Electronic Journal of Statistics
Vol. 10 (2016) 2825–2855
ISSN: 1935-7524
DOI: 10.1214/16-EJS1187

Reparameterized Birnbaum-Saunders regression models with varying precision

Manoel Santos-Neto

Department of Statistics
Universidade Federal de Campina Grande
Campina Grande, Brazil
e-mail: manoel.ferreira@ufcg.edu.br
url: www.santosnetoce.com.br

Francisco José A. Cysneiros

Department of Statistics
Universidade Federal de Pernambuco
Recife, Brazil
e-mail: cysneiros@de.ufpe.br

Víctor Leiva*

Faculty of Engineering and Sciences
Universidad Adolfo Ibáñez, Viña del Mar, Chile
Faculty of Administration, Accounting and Economics
Universidade Federal de Goiás, Goiânia, Brazil
e-mail: victorleivasanchez@gmail.com
url: www.victorleiva.cl

and

Michelli Barros

Department of Statistics
Universidade Federal de Campina Grande
Campina Grande, Brazil
e-mail: michelli.karinne@gmail.com

Abstract: We propose a methodology based on a reparameterized Birnbaum-Saunders regression model with varying precision, which generalizes the existing works in the literature on the topic. This methodology includes the estimation of model parameters, hypothesis tests for the precision parameter, a residual analysis and influence diagnostic tools. Simulation studies are conducted to evaluate its performance. We apply it to two real-world case-studies to show its potential with the R software.

MSC 2010 subject classifications: Primary 62J12, 62J20; secondary 62F03.
Keywords and phrases: Birnbaum-Saunders distribution, hypothesis testing, likelihood-based methods, local influence, Monte Carlo simulation, residuals, R software.

Received July 2014.

*Corresponding author.

Contents

1. Introduction

The Birnbaum-Saunders (BS) distribution has been widely studied and applied; see the seminal paper by Birnbaum and Saunders [2] and the books by Johnson et al. [12, pp. 651–663] and Leiva [13]. The BS distribution is skewed positively, has nice properties and a close relation with the normal distribution. It was derived in terms of shape and scale parameters, but the latter is also its median. The original parameterization of the BS distribution is useful in several settings, for example, when modeling biological, environmental and fatigue data; see Owen and Padgett [25], Owen [24], Qu and Xie [28], Villegas et al. [46], Ferreira et al. [8], Li et al. [20], Saulo et al. [36], Leiva et al. [15, 14, 19], Garcia-Papani et al. [10] and Marchant et al. [22, 23].

Santos-Neto et al. [34] proposed several parameterizations for the BS distribution by using different arguments. One of them indexes the BS distribution by its mean and precision, which we name the reparameterized BS (RBS) distribution. As mentioned, the original BS parameterization is based on its shape and median and then the associated BS modeling is formulated by its median instead of its mean. However, in statistics, it is usual to model the mean. Thus, because the RBS distribution is parameterized by its mean, it can be used as a competitor of the normal distribution, but also of well-known asymmetrical distributions, such as gamma and lognormal. Therefore, the RBS distribution is useful in settings for which the original parameters are limited. For example, when modeling economic, financial and management data; see the works by Jin and Kawczak [11], Bhatti [1], Paula et al. [27], Leiva et al. [16, 18, 17], Santos-Neto et al. [35], Rojas et al. [32] and Wanke and Leiva [47], applications that, such as in the original parameterization, were conducted by international, transdisciplinary groups of researchers.

Note also that the original BS parameterization describes data based on a logarithmic transformation, inducting to an interpretation problem of the obtained results. Regression models are often concerned on the mean response and in its original scale, because there the interpretations become simpler. By using the RBS distribution, one can model the mean with no transformations similarly as in generalized linear models (GLM), but the BS and RBS distributions do not belong to the exponential family. However, a GLM type modeling based on the RBS distribution can also be carried out; see Leiva et al. [16]. Thus, the mean response is related to a linear predictor by one of several possible link functions, encompassing parameters to be estimated. Differently from all the existing BS regression models studied until now, the approach proposed by Leiva et al. [16] allows data to be modeled in their natural scale with a wide flexibility. In addition, the RBS distribution has properties that its competitor distributions of the exponential family do not have; see Subsection 2.1.

The RBS distribution has a precision parameter. Variability is often measured by dispersion parameters, but it can also be described by precision parameters, which are inversely proportional to the dispersion. Variability modeling has been widely discussed in the literature related to heteroscedasticity;

see Van Keilegom and Wang [43] and Saumard [37]. For example, Cook and Weisberg [3] studied heteroscedastic normal models. Taylor and Verbyla [42] described jointly the location and dispersion parameters of the Student-t model. Lin et al. [21] considered tests for heteroscedasticity in Student-t regression models. Wu et al. [49] proposed a method to select variables describing the mean and dispersion of lognormal models. In the context of GLM, Smyth [39] defined sub-models to describe the mean and dispersion, whereas Smyth and Verbyla [40] proposed an extension of GLM allowing both the mean and dispersion to be modeled. Cysneiros et al. [5] considered heteroscedastic symmetric linear models and their diagnostics. However, there exists few works modeling heteroscedasticity by precision parameters. Ferrari et al. [7] considered beta regression models for which the precision parameter is not constant across data and described it as a function of explanatory variables (covariates). Simas et al. [38] assumed a non-linear regression structure for the precision parameter by using a beta distribution, whereas Rocha and Simas [31] derived local influence in this model. Paula [26] modeled simultaneously the mean and precision of the gamma distribution, and carried out diagnostics in double generalized linear models.

For BS regressions based on its original parameterization, Rieck and Nedelman [30] and Galea et al. [9] assumed that the corresponding shape parameter is homogeneous across data. Xie and Wei [50] proposed a test for homogeneity of this shape parameter. Heterogeneous BS log-linear and non-linear regressions with both shape and scale parameters modeled by covariates were studied by Qu and Xie [28], Li et al. [20] and Vanegas et al. [44]. For the RBS regression model proposed by Leiva et al. [16], it is assumed that the precision is constant across data. Modeling of precision based on the RBS distribution has not been studied.

The main objective of this paper is to propose an RBS regression model with precision varying, allowing heteroscedasticity to be described, extending the work by Leiva et al. [16]. The specific objectives are (i) to estimate the parameters with the maximum likelihood (ML) method; (ii) to introduce hypothesis tests for the precision parameter and evaluate their performance; (iii) to present four types of residuals for the RBS model and study their distributions by Monte Carlo (MC) simulations; (iv) to analyze the sensitivity of the ML estimators to perturbations by using local influence (LI) and generalized leverage (GL) methods [see 45]; and (v) to apply the obtained results to two real-world case-studies with the `R` software; see `www.R-project.org` and R-Team [29].

In Section 2, we formulate the RBS regression model with varying precision, estimate its parameters and discuss gradient (GR), likelihood ratio (LR), score (SC) and Wald (WA) tests for the precision parameter. In Section 3, we conduct a diagnostic analysis, including residuals, GL and LI under case-weight, response and covariate perturbations. In Section 4, we carry out MC simulation studies to evaluate the performance of the proposed hypothesis tests and the empirical distribution of the residuals. In Section 5, we illustrate the potential applications of the proposed methodology by means of two real-world case-studies. In Section 6, we provide our conclusions and possible future work.

2. Modeling and inference

2.1. RBS distribution

A random variable T follows a BS distribution with shape ($\varphi > 0$) and scale ($\rho > 0$) parameters, which is denoted by $T \sim \mathrm{BS}(\varphi, \rho)$, if its probability density function (PDF) is given by

$$f_T(t; \varphi, \rho) = \frac{1}{\sqrt{2\,\pi}} \exp\left(-\frac{1}{2\varphi^2}\left[\frac{t}{\rho} + \frac{\rho}{t} - 2\right]\right) \frac{[t+\rho]}{2\varphi\sqrt{\rho\, t^3}}, \quad t > 0.$$

The mean and variance of T are $\mathrm{E}[T] = \rho[1 + \varphi^2/2]$ and $\mathrm{Var}[T] = \rho^2\varphi^2[1 + 5\varphi^2/4]$. Santos-Neto et al. [34] defined the RBS distribution by the parameters $\mu = \rho\,[1 + \varphi^2/2] > 0$ and $\delta = 2/\varphi^2 > 0$. Under this parameterization, we use the notation $Y \sim \mathrm{RBS}(\mu, \delta)$, with $\mathrm{E}[Y] = \mu$ and $\mathrm{Var}[Y] = \varpi(\mu)\,\varsigma(\delta)^2$, where $\varpi(\mu) = \mu^2$ acts as a "variance function" and $\varsigma(\delta) = \sqrt{[2\delta + 5]}/[\delta + 1]$ is the coefficient of variation (CV) of Y ($0 < \mathrm{CV}(Y) < \sqrt{5}$), which is function only of δ. In addition, δ plays the role of a precision parameter in the sense that, for fixed μ, as δ increases, the corresponding variance decreases. If $Y \sim \mathrm{RBS}(\mu, \delta)$, then its PDF and log-PDF are given respectively by

$$f_Y(y; \mu, \delta) = \frac{\exp(\delta/2)\sqrt{\delta+1}}{4\sqrt{\pi\mu}\, y^{3/2}}\left[y + \frac{\delta\mu}{\delta+1}\right] \exp\left(-\frac{\delta}{4}\left[\frac{y\{\delta+1\}}{\delta\mu} + \frac{\delta\mu}{y\{\delta+1\}}\right]\right),\ y > 0 \quad (2.1)$$

$$\ell(y; \mu, \delta) = \frac{\delta}{2} + \frac{\log(\delta+1)}{2} + \log\left(y + \frac{\delta\mu}{\delta+1}\right) - \frac{y[\delta+1]}{4\mu} - \frac{\mu\delta^2}{4y[\delta+1]} - \frac{\log(16\pi\,\mu\, y^3)}{2}. \quad (2.2)$$

