\boldsymbol{\beta} = \hat{\boldsymbol{\beta}}$ and $\boldsymbol{\omega} = \boldsymbol{\omega}_0$, where the diagonal elements of the matrix $\{\partial \mathbf{u}_\omega / \partial \boldsymbol{\omega}^\top\}$ are given by

$$\frac{\partial u_{\omega_{ij}}}{\partial \omega_{ij}} = \left[\frac{1}{4}\text{sech}^2\left(\frac{y_{ij} - \mu_{\omega_{ij}}}{2}\right) - \frac{1}{\alpha^2}\left\{1 + 2\text{sinh}^2\left(\frac{y_{ij} - \mu_{\omega_{ij}}}{2}\right)\right\}\right] \beta_k s_{x_k},$$

for $i = 1, \ldots, n$ and $j = 1, \ldots, s_i$, and $\mathbf{A}$ is a $p \times \sum s_i$ matrix of zeros with 1's in the tth row.

5 Application

As illustration of the methodology developed in this paper, we will consider the data set described by Munnell (1990) on the productivity of public capital in the 48 continental states of the USA from 1970 to 1986. In particular, we will investigate the relationship between productivity of public capital and water utility capital across the years. The response refers to the gross state product (GSP) and the explanatory variable the water utility capital (WUC). Due to the time series structure of the data, we will assume a first-order autoregressive correlation among the responses of each state. This data set

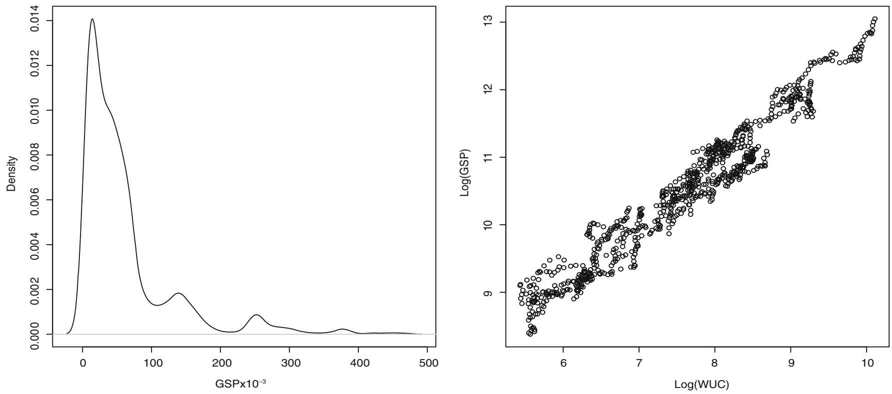

Fig. 1 Density of GSP (left) and scatter plot between the natural logarithm of WUC and the natural logarithm of GSP (right)

was analyzed by Greene (2012) and Manghi et al. (2016), respectively, under linear normal and linear elliptical mixed models. We will reanalyze this data set by applying generalized estimating equations with response BS.

Let $\mathbf{t}_i = (t_{i1}, \ldots, t_{i17})^\top$ and $\mathbf{x}_i = (x_{i1}, \ldots, x_{i17})^\top$ be the values of GSP and WUC, respectively, for the ith USA continental state across the 17 years, for $i = 1, \ldots, 48$. In Fig. 1 (left), one has the density of GSP (ignoring the explanatory variable values of WUC) with indication of skew distribution to the right, whereas in Fig. 1 (right) one has a linear tendency between the natural logarithm of WUC and the natural logarithm of GSP.

Thus, based on the descriptive analysis we propose the following BS-GEE model:

(i) $t_{ij}|x_{ij} \sim \mathrm{BS}(\alpha, \eta_{ij})$,
(ii) $\log(\eta_{ij}) = \beta_0 + \beta_1 \log(x_{ij})$,
(iii) $\mathbf{R}_i = \mathbf{R}_i(\rho)$ with $\mathrm{R}_{ijj'} = 1$, for $j = j'$, and $\mathrm{R}_{ijj'} = \rho^{|j-j'|}$, for $j \neq j'$,

where $\eta_{ij} > 0$ and $\alpha > 0$ denote, respectively, the gross state product median and the shape parameter, for $i = 1, \ldots, 48$ and $j = 1, \ldots, 17$. This regression is equivalent to a log-BS model with AR(1) correlation structure. Table 1 presents the parameter estimates from the generalized estimating equations with both coefficients being highly significant.

From Fig. 2 (left), one may notice that the standardized Pearson-type residuals are in general in the range $[-2, 2]$, whereas two measures for two different experimental units appear in Fig. 2 (right) as possible influential on the regression coefficients by the Cook's distance. Performing the local influence analysis, index plots of B_i under the case-weight, response and single-covariate perturbation schemes are described in Fig. 3. Secttion S.5 describes sensitivity studies for the shape parameter and the correlation coefficient estimates. We may notice from all graphs the following observations (state, year): (Delaware, 1972), (Wyoming, 1986), (SDakota, 1985), (SDakota, 1986), (Maine, 1977) and (Maine, 1978) as potentially influential. Elimination of each state pointed out in these graphs leads to relative changes up to 12% in the parameter estimates (see Table S14), but it does not change the inference. The normal probability plot

Table 1 Parameter estimates and approximate standard errors from the BS-GEE model fitted to the gross data set

Effect	Estimate	SE	z-Robust
Intercept	6.096	0.255	23.91
WUC	0.580	0.036	16.11
Shape	0.415	0.007	59.29
Correlation	0.979		

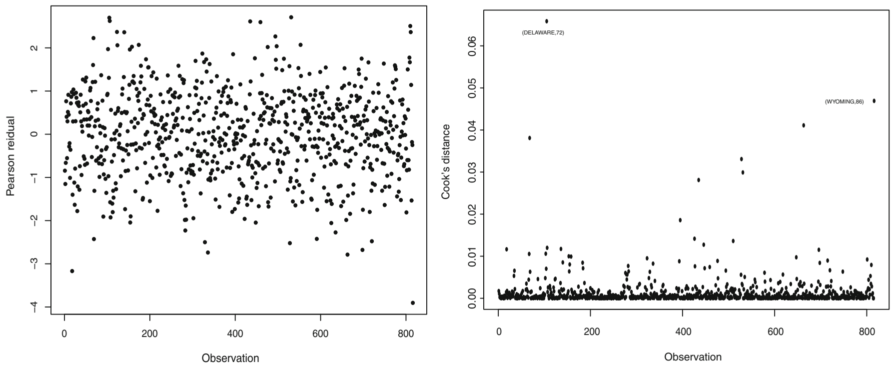

Fig. 2 Graphs of the standardized Pearson-type residual (left) and Cook distance (right) from the BS-GEE model fitted to the gross data set

with the generated confidence band of 95% for the standardized Pearson-type residual, described in Fig. 4 (left), does not present unusual features so that the assumption of BS distribution for the marginal response does not seem to be unsuitable.

For the purpose of comparison, we also fitted the following gamma-GEE model to the gross data set:

(i) $t_{ij}|x_{ij} \sim \mathrm{G}(\mu_{ij}, \phi)$,
(ii) $\log(\mu_{ij}) = \beta_0 + \beta_1 \log(x_{ij})$,
(iii) $\mathrm{Corr}(\mathbf{y}_i) = \mathbf{R}_i(\rho)$ with $\mathrm{R}_{ijj'} = 1$, for $j = j'$, and $\mathrm{R}_{ijj'} = \rho^{|j-j'|}$, for $j \neq j'$,

where $\mu_{ij} > 0$ and $\phi^{-1} > 0$ denote, respectively, the gross state product mean and the dispersion parameter, for $i = 1, \ldots, 48$ and $j = 1, \ldots, 17$. The normal probability plot of the respective standardized Pearson-type residual, see Fig. 4 (right), indicates some departures which may be indication that the gamma distribution is not suitable to explain the marginal response.

Therefore, it is reasonable to assume that marginally $t_{ij} \sim \mathrm{BS}(\alpha, \eta_{ij})$, where $\eta_{ij} = {x_{ij}}^{\beta_1} e^{\beta_0}$ denotes the median of GSP in the ith USA continental state in the jth year. To interpret the coefficient β_1, suppose that WUC increases $r \times 100\%$. Thus, the new value would be $\mathrm{WUC_N} = (1 + r)\mathrm{WUC}$ and the new median value for GSP becomes given by