In addition, the cumulative distribution function (CDF), quantile function (QF) and hazard rate of $Y \sim \mathrm{RBS}(\mu, \delta)$ are defined respectively as

$$F_Y(y; \mu, \delta) = \Phi(\sqrt{\{\delta/2\}}[\sqrt{\{\delta+1\}y/\{\mu\delta\}} - \sqrt{(\mu\delta/\{\delta+1\})y}]), \quad y > 0,$$

$$Q_Y(q; \mu, \delta) = [\delta\mu/\{\delta+1\}][Q_N(q)/\sqrt{2\delta} + \sqrt{\{Q_N(q)/\sqrt{2\delta}\}^2 + 1}]^2, \quad 0 < q < 1,$$

$$H_Y(y; \mu, \delta) = \frac{\exp(\delta/2)\sqrt{\delta+1}}{4\sqrt{\pi\mu y^3}}\left[y + \frac{\delta\mu}{\delta+1}\right] \frac{\exp\left(-\frac{\delta}{4}\left[\frac{y\{\delta+1\}}{\delta\mu} + \frac{\delta\mu}{y\{\delta+1\}}\right]\right)}{\Phi\left(-\sqrt{\frac{\delta}{2}}\left[\sqrt{\frac{\{\delta+1\}y}{\mu\delta}} - \sqrt{\frac{\mu\delta}{\{\delta+1\}y}}\right]\right)}, \quad y > 0,$$

where Φ and Q_N are the $\mathrm{N}(0, 1)$ CDF and QF, respectively. The RBS model has the properties: (i) $b\,Y \sim \mathrm{RBS}(b\,\mu, \delta)$, with $b > 0$; (ii) $1/Y \sim \mathrm{RBS}(\mu^\star, \delta)$, with $\mu^\star = [\delta+1]/[\delta\mu]$; (iii) as its QF has closed form, its median is $[\delta/\{\delta+1\}]\mu$; and (iv) its hazard rate has increasing, decreasing and upside-down shapes. All of these properties are shared by few distributions, in particular, several of them are not shared by exponential family distributions used in GLM. Moreover, note that, for modeling asymmetric data, the median might be more suitable than the mean. However, in the case of the RBS distribution, the median is proportional to its mean. Figure 1 presents the relation between the median and mean of the RBS distribution. From this figure, observe that, as the parameter δ increases, the two measures tend to be equal. For instance, from $\delta = 4$, we have that the median is 80% of the mean. Generally, in real-world applications, the estimate of δ is large, thus providing a good approximation between the median and mean. In the application of the work by Leiva et al. [16], the estimate of δ is 49.65, which means that the median represents 98.03% of the mean. In the two data analyses of the present paper, the estimates of δ are 82.38 and 35.32, representing 98.80% and 97.25% of mean, respectively. Thus, in practice, our model has no relevant loss of information by modeling the mean instead of the median. In addition, for the RBS regression model, we have the advantage of not transforming the response. Furthermore, a GLM type method, as that used in the RBS model [see 16], provides more flexibility in the functional relationship between the mean response and the linear predictor. More properties of the RBS distribution can be found in Santos-Neto et al. [35].

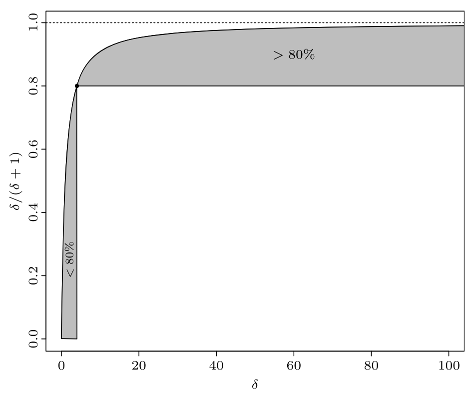

Fig 1. *Relation between median and mean of the RBS distribution.*

2.2. RBS regression model with varying precision

Let $\boldsymbol{Y} = [Y_1, \ldots, Y_n]^\top$ be a sample from an RBS population, that is, independent (IND) random variables but not independent identically distributed, such that, $Y_i \overset{\text{IND}}{\sim} \text{RBS}(\mu_i, \delta_i)$, for $i = 1, \ldots, n$, and $\boldsymbol{y} = [y_1, \ldots, y_n]^\top$ its observed value. The RBS regression model with varying precision can be written supposing that the corresponding mean and precision parameters satisfy, respectively, the functional relations

$$g(\mu_i) = \eta_i = \boldsymbol{x}_i^\top \boldsymbol{\beta}, \quad h(\delta_i) = \tau_i = \boldsymbol{z}_i^\top \boldsymbol{\alpha}, \quad i = 1, \ldots, n, \tag{2.3}$$

where $\mu_i = g^{-1}(\boldsymbol{x}_i^\top \boldsymbol{\beta})$, with $\boldsymbol{\beta} = [\beta_1, \ldots, \beta_p]^\top$ being a $p \times 1$ vector of unknown parameters to be estimated, and $\boldsymbol{x}_i = [1, x_{i2}, \ldots, x_{ip}]^\top$ being a $p \times 1$ vector that contains the values of p covariates, which can be summarized as $\boldsymbol{X} = [x_{ij}]$. In addition, $\delta_i = h^{-1}(\boldsymbol{z}_i^\top \boldsymbol{\alpha})$, with $\boldsymbol{\alpha} = [\alpha_1, \ldots, \alpha_q]^\top$ being also a $q \times 1$ vector of unknown parameters to be estimated, and $\boldsymbol{z}_i = [1, z_{i2}, \ldots, z_{iq}]^\top$ being a $q \times 1$ vector that contains the values of q covariates, which can be summarized as $\boldsymbol{Z} = [z_{ij}]$. Note that $p+q < n$. The link functions g: $\mathbb{R} \to \mathbb{R}^+$ and h: $\mathbb{R} \to \mathbb{R}^+$ defined in (2.3) must be strictly monotone, positive and at least twice differentiable. Table 1 lists the most common link functions for g and h along with their first and second derivatives.

TABLE 1
Derivatives for the indicated link function.

Link function	g	h	$\mathrm{d}\mu/\mathrm{d}\eta$	$\mathrm{d}\delta/\mathrm{d}\tau$	$\mathrm{d}^2\mu/\mathrm{d}\eta^2$	$\mathrm{d}^2\delta/\mathrm{d}\tau^2$
Identity	$\mu=\eta$	$\delta=\tau$	1	1	0	0
Logarithm	$\log(\mu)=\eta$	$\log(\delta)=\tau$	μ	δ	μ	δ
Squared root	$\sqrt{\mu}=\eta$	$\sqrt{\delta}=\tau$	$2\sqrt{\mu}$	$2\sqrt{\delta}$	2	2

2.3. Estimation

The log-likelihood function for the parameter $\boldsymbol{\theta}=[\boldsymbol{\beta}^\top,\boldsymbol{\alpha}^\top]^\top$ obtained from (2.2), related to $\boldsymbol{\mu}=[\mu_1,\ldots,\mu_n]^\top$ and $\boldsymbol{\delta}=[\delta_1,\ldots,\delta_n]^\top$ for the class of models with link functions in (2.3), is given by

$$\ell(\boldsymbol{\theta})=\sum_{i=1}^{n}\ell_i(\mu_i,\delta_i;y_i), \tag{2.4}$$

where (ignoring the constant term)

$$\begin{aligned}\ell_i(\mu_i,\delta_i;y_i) \quad = \quad & \frac{\delta_i}{2}-\frac{\log(\delta_i+1)}{2}-\frac{\log(\mu_i)}{2}-\frac{3\log(y_i)}{2}+\log(\delta_i y_i+y_i+\delta_i\mu_i)\\ & -\frac{y_i[\delta_i+1]}{4\mu_i}-\frac{\delta_i^2\mu_i}{4y_i[\delta_i+1]}.\end{aligned}$$

The $[p+q]\times 1$ score vector with first derivatives is obtained from (2.4) as

$$\dot{\boldsymbol{\ell}}(\boldsymbol{\theta})=\begin{bmatrix}\dot{\boldsymbol{\ell}}(\boldsymbol{\beta})\\ \dot{\boldsymbol{\ell}}(\boldsymbol{\alpha})\end{bmatrix}=\begin{bmatrix}\frac{\partial\ell(\boldsymbol{\theta})}{\partial\boldsymbol{\beta}}\\ \frac{\partial\ell(\boldsymbol{\theta})}{\partial\boldsymbol{\alpha}}\end{bmatrix}=\begin{bmatrix}\boldsymbol{X}^\top\boldsymbol{A}(\boldsymbol{y}^\star-\boldsymbol{\mu}^\star)\\ \boldsymbol{Z}^\top\boldsymbol{B}(\boldsymbol{y}^\star-\boldsymbol{\delta}^\star)\end{bmatrix};$$

see details in Appendix A.1. The $[p+q]\times[p+q]$ Hessian matrix with second derivatives also is obtained from (2.4) as

$$\ddot{\boldsymbol{\ell}}(\boldsymbol{\theta})=\begin{bmatrix}\ddot{\boldsymbol{\ell}}(\boldsymbol{\beta}) & \ddot{\boldsymbol{\ell}}(\boldsymbol{\beta\alpha})\\ \ddot{\boldsymbol{\ell}}(\boldsymbol{\alpha\beta}) & \ddot{\boldsymbol{\ell}}(\boldsymbol{\alpha})\end{bmatrix}=\begin{bmatrix}\frac{\partial^2\ell(\boldsymbol{\theta})}{\partial\boldsymbol{\beta}\partial\boldsymbol{\beta}^\top} & \frac{\partial^2\ell(\boldsymbol{\theta})}{\partial\boldsymbol{\beta}\partial\boldsymbol{\alpha}^\top}\\ \frac{\partial^2\ell(\boldsymbol{\theta})}{\partial\boldsymbol{\alpha}\partial\boldsymbol{\beta}^\top} & \frac{\partial^2\ell(\boldsymbol{\theta})}{\partial\boldsymbol{\alpha}\partial\boldsymbol{\alpha}^\top}\end{bmatrix}=\begin{bmatrix}\boldsymbol{X}^\top\boldsymbol{C}\boldsymbol{X} & \boldsymbol{X}^\top\boldsymbol{M}\boldsymbol{Z}\\ \boldsymbol{Z}^\top\boldsymbol{M}\boldsymbol{X} & \boldsymbol{Z}^\top\boldsymbol{W}\boldsymbol{Z}\end{bmatrix};$$

see details in Appendix A.2. The corresponding expected Fisher information matrix is given by

$$\boldsymbol{i}(\boldsymbol{\theta})=\begin{bmatrix}\boldsymbol{i}(\boldsymbol{\beta}) & \boldsymbol{i}(\boldsymbol{\beta\alpha})\\ \boldsymbol{i}(\boldsymbol{\alpha\beta}) & \boldsymbol{i}(\boldsymbol{\alpha})\end{bmatrix}=\begin{bmatrix}\boldsymbol{X}^\top\boldsymbol{V}\boldsymbol{X} & \boldsymbol{X}^\top\boldsymbol{S}\boldsymbol{Z}\\ \boldsymbol{Z}^\top\boldsymbol{S}\boldsymbol{X} & \boldsymbol{Z}^\top\boldsymbol{U}\boldsymbol{Z}\end{bmatrix};$$

see detail in Appendix A.2. The ML estimates of $\boldsymbol{\beta}$ and $\boldsymbol{\alpha}$ are computed as the solution of the non-linear system $\dot{\boldsymbol{\ell}}(\boldsymbol{\theta})=\mathbf{0}_{[p+q]\times 1}$, where $\mathbf{0}_{[p+q]\times 1}$ is $[p+$