$$\eta(\mathrm{WUC_N}) = e^{\beta_0} \beta_1^{\mathrm{WUC_N}}.$$

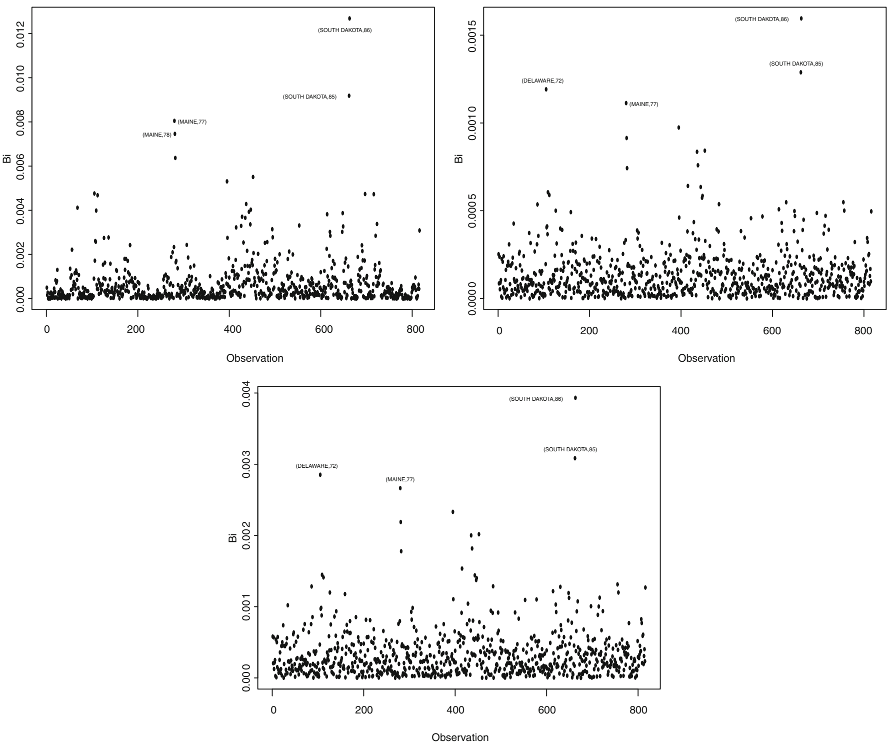

Fig. 3 Index plots of B_i under the case-weight (left), response (right) and single-covariate (bottom) perturbation schemes from the BS-GEE model fitted to the gross data set

The ratio between $\eta(\mathrm{WUC_N})$ and $\eta(\mathrm{WUC})$ yields

$$\begin{aligned}\frac{\eta(\mathrm{WUC_N})}{\eta(\mathrm{WUC})} &= (1+r)^{\beta_1}\\ &\cong (1+r\beta_1)\end{aligned}$$

for small r. Therefore, based on Table 1 we may conclude that for each change of 1% in WUC the median of GSP should increase about 0.58%. Since the shape parameter is same for all states and years, this interpretation is extended for all quantiles.

6 Median and shape modeling

Let now $\mathbf{t}_i = (t_{i1}, \ldots, t_{is_i})^\top$ be an $s_i \times 1$ vector containing positive responses for the ith experimental unit, for $i = 1, \ldots, n$. We will assume that $t_{ij} \sim \mathrm{BS}(\alpha_{ij}, \eta_{ij})$ with regression structure $\log(\eta_{ij}) = \mathbf{x}_{ij}^\top \boldsymbol{\beta}$ and $\log(\alpha_{ij}) = \mathbf{z}_{ij}^\top \boldsymbol{\gamma}$, where $\mathbf{x}_{ij} = (x_{ij1}, \ldots, x_{ijp})^\top$ and $\mathbf{z}_{ij} = (x_{ij1}, \ldots, x_{ijq})^\top$ contain values of explanatory variables, $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_p)^\top$ and $\boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_q)^\top$ are the regression coefficient vectors.

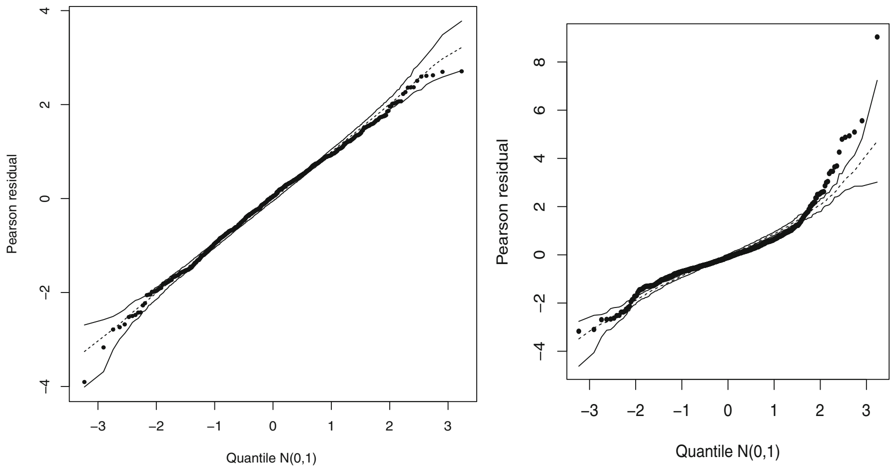

Fig. 4 Normal probability plot of the standardized Pearson-type residual with a 95% simulated confidence band from the BS-GEE model (left) and gamma-GEE model (right) fitted to the gross data set