$q] \times 1$ vector of zeros. In practice, the ML estimates can be obtained through a numerical maximization of the log-likelihood function. Consider the $2n \times 2n$ augmented matrix of the form

$$\tilde{\boldsymbol{W}} = \begin{bmatrix} \tilde{\boldsymbol{W}}(\boldsymbol{\beta}) & \tilde{\boldsymbol{W}}(\boldsymbol{\beta\alpha}) \\ \tilde{\boldsymbol{W}}(\boldsymbol{\alpha\beta}) & \tilde{\boldsymbol{W}}(\boldsymbol{\alpha}) \end{bmatrix} = \begin{bmatrix} \boldsymbol{V} & \boldsymbol{S} \\ \boldsymbol{S} & \boldsymbol{U} \end{bmatrix}, \tag{2.5}$$

where the elements of $\boldsymbol{V}$, $\boldsymbol{S}$ and $\boldsymbol{U}$ are given in Equation (A.2). Also consider $\tilde{\boldsymbol{X}}$ as being another $2n \times [p+q]$ augmented matrix of the form

$$\tilde{\boldsymbol{X}} = \begin{bmatrix} \boldsymbol{X} & \boldsymbol{0}_{p\times 1} \\ \boldsymbol{0}_{q\times 1} & \boldsymbol{Z} \end{bmatrix}. \tag{2.6}$$

Therefore, from (2.6) and (2.5), the Fisher information matrix given in Equation (A.2) and its inverse can be rewritten as $\boldsymbol{i}(\boldsymbol{\theta}) = \tilde{\boldsymbol{X}}^\top \tilde{\boldsymbol{W}} \tilde{\boldsymbol{X}}$ and $\boldsymbol{i}(\boldsymbol{\theta})^{-1} = [\tilde{\boldsymbol{X}}^\top \tilde{\boldsymbol{W}} \tilde{\boldsymbol{X}}]^{-1}$, respectively; see details of the matrix $\boldsymbol{i}(\boldsymbol{\theta})^{-1}$ in Appendix A.2. By using the Fisher scoring iterative procedure, the corresponding estimation algorithm for $\boldsymbol{\theta} = [\boldsymbol{\beta}^\top, \boldsymbol{\alpha}^\top]^\top$ is given by

$$\begin{aligned} \boldsymbol{\theta}^{(m+1)} &= \boldsymbol{\theta}^{(m)} + \Big[\boldsymbol{i}(\boldsymbol{\theta})^{(m)}\Big]^{-1} \dot{\boldsymbol{\ell}}(\boldsymbol{\theta})^{(m)} \\ &= \boldsymbol{\theta}^{(m)} + [\tilde{\boldsymbol{X}}^\top \tilde{\boldsymbol{W}}^{(m)} \tilde{\boldsymbol{X}}]^{-1} \begin{bmatrix} \boldsymbol{X}^\top \boldsymbol{A}^{(m)} (\boldsymbol{y}^\star - \boldsymbol{\mu}^\star)^{(m)} \\ \boldsymbol{Z}^\top \boldsymbol{B}^{(m)} (\boldsymbol{y}^\star - \boldsymbol{\delta}^\star)^{(m)} \end{bmatrix}, \\ &= [\tilde{\boldsymbol{X}}^\top \tilde{\boldsymbol{W}}^{(m)} \tilde{\boldsymbol{X}}]^{-1} \tilde{\boldsymbol{W}}^{(m)} \tilde{\boldsymbol{y}}^{\star(m)}, \end{aligned} \tag{2.7}$$

where

$$\tilde{\boldsymbol{y}}^{\star(m)} = \tilde{\boldsymbol{X}} \boldsymbol{\theta}^{(m)} + [\tilde{\boldsymbol{W}}^{(m)}]^{-1} \begin{bmatrix} \boldsymbol{A}^{(m)} & \boldsymbol{0}_{p\times 1} \\ \boldsymbol{0}_{q\times 1} & \boldsymbol{B}^{(m)} \end{bmatrix} \begin{bmatrix} (\boldsymbol{y}^\star - \boldsymbol{\mu}^\star)^{(m)} \\ (\boldsymbol{y}^\star - \boldsymbol{\delta}^\star)^{(m)} \end{bmatrix}.$$

Note that $\boldsymbol{\theta}^{(m+1)}$ given in (2.7) has the form of a reweighted least square (LS) estimate, where $\tilde{\boldsymbol{y}}^\star$ is the modified response variable. We propose to use the Cole-Green algorithm to conduct with the iterative procedure defined in (2.7). This algorithm can be better for distributions with highly correlated parameter estimators and equivalent to the Fisher scoring algorithm, when we are in the fully parametric case; see details about the Cole-Green algorithm in [41].

2.4. Inference

Under usual regularity conditions [see 4], the ML estimators of $\boldsymbol{\theta}$ and $\boldsymbol{i}(\boldsymbol{\theta})$, $\widehat{\boldsymbol{\theta}}$ and $\boldsymbol{i}(\widehat{\boldsymbol{\theta}})$ say, respectively, are consistent. Suppose that $\boldsymbol{j}(\boldsymbol{\theta}) = \lim_{n\to\infty}[1/n]\boldsymbol{i}(\boldsymbol{\theta})$ exists and is non-singular. Then, as $n \to \infty$,

$$\sqrt{n}[\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}] \xrightarrow{\mathcal{D}} \mathrm{N}_{p+q}(\boldsymbol{0}_{[p+q]\times 1}, \boldsymbol{j}(\boldsymbol{\theta})^{-1}), \tag{2.8}$$

where $\xrightarrow{\mathcal{D}}$ denotes convergence in distribution to. Thus, if θ_k denotes the kth element of $\boldsymbol{\theta}$, $[\widehat{\theta}_k - \theta_k]/\{i_{kk}\}^{1/2} \xrightarrow{\mathcal{D}} \mathrm{N}(0,1)$, where i_{kk} is kth diagonal element of the matrix $\boldsymbol{i}(\boldsymbol{\theta})^{-1}$, for $k = 1, \ldots, p+q$. Note that an asymptotic $100 \times [1-\gamma]\%$ confidence region for $\boldsymbol{\theta}$ is $[\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}]^\top \boldsymbol{i}(\widehat{\boldsymbol{\theta}})^{-1}[\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}] \leq \chi^2_{1-\gamma}(p+q)$, where $\boldsymbol{\theta} \in \mathbb{R}^{p+q}$ and $\chi^2_{1-\gamma}(p+q)$ is the $[1-\gamma]$th quantile of the chi-squared distribution with $p+q$ degrees of freedom. From the previous asymptotic distributional result given in (2.8), we have that

$$\sqrt{n}[\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}] \xrightarrow{\mathcal{D}} \mathrm{N}_p(\mathbf{0}_{p\times 1}, \boldsymbol{j}(\boldsymbol{\beta})^{-1}), \quad \sqrt{n}[\widehat{\boldsymbol{\alpha}} - \boldsymbol{\alpha}] \xrightarrow{\mathcal{D}} \mathrm{N}_q(\mathbf{0}_{q\times 1}, \boldsymbol{j}(\boldsymbol{\alpha})^{-1}).$$

Then, asymptotic $100[1-\gamma]\%$ confidence intervals for $\widehat{\beta}_j$ and $\widehat{\alpha}_r$ are $\widehat{\beta}_j \pm \xi_{1-\gamma/2}/\{i_{jj}\}^{1/2}$ and $\widehat{\alpha}_r \pm \xi_{1-\gamma/2}/\{i_{rr}\}^{1/2}$, respectively, with $j = 1, \ldots, p$, $r = 1, \ldots, q$ and ξ_γ being the γth N(0,1) quantile. We can obtain $100 \times [1-\gamma]\%$ confidence bands for the linear predictor $\mu(\boldsymbol{x}_{\text{pred}}) = g^{-1}(\boldsymbol{x}_{\text{pred}}^\top \boldsymbol{\beta})$, where $\boldsymbol{x}_{\text{pred}} \in \mathbb{R}^p$ is an arbitrary vector. Then, an asymptotic $100 \times [1-\gamma]\%$ confidence region for $\mu(\boldsymbol{x}_{\text{pred}})$ is obtained as

$$\Bigg[g^{-1}\left(\boldsymbol{x}_{\text{pred}}^\top \widehat{\boldsymbol{\beta}} - \sqrt{\chi^2_{1-\gamma}(p)}\left\{\boldsymbol{x}_{\text{pred}}^\top \boldsymbol{i}(\widehat{\boldsymbol{\beta}})\, \boldsymbol{x}_{\text{pred}}\right\}^{\frac{1}{2}}\right), \quad g^{-1}\left(\boldsymbol{x}_{\text{pred}}^\top \widehat{\boldsymbol{\beta}} + \sqrt{\chi^2_{1-\gamma}(p)}\left\{\boldsymbol{x}_{\text{pred}}^\top \boldsymbol{i}(\widehat{\boldsymbol{\beta}})\, \boldsymbol{x}_{\text{pred}}\right\}^{\frac{1}{2}}\right)\Bigg],$$