Equivalently one has the following double model: (i) $y_{ij}|(\mathbf{x}_{ij}, \mathbf{z}_{ij}) \sim \text{log-BS}(\alpha_{ij}, \mu_{ij})$ with (ii) $\mu_{ij} = \mathbf{x}_{ij}^\top \boldsymbol{\beta}$ and $\log(\alpha_{ij}) = \mathbf{z}_{ij}^\top \boldsymbol{\gamma}$, for $i = 1, \ldots, n$ and $j = 1, \ldots, s_i$.

From Sects. S.1 and S.7, the optimum estimating function for $\boldsymbol{\theta} = (\boldsymbol{\beta}^\top, \boldsymbol{\gamma}^\top)^\top$ takes the form

$$\boldsymbol{\Gamma}^*(\boldsymbol{\theta}) = \sum_{i=1}^{n} \mathrm{E}\begin{pmatrix} \frac{\partial \mathbf{u}_i}{\partial \boldsymbol{\beta}^\top} & \frac{\partial \mathbf{u}_i}{\partial \boldsymbol{\gamma}^\top} \\ \frac{\partial \mathbf{v}_i}{\partial \boldsymbol{\beta}^\top} & \frac{\partial \mathbf{v}_i}{\partial \boldsymbol{\gamma}^\top} \end{pmatrix}^\top \mathrm{Cov}(\mathbf{d}_i)^{-1}\mathbf{d}_i,$$

where $\mathbf{d}_i = (\mathbf{u}_i^\top, \mathbf{v}_i^\top)^\top$, $\mathbf{u}_i$ is defined as in Sect. 3.1 and $\mathbf{v}_i = (v_{i1}, \ldots, v_{is_i})^\top$ with $v_{ij} = v_i(y_{ij}; \boldsymbol{\theta})$ so that $\mathrm{E}(v_{ij}) = 0$ and $\mathrm{Cov}(\mathbf{d}_i)$ denotes the variance–covariance matrix of $\mathbf{d}_i$. In addition, one has that $\mathrm{E}(\partial \mathbf{u}_i/\partial \boldsymbol{\beta}^\top) = \mathbf{N}_i\mathbf{X}_i$, $\mathrm{E}(\partial \mathbf{v}_i/\partial \boldsymbol{\gamma}^\top) = \boldsymbol{\Omega}_i\mathbf{M}_i\mathbf{Z}_i$, $\mathrm{E}(\partial \mathbf{u}_i/\partial \boldsymbol{\gamma}^\top) = \mathbf{0}$ and $\mathrm{E}(\partial \mathbf{v}_i/\partial \boldsymbol{\beta}^\top) = \mathbf{0}$, where $\mathbf{N}_i$ and $\mathbf{X}_i$ were defined in Sect. 3.1, $\mathbf{Z}_i$ is an $s_i \times q$ matrix with rows $\mathbf{z}_{ij}^\top$, $\mathbf{M}_i = \mathrm{E}(\partial \mathbf{v}_i/\partial \boldsymbol{\alpha}_i^\top) = \mathrm{diag}\{\mathrm{E}(\dot{v}_{i1}), \ldots, \mathrm{E}(\dot{v}_{is_i})\}$ and $\boldsymbol{\Omega}_i = \mathrm{diag}\{\alpha_{i1}, \ldots, \alpha_{is_i}\}$, for $i = 1, \ldots, n$ and $j = 1, \ldots, s_i$.