where $\boldsymbol{i}(\widehat{\boldsymbol{\beta}}) = [\boldsymbol{X}^\top \widehat{\boldsymbol{W}_3} \boldsymbol{X}]^{-1}$, with $\widehat{\boldsymbol{W}_3} = \widehat{\boldsymbol{V}} - \widehat{\boldsymbol{S}}\boldsymbol{Z}[\boldsymbol{Z}^\top \widehat{\boldsymbol{U}} \boldsymbol{Z}]^{-1}\boldsymbol{Z}^\top \widehat{\boldsymbol{S}}$. Analogously, an asymptotic $100 \times [1-\gamma]\%$ confidence region for $\delta(\boldsymbol{z}_{\text{pred}}) = h^{-1}(\boldsymbol{z}_{\text{pred}}^\top \boldsymbol{\alpha})$ is given by

$$\Bigg[h^{-1}\left(\boldsymbol{z}_{\text{pred}}^\top \widehat{\boldsymbol{\alpha}} - \sqrt{\chi^2_{1-\gamma}(q)}\left\{\boldsymbol{z}_{\text{pred}}^\top \boldsymbol{i}(\widehat{\boldsymbol{\alpha}})\, \boldsymbol{z}_{\text{pred}}\right\}^{\frac{1}{2}}\right), \quad h^{-1}\left(\boldsymbol{z}_{\text{pred}}^\top \widehat{\boldsymbol{\alpha}} + \sqrt{\chi^2_{1-\gamma}(q)}\left\{\boldsymbol{z}_{\text{pred}}^\top \boldsymbol{i}(\widehat{\boldsymbol{\alpha}})\, \boldsymbol{z}_{\text{pred}}\right\}^{\frac{1}{2}}\right)\Bigg],$$

where $\boldsymbol{z}_{\text{pred}} \in \mathbb{R}^q$ and $\boldsymbol{i}(\widehat{\boldsymbol{\alpha}}) = [\boldsymbol{Z}^\top \widehat{\boldsymbol{W}_4} \boldsymbol{Z}]^{-1}$, with $\widehat{\boldsymbol{W}_4} = \widehat{\boldsymbol{U}} - \widehat{\boldsymbol{S}}\boldsymbol{X}[\boldsymbol{X}^\top \widehat{\boldsymbol{V}} \boldsymbol{X}]^{-1}\boldsymbol{X}^\top \widehat{\boldsymbol{S}}$.

2.5. Hypothesis testing

We want to test if the model precision is constant, that is, our hypotheses are H_0: $\boldsymbol{\alpha}_0 \in \boldsymbol{\Omega}_0$ versus H_A: $\boldsymbol{\alpha}_0 \notin \boldsymbol{\Omega}_0$, where $\boldsymbol{\Omega}_0 = \{\boldsymbol{\theta}\colon \alpha_i = 0, i = 2, \ldots, q\}$, $\boldsymbol{\Omega} = \boldsymbol{\Omega}_0 \cup \boldsymbol{\Omega}_0^c$ and $\boldsymbol{\alpha}_0 = [\alpha_2, \ldots, \alpha_q]^\top$. For large n, under H_0 and usual regularity conditions, the GR, LR, SC and WA statistics follow the χ^2_{q-1} distribution. Thus, the above hypotheses can be tested with critical values based on this distribution.

We use $\widetilde{\boldsymbol{\theta}}$ for denoting the ML estimator of $\boldsymbol{\theta}$ evaluated at H_0 (restricted case), or for any function or element of this vector, and $\widehat{\boldsymbol{\theta}}$ for the ML estimator of $\boldsymbol{\theta}$ under H_A (unrestricted case).

2.5.1. The GR test

Its statistic to test H_0 versus H_A is given by

$$\mathrm{GR} = [\boldsymbol{y}^\star - \widetilde{\boldsymbol{\delta}}^\star]^\top \widetilde{\boldsymbol{B}} \boldsymbol{Z}_0 \widehat{\boldsymbol{\alpha}}_0, \tag{2.9}$$

where $\boldsymbol{Z}_0$ is composed by the elements of the $q-1$ last columns of $\boldsymbol{Z}$ and $\boldsymbol{y}^\star, \widetilde{\boldsymbol{\delta}}^\star$ are defined in Appendix A.1.

2.5.2. The LR test

Its statistics to test H_0 versus H_A is defined as

$$\mathrm{LR} = 2\Big[\sup_{\boldsymbol{\theta}\in\boldsymbol{\Omega}} \ell(\boldsymbol{\theta}) - \sup_{\boldsymbol{\theta}\in\boldsymbol{\Omega}_0} \ell(\boldsymbol{\theta})\Big] = 2[\ell(\widehat{\boldsymbol{\theta}}) - \ell(\widetilde{\boldsymbol{\theta}})] = 2\sum_{i=1}^{n} \Lambda_i, \tag{2.10}$$

where

$$\Lambda_i = \tfrac{\widehat{\delta}_i - \tilde{\delta}_i}{2} + \log\left(\tfrac{\tilde{\mu}_i}{\widehat{\mu}_i}\right) + \log\left(\tfrac{\widehat{\delta}_i y_i + y_i + \widehat{\delta}_i \widehat{\mu}_i}{\tilde{\delta}_i y_i + y_i + \tilde{\delta}_i \tilde{\mu}_i}\right) + \tfrac{y_i}{4}\left[\tfrac{\tilde{\delta}_i + 1}{\tilde{\mu}_i} - \tfrac{\widehat{\delta}_i + 1}{\widehat{\mu}_i}\right] + \tfrac{1}{4y_i}\left[\tfrac{\tilde{\delta}_i \tilde{\mu}_i}{\tilde{\delta}_i + 1} - \tfrac{\widehat{\delta}_i \tilde{\mu}_i}{\widehat{\delta}_i + 1}\right].$$

2.5.3. The SC test

Its statistic to test H_0 versus H_A can be expressed as

$$\mathrm{SC} = \dot{\boldsymbol{\ell}}(\widetilde{\boldsymbol{\alpha}}_0)^\top \widetilde{\mathrm{Var}}(\widetilde{\boldsymbol{\alpha}}_0) \dot{\boldsymbol{\ell}}(\widetilde{\boldsymbol{\alpha}}_0) = [\boldsymbol{y}^\star - \widetilde{\boldsymbol{\delta}}^\star]^\top \widetilde{\boldsymbol{B}} \boldsymbol{Z}_0 \left[\boldsymbol{Z}_0^\top \widetilde{\boldsymbol{W}}_4 \boldsymbol{Z}_0\right]^{-1} \boldsymbol{Z}_0^\top \widetilde{\boldsymbol{B}} [\boldsymbol{y}^\star - \widetilde{\boldsymbol{\delta}}^\star]. \tag{2.11}$$

2.5.4. The WA test

Its statistic to test H_0 versus H_A is obtained as

$$\mathrm{WA} = \widehat{\boldsymbol{\alpha}}_0^\top \widehat{\mathrm{Var}}(\widehat{\boldsymbol{\alpha}}_0)^{-1} \widehat{\boldsymbol{\alpha}}_0 = \widehat{\boldsymbol{\alpha}}_0^\top \boldsymbol{Z}_0^\top \widehat{\boldsymbol{W}}_4 \boldsymbol{Z}_0 \widehat{\boldsymbol{\alpha}}_0. \tag{2.12}$$

3. Diagnostic analysis

3.1. Residuals

We introduce residuals for the model with link function defined in (2.3) based on the standardized Pearson, score and quantile residuals given respectively by

$$r_i^{\mathrm{p}} = \widehat{\phi}_i^{1/2}[y_i - \widehat{\mu}_i]/[\varpi(\widehat{\mu}_i)]^{1/2}, \quad r_i^{\mathrm{s}} = [y_i^\star - \widehat{\mu}_i^\star]/\widehat{\nu}_i^{1/2}, \quad r_i^{\mathrm{q}} = \Phi^{-1}(F_Y(y_i; \widehat{\mu}_i, \widehat{\delta}_i)),$$

where $\varpi(\mu_i) = \mu_i^2$ and $\phi_i = [\delta_i + 1]^2/[2\delta_i + 5]$, with $\widehat{\mu}_i = g^{-1}(\widehat{\eta}_i)$ and $\widehat{\delta}_i = h^{-1}(\widehat{\tau}_i)$ being the ML estimates of μ_i and δ_i, respectively. In addition, $\widehat{\nu}_i = \widehat{\delta}_i/[2\widehat{\mu}_i^2] + \widehat{\delta}_i^{\,2}/[\widehat{\delta}_i + 1]^2 I(\widehat{\mu}, \widehat{\delta})$, with $I(\widehat{\mu}, \widehat{\delta}) = \int_0^\infty [y + \{\widehat{\mu}\widehat{\delta}\}/\{\widehat{\delta} + 1\}]^{-2} \cdot f_Y(y; \widehat{\mu}, \widehat{\delta})\,\mathrm{d}y$. Furthermore, F_Y is the RBS CDF; see details about the quantile residual in Dunn and Smyth [6]. For the score residual, see details in Appendix A.3. Moreover, we propose a deviance component type residual, assuming that the precision parameter has a structure of regression known, by

$$r_i^{\mathrm{d}} = \mathrm{sign}(y_i - \widehat{\mu}_i)\sqrt{2}\left[\log(2) - \frac{\widehat{\delta}_i}{2} + \frac{\{\widehat{\delta}_i + 1\}y_i}{4\widehat{\mu}_i} + \frac{\widehat{\delta}_i^2\widehat{\mu}_i}{4\{\widehat{\delta}_i + 1\}y_i} + \frac{1}{2}\log\left(\frac{\widehat{\delta}_i\{\widehat{\delta}_i + 1\}y_i\,\widehat{\mu}_i}{\{\widehat{\delta}_i\,y_i + y_i + \widehat{\delta}_i\widehat{\mu}_i\}^2}\right)\right]^{1/2},$$

where the sign function is defined as $\mathrm{sign}(x) = \{-1, 0, 1\}$, for $x < 0$, $x = 0$ and $x > 0$, respectively. Residual are used to verify model adequacy, to test nonlinearity and to detect outliers, autocorrelation and heteroscedasticity, plotting them versus covariates individually, or versus $\widehat{\mu}_i$ or $\widehat{\eta}_i$.