A candidate for v_{ij} is the score function of α_{ij} given by

$$\begin{aligned} v_{ij} &= \partial \log\{f(y_{ij}; \alpha_{ij}, \mu_{ij})\}/\partial \alpha_{ij} \\ &= \frac{1}{\alpha_{ij}}\left\{\frac{4}{\alpha_{ij}^2}\sinh^2\left(\frac{y_{ij}-\mu_{ij}}{2}\right) - 1\right\}, \end{aligned}$$

and from Rieck and Nedelman (1991) $(2/\alpha_{ij})\sinh\{(y_{ij} - \mu_{ij})/2\} \sim \mathrm{N}(0, 1)$ then $\mathrm{E}[\sinh^2\{(y_{ij} - \mu_{ij})/2\}] = \alpha_{ij}^2/4$, which implies $\mathrm{E}(v_{ij}) = 0$ and $\mathrm{Var}(v_{ij}) = \mathrm{E}(v_{ij}^2) = -\mathrm{E}(\dot{v}_{ij}) = 2\alpha_{ij}^{-2}$, for $i = 1, \ldots, n$ and $j = 1, \ldots, s_i$. Thus, the optimum estimating function for $\boldsymbol{\theta}$ reduces to

$$\boldsymbol{\Gamma}^*(\boldsymbol{\theta}) = \sum_{i=1}^{n} \begin{pmatrix} \mathbf{X}_i^\top \mathbf{N}_i & \mathbf{0} \\ \mathbf{0} & \mathbf{Z}_i^\top \mathbf{M}_i \boldsymbol{\Omega}_i \end{pmatrix} \mathrm{Cov}(\mathbf{d}_i)^{-1} \mathbf{d}_i .$$

Under the conditions that $\mathrm{Cov}(\mathbf{u}_i, \mathbf{v}_i) = \mathbf{0}$ and $\mathrm{Corr}(\mathbf{v}_i) = \mathbf{I}_{s_i}$ (see, for instance, Artes and Jorgensen 2000) one has $\mathrm{Cov}(\mathbf{d}_i) = \mathrm{blockdiag}\{\mathbf{C}_i, \boldsymbol{\Sigma}_{v_i}\}$, where $\boldsymbol{\Sigma}_{v_i} = \mathrm{diag}\{\mathrm{Var}(v_{i1}), \ldots, \mathrm{Var}(v_{is_i})\} = -\mathbf{M}_i$. Then, by assuming that the Pearson correlation matrix $\mathbf{R}(\mathbf{u}_i)$ has a parametric form $\mathbf{R}_i(\boldsymbol{\rho})$, we obtain the estimating function

$$\begin{aligned} \boldsymbol{\Gamma}(\boldsymbol{\theta}) &= (\boldsymbol{\Gamma}_\beta^\top(\boldsymbol{\theta}), \boldsymbol{\Gamma}_\alpha^\top(\boldsymbol{\theta}))^\top \\ &= \sum_{i=1}^{n} \begin{pmatrix} \mathbf{X}_i^\top \mathbf{N}_i \mathbf{C}_i^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{Z}_i^\top \boldsymbol{\Omega}_i \end{pmatrix} \mathbf{d}_i \\ &= \sum_{i=1}^{n} \mathbf{Q}_i^\top \mathbf{W}_i \mathbf{A}_i^{-1} \mathbf{d}_i , \end{aligned} \tag{6}$$

where $\mathbf{Q}_i = \mathrm{blockdiag}\{\mathbf{X}_i, \mathbf{Z}_i\}$, $\mathbf{W}_i = \mathrm{blockdiag}\{\mathbf{N}_i \mathbf{C}_i^{-1} \mathbf{N}_i, \boldsymbol{\Omega}_i \boldsymbol{\Sigma}_{v_i} \boldsymbol{\Omega}_i\}$ and $\mathbf{A}_i = \mathrm{blockdiag}\{\mathbf{N}_i, -2\boldsymbol{\Omega}_i^{-1}\}$, for $i = 1, \ldots, n$.

To solve $\boldsymbol{\Gamma}(\hat{\boldsymbol{\theta}}) = \mathbf{0}$, we may apply the Newton scoring method with $\boldsymbol{\Gamma}'(\boldsymbol{\theta})$ being replaced by its expectation, $\mathrm{E}\{\boldsymbol{\Gamma}'(\boldsymbol{\theta})\} = \sum_{i=1}^{n} \mathbf{Q}_i^\top \mathbf{W}_i \mathbf{Q}_i$, obtaining the following iterative process:

$$\boldsymbol{\theta}^{(m+1)} = \left\{ \sum_{i=1}^{n} \mathbf{Q}_i^\top \mathbf{W}_i^{(m)} \mathbf{Q}_i \right\}^{-1} \left\{ \sum_{i=1}^{n} \mathbf{Q}_i^\top \mathbf{W}_i^{(m)} \boldsymbol{\xi}_i^{(m)} \right\}, \tag{7}$$

for $m = 0, 1, 2 \ldots$, where $\boldsymbol{\xi}_i = \mathbf{Q}_i \boldsymbol{\theta} - \mathbf{A}_i^{-1} \mathbf{d}_i$ is a modified dependent variable, for $i = 1, \ldots, n$. The iterative process (7) may be started with the respective maximum likelihood estimates of $\boldsymbol{\beta}$ and $\boldsymbol{\alpha}$ from the independent case and $\boldsymbol{\rho} = \mathbf{0}$. Then, we should update (7) alternating with the moment estimate of $\boldsymbol{\rho}$ until the convergence.