3.2. Generalized leverage

GL of an estimator is defined in regression as a measure of the importance of individual cases, evaluating the influence of the observed response on its own estimated value. GL can be obtained in a general form as $\partial\widehat{\boldsymbol{y}}/\partial\boldsymbol{y} = \boldsymbol{D}(\boldsymbol{\theta})[-\ddot{\boldsymbol{\ell}}(\boldsymbol{\theta})]^{-1}\ddot{\boldsymbol{\ell}}(\boldsymbol{\theta}\boldsymbol{y})|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}}$, where $\boldsymbol{D}(\boldsymbol{\theta}) = \partial\boldsymbol{\mu}/\partial\boldsymbol{\theta}$ and $\ddot{\boldsymbol{\ell}}(\boldsymbol{\theta}\boldsymbol{y}) = \partial^2\ell(\boldsymbol{\theta})/\partial\boldsymbol{\theta}\partial\boldsymbol{y}^\top$; see details of the matrix $\ddot{\boldsymbol{\ell}}(\boldsymbol{\theta})^{-1}$ in Appendix A.2. In the RBS regression model with varying precision, the GL matrix is given by

$$\mathrm{GL}(\widehat{\boldsymbol{\theta}}) = \widehat{\boldsymbol{A}}\boldsymbol{X}[-\ddot{\boldsymbol{\ell}}(\widehat{\boldsymbol{\beta}})]\boldsymbol{X}^\top\widehat{\boldsymbol{A}}\widehat{\boldsymbol{E}} + \widehat{\boldsymbol{A}}\boldsymbol{X}[-\ddot{\boldsymbol{\ell}}(\widehat{\boldsymbol{\beta}}\widehat{\boldsymbol{\alpha}})]\boldsymbol{Z}^\top\widehat{\boldsymbol{B}}\widehat{\boldsymbol{L}}, \tag{3.1}$$

where $\ddot{\boldsymbol{\ell}}(\boldsymbol{\beta})$ and $\ddot{\boldsymbol{\ell}}(\boldsymbol{\beta}\boldsymbol{\alpha})$ are detailed in Appendix A.2, $\boldsymbol{E} = [d_{y\mu}^{(i)}\,\delta_{ij}^{n}]$ and $\boldsymbol{L} = [d_{y\delta}^{(i)}\,\delta_{ij}^{n}]$, with

$$\begin{aligned} d_{y\mu}^{(i)} &= \frac{\delta_i + 1}{4\,\mu_i^2} + \frac{\delta_i^2}{4\,y_i^2\,[\delta_i + 1]} - \frac{\delta_i\,[\delta_i + 1]}{[\delta_i\,y_i + y_i + \delta_i\,\mu_i]^2}, \\ d_{y\delta}^{(i)} &= \frac{\mu_i\,\delta_i\,[\delta_i + 2]}{4\,y_i^2\,[\delta_i + 1]^2} - \frac{1}{4\,\mu_i} - \frac{\mu_i}{[\delta_i\,y_i + y_i + \delta_i\,\mu_i]^2}. \end{aligned} \tag{3.2}$$

Considering the results presented in (3.1), the GL given in (3.2) can be rewritten as $\mathrm{GL}(\widehat{\boldsymbol{\theta}}) = \mathrm{GL}(\widehat{\boldsymbol{\beta}})[[\widehat{\boldsymbol{A}}\widehat{\boldsymbol{E}}]^{-1}\widehat{\boldsymbol{M}}\boldsymbol{Z}[\boldsymbol{Z}^\top\widehat{\boldsymbol{W}}\boldsymbol{Z}]^{-1}\boldsymbol{Z}^\top\widehat{\boldsymbol{B}}\widehat{\boldsymbol{L}} - \boldsymbol{I}_n]$, where $\mathrm{GL}(\widehat{\boldsymbol{\beta}}) = \widehat{\boldsymbol{A}}\boldsymbol{X}[\boldsymbol{X}^\top\widehat{\boldsymbol{W}}_1\boldsymbol{X}]^{-1}\boldsymbol{X}^\top\widehat{\boldsymbol{A}}\widehat{\boldsymbol{E}}$ is the GL considering only the vector $\boldsymbol{\beta}$ and $\boldsymbol{I}_n$ is the $n \times n$ identity matrix.

3.3. Local influence

For the RBS model given in (2.3), let $\ell(\boldsymbol{\theta}|\boldsymbol{\omega})$ be log-likelihood function corresponding to this model perturbed by $\boldsymbol{\omega}$. The perturbation vector $\boldsymbol{\omega}$ belongs to a subset of $\mathbb{R}^n$ and $\boldsymbol{\omega}_0$ is an $n \times 1$ non-perturbation vector, such that $\ell(\boldsymbol{\theta}|\boldsymbol{\omega}_0) = \ell(\boldsymbol{\theta})$, for all $\boldsymbol{\theta}$. The likelihood displacement (LD) defined as $\mathrm{LD}(\boldsymbol{\omega}) = 2[\ell(\widehat{\boldsymbol{\theta}}) - \ell(\widehat{\boldsymbol{\theta}}_\omega)]$, where $\widehat{\boldsymbol{\theta}}_\omega$ denotes the ML estimate of $\boldsymbol{\theta}$ upon the perturbed RBS model, can be used to assess the influence of the perturbation on the ML estimate. The normal curvature for $\boldsymbol{\theta}$ in the direction vector $\boldsymbol{l}$, with $||\boldsymbol{l}|| = 1$, is expressed as $C_l(\boldsymbol{\theta}) = 2|\boldsymbol{l}^\top \boldsymbol{\Delta}^\top \ddot{\boldsymbol{\ell}}(\boldsymbol{\theta})^{-1} \boldsymbol{\Delta}\, \boldsymbol{l}|$, where $\boldsymbol{\Delta}$ is a $[p+q] \times n$ matrix of perturbations with elements $\Delta_{ji} = \partial^2 \ell(\boldsymbol{\theta}|\boldsymbol{\omega})/\partial\theta_j \partial\omega_i$ and $\ddot{\boldsymbol{\ell}}(\boldsymbol{\theta})$ is detailed in (A.1) of Appendix B. Note that $\boldsymbol{\Delta}$ must be evaluated at $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}$ and $\boldsymbol{\omega} = \boldsymbol{\omega}_0$, for $j = 1, \ldots, p+1$ and $i = 1, \ldots, n$. An LI diagnostic is generally based on index plots. For instance, the index graph of the eigenvector $\boldsymbol{l}_{\max}$ corresponding to the maximum eigenvalue of

$$\boldsymbol{F}(\boldsymbol{\theta}) = \boldsymbol{\Delta}^\top \ddot{\boldsymbol{\ell}}(\boldsymbol{\theta})^{-1} \boldsymbol{\Delta}, \tag{3.3}$$

$C_{l_{\max}}(\boldsymbol{\theta})$ say, evaluated at $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}$, can detect those cases that under small perturbations exercise a great influence on $\mathrm{LD}(\boldsymbol{\omega})$. In addition to the direction vector of maximum normal curvature, $\boldsymbol{l}_{\max}$ say, another direction of interest is $\boldsymbol{l}_i = \boldsymbol{e}_{in}$, which corresponds to the direction of the case i, where $\boldsymbol{e}_{in}$ is an $n \times 1$ vector of zeros with a value equal to one at the ith position, that is, $\{\boldsymbol{e}_{in}, 1 \le i \le n\}$ is the canonical basis of $\mathbb{R}^n$. In this case, the normal curvature is given by $C_i(\boldsymbol{\theta}) = 2|f_{ii}|$, where f_{ii} is the ith diagonal element of $\boldsymbol{F}(\boldsymbol{\theta})$ defined in (3.3), for $i = 1, \ldots, n$, evaluated at $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}$. Those cases when $C_i(\widehat{\boldsymbol{\theta}}) > 2\overline{C}(\widehat{\boldsymbol{\theta}})$, where $\overline{C}(\widehat{\boldsymbol{\theta}}) = \sum_{i=1}^n C_i(\widehat{\boldsymbol{\theta}})/n$, are considered as potentially influential. This procedure is called the total LI of the case i. For the indicated scheme, Table 2 presents the respective perturbation matrix given in general by

$$\boldsymbol{\Delta} = \begin{bmatrix} \boldsymbol{\Delta}(\boldsymbol{\beta}) \\ \boldsymbol{\Delta}(\boldsymbol{\alpha}) \end{bmatrix}, \tag{3.4}$$

which must be evaluated at the non-perturbation vector $\boldsymbol{\omega}_0$ and at $\widehat{\boldsymbol{\theta}}$.

TABLE 2
Matrices of perturbations for the indicated scheme in the RBS model with varying precision.

Scheme	Form	Matrices
Case-weight	$\sum_{i=1}^n \omega_i \ell_i(\boldsymbol{\theta})$	$\boldsymbol{\Delta}(\boldsymbol{\beta}) = \boldsymbol{X}^\top\, a_i\, d_\mu^{(i)} \delta_{ij}^n$ $\boldsymbol{\Delta}(\boldsymbol{\alpha}) = \boldsymbol{Z}^\top\, b_i\, d_\delta^{(i)} \delta_{ij}^n$
Response	$Y_i + \omega_i S_{Y_i}$	$\boldsymbol{\Delta}(\boldsymbol{\beta}) = \boldsymbol{X}^\top\, a_i\, d_{y\mu}^{(i)}\, S_{Y_i}\, \delta_{ij}^n$ $\boldsymbol{\Delta}(\boldsymbol{\alpha}) = \boldsymbol{Z}^\top\, b_i\, d_{y\delta}^{(i)}\, S_{Y_i}\, \delta_{ij}^n$
Covariate	$x_l + \omega S_{X_l}$	$\boldsymbol{\Delta}(\boldsymbol{\beta}) = \beta_l\, S_{X_l}\, \boldsymbol{X}^\top\, c_i \delta_{ij}^n + S_{X_l}\, \boldsymbol{1}_{n\times p}^{(k)\top}\, d_\mu^{(i)}\, a_i\, \delta_{ij}^n$ $\boldsymbol{\Delta}(\boldsymbol{\alpha}) = \beta_l\, S_{X_l}\, \boldsymbol{Z}^\top\, m_i\, \delta_{ij}^n$
	$z_k + \omega S_{Z_k}$	$\boldsymbol{\Delta}(\boldsymbol{\beta}) = \alpha_k\, S_{Z_k}\, \boldsymbol{X}^\top\, m_i\, \delta_{ij}^n$ $\boldsymbol{\Delta}(\boldsymbol{\alpha}) = \alpha_k S_{Z_k} \boldsymbol{Z}^\top w_i \delta_{ij}^n + S_{X_l} \boldsymbol{1}_{n\times q}^{(l)\top} d_\delta^{(i)} b_i \delta_{ij}^n$
Joint covariate	$\tau_i(\omega)$	$\boldsymbol{\Delta}(\boldsymbol{\beta}) = S_{X_l}[\boldsymbol{X}^\top \{\beta_l c_i \delta_{ij}^n + \alpha_k m_i \delta_{ij}^n\} + \boldsymbol{1}_{n\times p}^{(k)\top} a_i d_\mu^{(i)} \delta_{ij}^n]$ $\boldsymbol{\Delta}(\boldsymbol{\alpha}) = S_{X_l}[\boldsymbol{Z}^\top \{\alpha_k w_i \delta_{ij}^n + \beta_l m_i \delta_{ij}^n\} + \boldsymbol{1}_{n\times q}^{(l)\top} b_i d_\delta^{(i)} \delta_{ij}^n]$

where $\tau_i(\omega) = \alpha_1 + \cdots + \alpha_k[x_{il} + \omega_i S_{X_l}] + \cdots + \alpha_q z_{iq}$, $\boldsymbol{1}_{n\times p}^{(k)}$ and $\boldsymbol{1}_{n\times q}^{(l)}$ are $n \times p$ and $n \times q$ matrices, respectively, of zeros except for the kth and lth columns, which only contains ones.

4. Simulation studies

4.1. Hypothesis testing

We report MC simulations of size 5000 to compare the performance of the GR, LR, SC and WA statistics given in (2.9), (2.10), (2.11) and (2.12), respectively, by four `R` functions. We estimate the coefficients of the RBS regression model with varying precision defined in (2.3) using the `gamlss` function, contained in the `R` package of the same name. We implement the RBS distribution inside the distributions of this `R` package; see Stasinopoulos and Rigby [41]. To use the `gamlss` function, we introduce the RBS distribution defined in (2.1) in the same structure of the distributions defined in the `gamlss.dist` library. Based on the proposed methodology, we develop a set of computational routines in `R`, which form part of the `rbs` package; see Santos-Neto et al. [33]. To install this package, the code `devtools::install_github(''santosneto/RBS'')` must be used. We consider $\lambda = \max \delta_i / \min \delta_i$, for $i = 1, \ldots, n$, to measure the heterogeneity degree of the data. Note that if $\lambda = 1$, the precision is constant obtaining the model proposed by Leiva et al. [16]. The simulations are based on the link functions $\log(\mu_i) = \beta_0 + \beta_1 x_{i1}$ and $\log(\delta_i) = \alpha_0 + \alpha_1 z_{i1}$, for $i = 1, \ldots, n$. The values of the covariates x_{i1} and z_{i1} are generated from the $\mathrm{U}(0,1)$ distribution and, for fixed n, those values are kept constant throughout the MC experiment. To perform the simulations, we implement an `R` routine that calculates the rejection rates of each test (varying the values of δ) and its empirical power, considering $n \in \{30, 50, 100\}$, whereas the values of β coefficients are $\beta_0 = 2.0$ and $\beta_1 = -1.7$. First, we obtain the empirical sizes of the test and compare them with the nominal sizes $\gamma \in \{0.01, 0.05, 0.10\}$. The null hypothesis is H_0: $\alpha_1 = 0$, which is tested against a two-sided alternative. Considering $\alpha_0 \in \{1.5, 2.0, 3.0\}$, we have $\delta \in \{4.5, 7.4, 20.1\}$. The results of this study are shown in Table 3, from where it is possible to note that the GR, LR and WA tests are markedly liberal, whereas the SC test is, in general, conservative. Note that, when $\delta = 20.1$, $n = 100$ and $\gamma = 0.01$, the rejection rates are 0.96% (SC), 1.26% (GR), 1.38% (LR) and 1.76% (WA). As expected and for all the tests, these rates converge to the assumed levels as n increases. Note that the GR and SC tests present the best performances. Figure 2 displays empirical quantiles versus theoretical quantiles (QQ) plots of these statistics compared to the χ^2_{q-1} distribution. From this figure, observe that the empirical distributions of the four studied test statistics agree very well with their asymptotic distributions. Second, we compute the rejection rates under the alternative hypotheses $\alpha_1 \in \{-3.0, -2.0, -1.0\}$ against $\alpha_0 = 1.5$, which implies values of $\lambda \approx \{2.6, 6.8, 18.4\}$. Table 4 presents the empirical power of the tests for different n and λ. Note that the GR and SC tests present smaller power values in relation to the LR and WA tests. For instance, when $\lambda = 18.4$, $n = 50$ and $\gamma = 0.05$, these powers are 99.20 (GR), 99.20 (LR), 98.62 (SC) and 99.44 (WA). As expected, the powers of the four tests increase with λ. For small to moderate n, the best performance is detected for the GR and SC tests, which are more powerful than the LR and WA tests. Hence, GR and SC tests may be recommended to test hypotheses on the precision parameter in RBS regression models. The GR test has a slight advantage over the SC test, because the GR statistic is simpler to calculate.

4.2. Residuals

Now, we report the results of MC simulations of size 5000 to study the empirical distributions of the residuals r_i^{p}, r_i^{s}, r_i^{d} and r_i^{q} for the RBS regression model with

Table 3
Rejection rate of the indicated n, statistic, δ and level.

n	Statistic	$\delta = 4.5$ 1%	5%	10%	$\delta = 7.4$ 1%	5%	10%	$\delta = 20.1$ 1%	5%	10%
30	GR	1.24	6.12	11.74	1.22	6.08	11.70	1.20	6.22	11.82
	LR	1.52	6.62	12.34	1.42	6.60	12.34	1.44	6.58	12.46
	SC	0.72	4.38	9.60	0.72	4.40	9.74	0.72	4.34	9.82
	WA	2.76	8.72	15.34	2.76	8.60	15.38	2.76	8.58	15.24
50	GR	1.10	5.70	10.40	1.12	5.66	10.40	1.14	5.70	10.40
	LR	1.28	5.90	10.64	1.30	5.80	10.72	1.26	5.80	10.76
	SC	0.78	4.64	9.38	0.78	4.66	9.46	0.76	4.70	9.48
	WA	1.90	7.06	12.40	1.88	7.02	12.38	1.86	6.96	12.42
100	GR	1.26	5.38	10.62	1.28	5.40	10.58	1.26	5.38	10.62
	LR	1.36	5.54	10.80	1.36	5.54	10.72	1.38	5.64	10.72
	SC	0.92	4.96	9.62	0.94	5.02	9.76	0.96	5.02	9.78
	WA	1.80	6.34	11.62	1.84	6.36	11.78	1.76	6.30	11.60

Table 4
Power of the indicated n, statistic, λ and level.

n	Statistic	$\lambda = 2.6$ 1%	5%	10%	$\lambda = 6.8$ 1%	5%	10%	$\lambda = 18.4$ 1%	5%	10%
30	GR	7.30	20.46	31.84	37.48	62.38	74.72	75.72	91.40	95.42
	LR	7.40	20.82	32.26	38.20	63.38	75.00	76.68	91.78	95.48
	SC	5.80	19.66	30.24	30.24	57.70	70.96	61.82	86.32	92.34
	WA	7.36	20.78	31.92	36.50	63.80	75.60	77.44	92.82	96.20
50	GR	13.48	31.88	45.60	65.94	85.12	92.00	96.34	99.20	99.70
	LR	13.24	32.36	45.46	65.56	85.24	92.00	96.22	99.20	99.68
	SC	11.92	31.66	43.80	60.00	82.38	89.92	92.22	98.62	99.34
	WA	14.20	32.72	45.58	68.06	86.18	92.50	97.40	99.44	99.80
100	GR	28.12	54.80	68.50	92.82	98.20	99.16	99.94	100.00	100.00
	LR	28.38	54.86	68.58	92.90	98.18	99.18	99.94	100.00	100.00
	SC	29.72	54.14	67.00	91.58	97.56	98.98	99.80	100.00	100.00
	WA	27.36	54.78	68.80	93.20	98.44	99.30	100.00	100.00	100.00

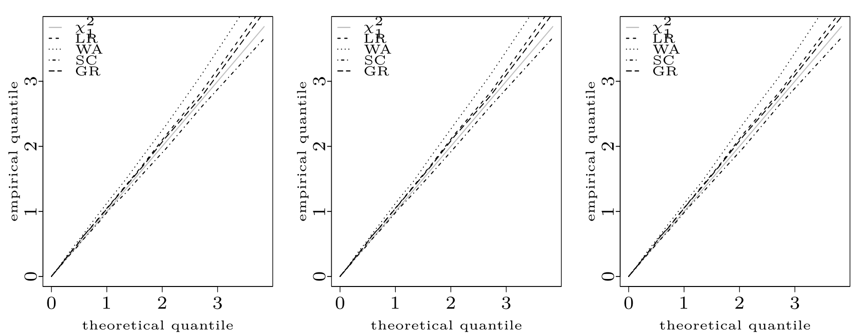

Fig 2. *QQ-plots for the indicated test statistic with $\delta = 4.5$ (left), $\delta = 7.4$ (center) and $\delta = 20.1$ (right) when $n = 50$.*

varying precision. The link functions adopted are $\log(\mu_i) = \beta_0 + \beta_1 x_i$ and $h(\delta_i) = \alpha_0 + \alpha_1 z_i$, for $i = 1, \ldots, 20$. The simulations are performed considering: [Case

I] $h(u) = \log(u)$, and [Case II] $h(u) = \sqrt{u}$, where the vectors $\boldsymbol{x} = [x_1, \dots, x_{20}]^\top$ and $\boldsymbol{z} = [z_1, \dots, z_{20}]^\top$ are obtained from the U(30, 60) distribution, with $\boldsymbol{\beta} = [2.14, 0.10]^\top$ and $\boldsymbol{\alpha} = [3.30, -0.01]^\top$ as the true values of the model coefficients. We compute the mean ($\bar{r}$), standard deviation (SD) and coefficients of skewness (CS) and kurtosis (CK), whose results are presented in Tables 5 and 6. Note that all of the residuals have mean approximately equal to zero and SD close to one. Observe also that the empirical distribution of r_i^{p} has positive asymmetry. Also, notice that r_i^{s}, r_i^{d} and r_i^{q} present a CS close to zero. The residuals r_i^{p} and r_i^{s} presents, generally, CK closer to three for both cases. Observe that r_i^{d} and r_i^{q} display in general a similar behavior. Figure 3 presents the QQ-plots with simulated envelopes of empirical quantiles versus theoretical quantiles of r_i^{p}, r_i^{s}, r_i^{d} and r_i^{q}, based on 5000 residuals. From this figure, note that r_i^{s}, r_i^{d} and r_i^{q} are more distant from the diagonal line. However, r_i^{d} presents a behavior approximately linear.