Using results from Sect. S.1, one has that

$$\sqrt{n}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \xrightarrow{D} \mathrm{N}_{(p+q)}(\mathbf{0}, \mathbf{J}_\Gamma^{-1}(\boldsymbol{\theta})),$$

where $\mathbf{J}_\Gamma(\boldsymbol{\theta}) = \lim_{n\to\infty} n^{-1} \mathbf{J}_{n\Gamma}(\boldsymbol{\theta})$, with

$$\mathbf{J}_{n\Gamma}(\boldsymbol{\theta}) = \left\{ \sum_{i=1}^{n} \mathbf{S}_i(\boldsymbol{\theta}) \right\} \left\{ \sum_{i=1}^{n} \mathbf{V}_i(\boldsymbol{\theta}) \right\}^{-1} \left\{ \sum_{i=1}^{n} \mathbf{S}_i(\boldsymbol{\theta}) \right\},$$

$\mathbf{S}_i(\boldsymbol{\theta}) = \mathbf{Q}_i^\top \mathbf{W}_i \mathbf{Q}_i$ and $\mathbf{V}_i(\boldsymbol{\theta}) = \mathbf{Q}_i^\top \mathbf{W}_i \mathbf{A}_i^{-1} \mathrm{Cov}(\mathbf{d}_i) \mathbf{A}_i^{-1} \mathbf{W}_i \mathbf{Q}_i$ (see Sect. S.7), for $i = 1, \ldots, n$. Therefore, we may derive a sandwich consistent estimator for the variance–covariance matrix of $\hat{\boldsymbol{\theta}}$ given by

$$\{\widehat{\mathbf{J}}_{\Gamma}(\boldsymbol{\theta})\}^{-1} = \left\{\sum_{i=1}^{n} \mathbf{Q}_i^{\top} \widehat{\mathbf{W}}_i \mathbf{Q}_i\right\}^{-1} \left\{\sum_{i=1}^{n} \mathbf{Q}_i^{\top} \widehat{\mathbf{W}}_i \widehat{\mathbf{A}}_i^{-1} \hat{\mathbf{d}}_i \hat{\mathbf{d}}_i^{\top} \widehat{\mathbf{A}}_i^{-1} \widehat{\mathbf{W}}_i \mathbf{Q}_i\right\} \left\{\sum_{i=1}^{n} \mathbf{Q}_i^{\top} \widehat{\mathbf{W}}_i \mathbf{Q}_i\right\}^{-1},$$

where $\widehat{\mathbf{W}}_i = \text{blockdiag}\{\widehat{\mathbf{N}}_i \widehat{\mathbf{C}}_i^{-1} \widehat{\mathbf{N}}_i, \widehat{\boldsymbol{\Omega}}_i \widehat{\boldsymbol{\Sigma}}_{v_i} \widehat{\boldsymbol{\Omega}}_i\}$ and $\widehat{\mathbf{A}}_i = \text{blockdiag}\{\widehat{\mathbf{N}}_i, -2\widehat{\boldsymbol{\Omega}}_i^{-1}\}$, for $i = 1, \ldots, n$. In Sect. S.8, some diagnostic measures are derived for the double case.

6.1 Homogeneity of shape parameter

Consider the particular case: (i) $y_{ij}|\mathbf{x}_{ij} \sim \text{log-BS}(\alpha_{ij}, \mu_{ij})$ with (ii) $\mu_{ij} = \mathbf{x}_{ij}^{\top}\boldsymbol{\beta}$ and $\alpha_{ij} = \alpha$, for $i = 1, \ldots, n$ and $j = 1, \ldots, s_i$. Here, one has $\mathbf{Z}_i = \mathbf{1}_{s_i}$, $\boldsymbol{\Sigma}_{v_i} = \frac{2}{\alpha^2}\mathbf{I}_{s_i}$ and $\boldsymbol{\Omega}_i = \alpha \mathbf{I}_{s_i}$; consequently $\mathbf{Q}_i = \text{blockdiag}\{\mathbf{X}_i, \mathbf{1}_{s_i}\}$, $\boldsymbol{\Sigma}_{v_i} = \frac{2}{\alpha^2}\mathbf{I}_{s_i}$, $\boldsymbol{\Omega}_i = \alpha\mathbf{I}_{s_i}$, $\mathbf{W}_i = \text{blockdiag}\{\mathbf{N}_i \mathbf{C}_i^{-1} \mathbf{N}_i, 2\mathbf{I}_{s_i}\}$ and $\mathbf{A}_i = \text{blockdiag}\{\mathbf{N}_i, -\frac{2}{\alpha}\mathbf{I}_{s_i}\}$, for $i = 1, \ldots, n$.