TABLE 5

Descriptive summary of the indicated residual for Case I.

	$r_i^{(\cdot)}$	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
$\bar{r}$	r_i^{p}	0.01	0.01	-0.00	-0.02	0.02	-0.00	0.00	-0.01	-0.00	-0.00	0.02	-0.00	-0.01	0.00	0.00	-0.01	0.02	0.00	-0.02	0.00
	r_i^{s}	0.01	0.03	0.02	0.00	0.00	0.00	-0.01	0.01	-0.02	0.02	0.02	-0.00	-0.02	-0.01	-0.00	-0.02	-0.00	-0.01	-0.01	0.00
	r_i^{d}	0.00	0.02	0.01	-0.01	-0.01	-0.01	-0.02	-0.00	-0.03	0.01	0.01	-0.01	-0.03	-0.02	-0.01	-0.03	-0.01	-0.02	-0.02	-0.01
	r_i^{q}	0.01	0.03	0.02	0.00	0.00	0.00	-0.01	0.01	-0.02	0.02	0.02	-0.00	-0.02	-0.01	-0.01	-0.02	-0.00	-0.01	-0.01	0.00
SD	r_i^{p}	0.98	0.92	0.91	0.93	1.05	0.96	1.05	0.93	1.06	0.91	0.97	0.98	1.00	1.04	1.02	1.01	1.06	1.03	0.96	1.01
	r_i^{s}	1.00	0.94	0.93	0.94	1.05	0.98	1.03	0.95	1.03	0.93	0.98	1.00	1.03	1.02	1.03	1.04	1.05	1.03	0.97	1.01
	r_i^{d}	0.99	0.92	0.92	0.93	1.03	0.97	1.04	0.93	1.05	0.92	0.97	0.98	1.01	1.04	1.01	1.01	1.04	1.03	0.96	0.99
	r_i^{q}	1.00	0.93	0.93	0.94	1.05	0.98	1.05	0.95	1.06	0.93	0.98	0.99	1.02	1.05	1.02	1.03	1.05	1.04	0.97	1.01
CS	r_i^{p}	0.59	0.60	0.50	0.55	0.86	0.56	0.73	0.61	0.85	0.60	0.57	0.70	0.67	0.75	0.76	0.73	0.84	0.81	0.62	0.76
	r_i^{s}	-0.03	-0.01	-0.05	-0.00	0.06	0.01	0.00	0.01	0.04	-0.01	-0.03	0.05	-0.05	0.03	0.04	-0.04	0.05	0.01	0.04	0.08
	r_i^{d}	-0.01	0.02	-0.02	0.02	0.08	0.04	0.02	0.05	0.07	0.02	0.00	0.07	-0.01	0.05	0.06	-0.01	0.06	0.04	0.06	0.10
	r_i^{q}	-0.03	-0.02	-0.05	-0.00	0.05	0.01	0.00	0.01	0.04	-0.02	-0.03	0.04	-0.04	0.03	0.04	-0.03	0.04	0.02	0.04	0.07
CK	r_i^{p}	2.90	2.87	2.72	2.82	3.75	2.73	3.33	2.80	3.70	2.83	2.84	3.06	3.13	3.31	3.30	3.36	3.75	3.62	2.92	3.26
	r_i^{s}	2.52	2.48	2.44	2.44	2.85	2.38	2.73	2.41	2.87	2.45	2.53	2.53	2.67	2.75	2.67	2.80	2.84	2.89	2.48	2.62
	r_i^{d}	2.40	2.35	2.34	2.34	2.66	2.29	2.58	2.29	2.69	2.32	2.42	2.40	2.51	2.60	2.52	2.62	2.65	2.68	2.39	2.48
	r_i^{q}	2.40	2.35	2.34	2.34	2.66	2.28	2.57	2.29	2.68	2.32	2.42	2.40	2.51	2.60	2.51	2.62	2.64	2.68	2.38	2.48

TABLE 6

Descriptive summary of the indicated residual for Case II.

	$r_i^{(\cdot)}$	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
$\bar{r}$	r_i^{p}	0.00	0.01	-0.01	-0.02	0.02	0.00	0.01	-0.01	0.00	0.00	0.01	0.00	-0.02	0.00	0.00	-0.01	0.02	0.01	-0.02	0.00
	r_i^{s}	0.01	0.04	0.03	0.01	0.00	0.01	-0.02	0.02	-0.03	0.04	0.02	0.00	-0.03	-0.02	-0.01	-0.03	-0.01	-0.02	-0.01	0.00
	r_i^{d}	-0.01	0.02	0.01	-0.01	-0.02	-0.01	-0.03	0.00	-0.04	0.01	0.00	-0.02	-0.05	-0.04	-0.03	-0.04	-0.03	-0.03	-0.03	-0.02
	r_i^{q}	0.01	0.04	0.03	0.01	-0.01	0.01	-0.02	0.02	-0.03	0.03	0.02	0.00	-0.03	-0.02	-0.01	-0.03	-0.01	-0.02	-0.01	0.00
SD	r_i^{p}	0.96	0.90	0.88	0.90	1.07	0.94	1.04	0.91	1.06	0.88	0.95	0.98	0.99	1.04	1.02	1.01	1.07	1.04	0.93	1.00
	r_i^{s}	0.97	0.89	0.89	0.90	1.06	0.94	1.06	0.91	1.07	0.88	0.96	0.98	1.02	1.05	1.02	1.03	1.07	1.05	0.94	1.00
	r_i^{d}	0.96	0.90	0.89	0.90	1.03	0.94	1.03	0.91	1.04	0.88	0.95	0.97	1.00	1.03	1.00	1.01	1.04	1.02	0.93	0.98
	r_i^{q}	0.99	0.92	0.91	0.93	1.06	0.97	1.06	0.94	1.07	0.91	0.98	1.00	1.03	1.06	1.03	1.04	1.07	1.05	0.96	1.01
CS	r_i^{p}	0.87	0.84	0.76	0.82	1.20	0.82	1.04	0.84	1.18	0.82	0.83	0.96	0.96	1.05	1.05	1.04	1.17	1.14	0.89	1.03
	r_i^{s}	-0.03	0.01	-0.04	0.01	0.08	0.03	0.00	0.03	0.05	0.01	-0.03	0.05	-0.06	0.03	0.05	-0.05	0.07	0.02	0.05	0.08
	r_i^{d}	0.02	0.06	0.01	0.05	0.11	0.07	0.05	0.09	0.09	0.07	0.02	0.09	0.01	0.08	0.09	0.02	0.09	0.07	0.09	0.12
	r_i^{q}	-0.03	-0.01	-0.05	0.00	0.06	0.02	0.01	0.02	0.05	0.00	-0.03	0.04	-0.04	0.04	0.04	-0.03	0.05	0.02	0.04	0.07
CK	r_i^{p}	3.46	3.33	3.20	3.33	4.71	3.24	4.07	3.22	4.68	3.24	3.34	3.63	3.80	3.97	3.99	4.10	4.67	4.49	3.48	3.87
	r_i^{s}	2.68	2.61	2.60	2.60	3.07	2.52	2.91	2.51	3.09	2.56	2.68	2.67	2.84	2.90	2.84	3.00	3.05	3.12	2.63	2.77
	r_i^{d}	2.42	2.36	2.38	2.38	2.66	2.31	2.58	2.30	2.68	2.32	2.45	2.40	2.51	2.59	2.52	2.62	2.65	2.67	2.41	2.48
	r_i^{q}	2.42	2.35	2.37	2.37	2.64	2.30	2.56	2.28	2.66	2.31	2.44	2.39	2.50	2.57	2.51	2.61	2.63	2.66	2.39	2.46

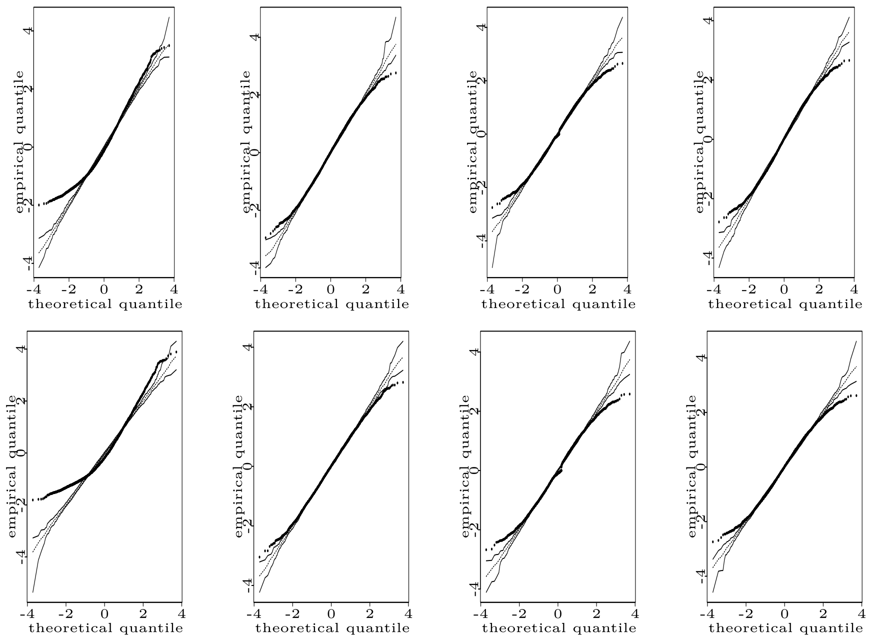

FIG 3. *QQ-plots with simulated envelopes for the residuals r_i^p, r_i^s, r_i^d and r_i^q (from left to right) in Case I (first panel) and Case II (second panel).*

5. Illustrative examples with two real-world data sets

5.1. Data set I

5.1.1. Case-study description

Alfalfa is a high protein crop and appropriate feed for dairy cows. Then, rent for land planted with alfalfa in relation to rent for other agricultural uses should be higher in zones with a high density of dairy cows. However, these rents should be lower in zones where a fertilizer is required, due to further expenses. In this line, Weisberg [48, Problem 9.10, p. 208] reported a study on the variation in rent paid for agricultural land planted with alfalfa in Minnesota. One of the objectives of that study was to investigate how the rent for land planted with alfalfa crops in relation to rent for other agricultural uses is affected by the density of dairy cows and by the proportion of farmland used as pasture, when land requires a fertilizer to increase the productivity of alfalfa or not.