From (6), it follows that the estimating equation for obtaining $\hat{\alpha}$, for μ_{ij} fixed, becomes given by

$$\boldsymbol{\Gamma}_{\alpha}(\boldsymbol{\theta}) = \sum_{i=1}^{n} \sum_{j=1}^{s_i} \left\{4 \sinh^2\left(\frac{y_{ij} - \mu_{ij}}{2}\right) - \hat{\alpha}\right\},$$

which leads to the same estimator proposed in (5). On the other hand, the solution of the estimating equations $\boldsymbol{\Gamma}_{\beta}(\hat{\boldsymbol{\theta}}) = \mathbf{0}$ leads to the same iterative process presented in (4). Thus, the iterative process (7), under $\alpha_{ij} = \alpha$, alternating with the moment estimate of $\boldsymbol{\rho}$ is equivalent to the iterative procedure presented in Sect. 3.2.

Developing $\{\widehat{\mathbf{J}}_{\Gamma}(\boldsymbol{\theta})\}^{-1}$, we find the asymptotic estimator $\widehat{\text{Var}}(\hat{\boldsymbol{\beta}}) = \{\widehat{\mathbf{J}}_{\Psi}(\boldsymbol{\beta})\}^{-1}$, that is, the same expression derived in Sect. 3.4. In addition, we obtain the following asymptotic estimators:

$$\widehat{\text{Var}}(\hat{\alpha}) = \frac{\hat{\alpha}^2 \sum_{i=1}^{n} \mathbf{1}_{s_i}^{\top} \hat{\mathbf{v}}_i \hat{\mathbf{v}}_i^{\top} \mathbf{1}_{s_i}}{4\left(\sum_{i=1}^{n} s_i\right)^2}$$

and

$$\widehat{\text{Cov}}(\hat{\boldsymbol{\beta}}, \hat{\alpha}) = -\frac{\hat{\alpha}}{\sum_{i=1}^{n} s_i} \left\{\sum_{i=1}^{n} \mathbf{X}_i^{\top} \widehat{\mathbf{W}}_i \mathbf{X}_i\right\}^{-1} \left\{\sum_{i=1}^{n} \mathbf{X}_i^{\top} \widehat{\mathbf{N}}_i \widehat{\mathbf{C}}_i^{-1} \hat{\mathbf{u}}_i \hat{\mathbf{v}}_i^{\top} \mathbf{1}_{s_i}\right\}.$$

Thus, the joint iterative process for obtaining $(\hat{\boldsymbol{\beta}}, \hat{\alpha}, \hat{\rho})$ given in Sect. 6 leads to the same parameter estimates of the iterative process given in Sect. 3.2. However, in the numerical studies we have performed the last iterative process has been much faster than the one described in Sect. 6.

7 Conclusions

Generalized estimating equations for modeling the median of the life time until the failure of BS correlated data by a regression structure are derived in this paper. A reweighted iterative process for the parameter estimation, consistent estimators and various diagnostic procedures are also proposed as well as simulation studies to perform the empirical distribution of the derived estimators. An application in which a BS-GEE model is fitted to a real data set is given, and a comparison is presented with a gamma-GEE model. Extensions for the double case, in which the median is jointly modeled with the shape parameter, are also summarized, with the particular case of shape parameter homogeneity being discussed. R codes (R Core Team 2017) developed by authors to fit the proposed models and to derive the diagnostic graphs are available upon request.

Acknowledgements This work was supported by FAPESP and CNPq, Brazil. The authors are grateful to the Associate Editor and reviewers for their helpful comments.

Funding Funding was provided by Conselho Nacional de Desenvolvimento Científico e Tecnológico (Grant No. 305757/2014-8 and No. 310359/2017-1).

References

Alencar AP, Singer JM, Rocha FMM (2012) Competing regression models for longitudinal data. Biom J 54:214–229
Artes R (1997) Extensions of the Generalized Estimating Equation Theory to Circular Data and Dispersion Models [unpublished Ph.D. thesis in Portuguese]. Department of Statistics, University of São Paulo, Brazil
Artes R, Jorgensen B (2000) Longitudinal data estimating equations for dispersion model. Scand J Stat 27:321–334
Atkinson AC (1981) Two graphical displays for outlying and influential observations in regression. Biometrika 68:13–20
Balakrishnan N, Kundu D (2019) Birnbaum-Saunders distribution: a review of models, analysis, and applications. Appl Stoch Model Bus 35:4–49
Birnbaum ZW, Saunders SC (1969) A new family of life distributions. J Appl Probab 6:319–327
Cadigan NG, Farrell PJ (2002) Generalized local influence with applications to fish stock cohort analysis. Appl Stat 51:469–483
Cook RD (1977) Detection of influential observations in linear regressions. Technometrics 19:15–18
Cook RD (1986) Assessment of local influence. J R Stat Soc B 48:133–169
Crowder M (1987) On linear and quadratic estimating functions. Biometrika 74:591–597
Fitzmaurice GM, Laird NM, Rotnitzky AG (1993) Regression models for discrete longitudinal responses. Stat Sci 8:284–299
Godambe VP (1997) Estimating functions: a synthesis of least squares and maximum likelihood methods. In: Selected proceedings of the symposium on estimating functions. Institute of Mathematical Statistics, Hayward, CA, pp 5–16
Greene WH (2012) Econometric analysis. Prentice Hall, New York
Hardin JW, Hilbe JM (2012) Generalized estimating equations. Chapman & Hall, Cambridge
Jorgensen B, Lundbye-Christensen S, Song PX-K, Sun L (1996) State-space models for multivariate longitudinal data of mixed types. Can J Stat 24:385–402
Kim SK, Huggins R (1998) Diagnostic for autocorrelation regression models. Aust N Z J Stat 40:65–71
Kundu D, Balakrishnan N, Jamalizadeh A (2013) Generalized multivariate Birnbaum–Saunders distributions and related inferential results. J Multivar Anal 116:230–244
Leiva V (2016) The Birnbaum–Saunders distribution. Elsevier, Amsterdam

Leiva V, Barros M, Paula GA, Galea M (2007) Influence diagnostics in log-Birnbaum–Saunders regression models with censored data. Comput Stat Data Anal 51:5694–5707

Liang KY, Zeger SL (1986) Longitudinal analysis using generalized linear models. Biometrika 73:13–22

Liang NY, Zeger SL, Qaqish (1992) Multivariate regression analyses for categorical data. J R Stat Soc B 54:3–40

Manghi RF, Paula GA, Cysneiros JFA (2016) On elliptical multilevel models. J Appl Stat 43:2150–2171

Marchant C, Leiva V, Cysneiros FJA (2016) A multivariate log-linear model for Birnbaum–Saunders distributions. IEEE Trans Reliab 652:816–827

Marchant C, Leiva V, Cysneiros FJA, Vivanco JF (2016) Diagnostics in multivariate Birnbaum–Saunders regression models. J Appl Stat 43:2829–2849

Moral RA, Hinde J, Demétrio CGB (2017) Half-normal plots and overdispersed models in R: the hnp package. J Stat Softw 81:1–23

Munnell AH (1990) Why has productivity declined? Productivity and public investment. N Engl Econ Rev Jan/Feb 3–22

Paula GA, Medeiros M, Vilca-Labra FE (2009) Influence diagnostics for linear models with first-order autoregressive elliptical errors. Stat Probab Lett 79:339–346

Poon W, Poon YS (1999) Conformal normal curvature and assessment of local influence. J R Stat Soc B 61:51–61

Pregibon D (1981) Logistic regression diagnostics. Ann Stat 9:705–724

Preisser JS, Qaqish BF (1996) Deletion diagnostics for generalised estimating equations. Biometrika 83:551–562

R Core Team (2017) R: a language and environment for statistical computing. https://www.R-project.org/

Rieck JR, Nedelman JR (1991) A log-linear model for the Birnbaum–Saunders distribution. Technometrics 33:51–60

Venezuela MK, Botter DA, Sandoval MC (2007) Diagnostic techniques in generalized estimating equations. J Stat Comput Simul 77:879–888

Venezuela MK, Sandoval MC, Botter DA (2011) Local influence in estimating equations. Comput Stat Data Anal 55:1867–1883

Villegas C, Paula GA, Leiva V (2011) Birnbaum–Saunders mixed models for censored reliability data analysis. IEEE Trans Reliab 60:748–758