5.1.2. Previous studies on these data

To evaluate the objective of the case-study described in Subsection 5.1, data were collected for 67 counties of Minnesota in 1977; see Weisberg [48, Problem 9.10, p.

208]. These data are available from an `R` package named `alr3` by the command `data(landrent)`. Landrent data were analyzed by Taylor and Verbyla [42], Lin et al. [21] and Wu et al. [49]. Non-constant variance in a linear regression model can be diagnosed by residual plots. Taylor and Verbyla [42] discussed the joint modeling of location and dispersion parameters and derived a methodology to detect heteroscedasticity, when the response is Student-t distributed. Lin et al. [21] considered tests for heteroscedasticity in Student-t linear regression models. All of these studies and also Wu et al. [49] showed evidence of heteroscedasticity in landrent data, but this last recent work showed the convenience of modeling simultaneously the mean and dispersion of the lognormal distribution, which is an asymmetrical distribution, such as the RBS distribution is.

5.1.3. Variables to be modeled

The considered variables are: (i) the ratio between the average rent per acre planted with alfalfa and the corresponding average rent for other agricultural uses ('ratio'); (ii) the density of dairy cows in number per square mile ('density'); (iii) the proportion of farmland used as pasture ('proportion'); and (iv) if the fertilizer 'liming' is required to increase the productivity of alfalfa or not. We discard the covariate 'proportion' due to it is correlated with the covariate 'density' and concentrate on the group 'liming'. Note that the response variable 'ratio' is strongly correlated with 'density' indicating a linear relation between these two variables; see Figure 4 (1st panel left). Then, our study is concentrated on the response variable 'ratio' (Y) and the covariate 'density' for the group 'liming' (X).

5.1.4. Estimation and model cheking

First, we consider the RBS model with fixed precision $Y_i \sim \text{RBS}(\mu_i, \delta)$ and link function $\mu_i = \beta_0 + \beta_1 x$, for $i = 1, \dots, 33$; see Figure 4 (1st panel left). The ML estimates of its parameters, with estimated asymptotic standard errors (SE) in parenthesis, are $\widehat{\beta}_0 = 0.67810(0.0363)$, $\widehat{\beta}_1 = 0.0166(0.0031)$ and $\widehat{\delta} = 82.3752(20.2798)$. From Figure 4 (1st panel center), observe that the assumption of the RBS distribution seems to be reasonable. From Figure 4 (1st panel right), note that the residual plot shows a pattern that indicates an evidence of a non-constant precision. Observing at Figure 4 (2nd panel left), note that the precision depends on the value x of the covariate. Then, based on this last figure, we propose the RBS model with varying precision $Y_i \sim \text{RBS}(\mu_i, \delta_i)$, where $\mu_i = \beta_0 + \beta_1 x$ and $\sqrt{\delta_i} = \alpha_0 + \alpha_1 x$, with $i = 1, \dots, 33$. For the precision parameter, we choice the square root link function based on usual selection model criteria (AIC, BIC) and the tests for varying precision proposed in Subsection 2.5. For the RBS model with varying precision, the ML estimates of its parameters and estimated asymptotic SEs in parenthesis are shown in Table 7. Model assumptions are verified using a residual analysis based on landrent data provided in Figure 4 (2nd panel center), because unusual features are not observed. In addition, the independence assumption also is verified by the QQ-plot with simulated envelope and by the plot of residuals displayed in Figure 4 (2nd panel right), from which an outlying case (#33) is detected.

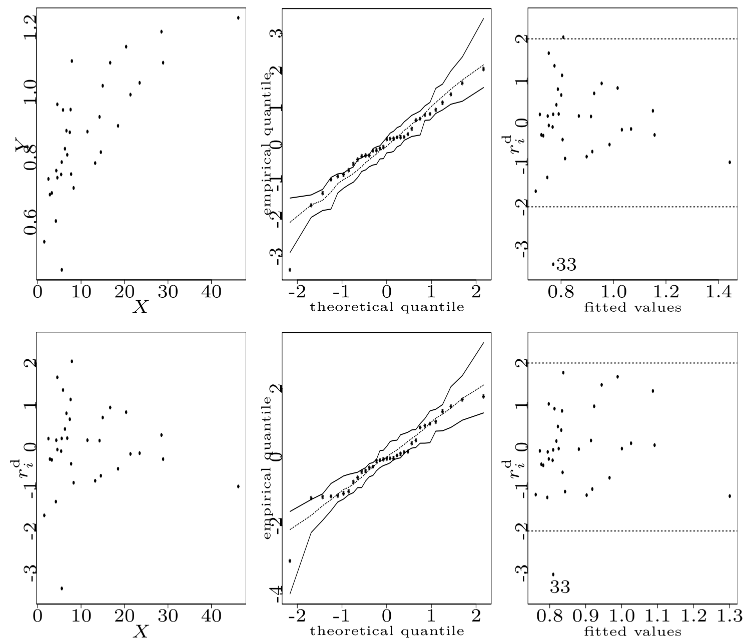

Fig 4. *Scatter-plot of X vs. Y (1st panel left), QQ-plots with simulated envelope of r_i^d for fixed precision (1st panel center) and varying precision (2nd panel center), plots of fitted values vs. r_i^d for fixed precision (1st panel right) and varying precision (2nd panel right) and X vs. r_i^d for fixed precision (1st panel right) based on landrent data and the RBS regression model.*

Table 7

ML estimates, estimated SE, t-and-p-values for the coefficients of the model with landrent data.

Coefficient	Estimate	SE	t-value	p-value
β_0	0.7426	0.0426	17.4496	< 0.001
β_1	0.0121	0.0019	6.4517	< 0.001
α_0	4.3483	2.0545	2.1165	0.043
α_1	0.5712	0.2398	2.3825	0.024

5.1.5. Influence local

Diagnostics for the RBS regression model with varying precision are in Figure 5. Specifically, this figure displays index plots of C_i, from where cases #26 and #33 are detected as potentially influential. We investigate their impact on the model inference when they are removed. Then, we again estimate the model after removing cases #26 and #33. To evaluate the impact of the removed cases, we define the relative change (RC) by $\mathrm{RC}_{\theta_{j(i)}} = |[\{\widehat{\theta}_j - \widehat{\theta}_{j(i)}\}/\widehat{\theta}_j]| \times 100\%$, where $\widehat{\theta}_{j(i)}$

denote the ML estimate of θ_j obtained after removing the case i, for $j = 1, 2, 3$ and $i = 1, \ldots, 33$. The corresponding RCs are in Table 8. Inferential changes are not detected when the cases are removed. The case #26 is a county with a higher average rent per acre planted with alfalfa, even not being the density of dairy cows the largest one. The case #33 is a county with small average rent per acre planted with alfalfa, despite possessing a density of dairy cows larger than 25% of the counties. We have that the mean of the ratio between the average rent per acre planted with alfalfa and the corresponding average rent for other agricultural uses can be described by $\widehat{\mu}(x) = 0.7426 + 0.0121\,x$. In this model, 0.7547 is the estimated mean ratio when the dairy cow density per square mile is zero, whereas the estimated mean ratio increases in 0.0121 if we increase in one unit the dairy cow density per square mile.

TABLE 8
RC for the indicated removed case(s) with landrent data.

Removed subset	$\widehat{\beta_0}$	$\widehat{\beta_1}$	$\widehat{\alpha_0}$	$\widehat{\alpha_1}$
{26}	6.4069	30.8475	1.7622	8.9307
{33}	0.4237	0.2543	54.5430	12.6125
{26,33}	4.4108	25.9012	58.0622	5.0616

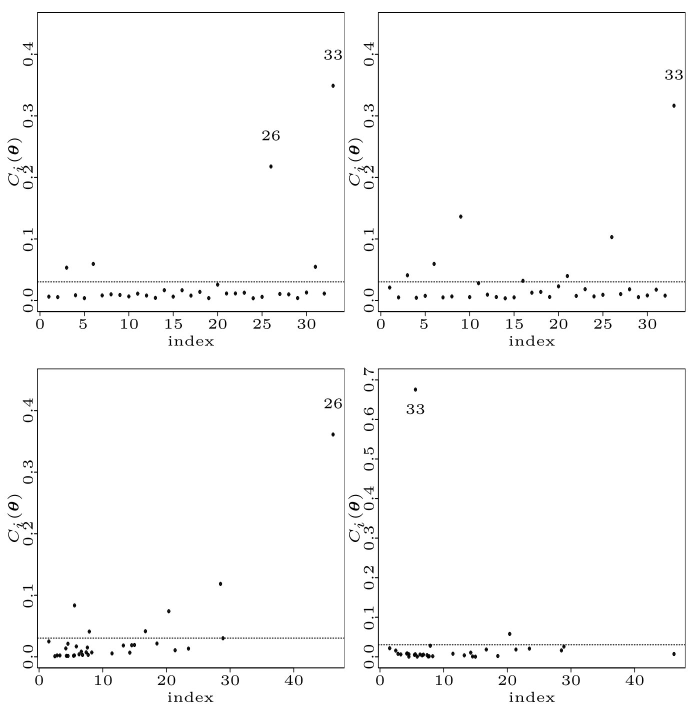

FIG 5. *Index plots of C_i for $\boldsymbol{\theta}$ under case-weight (a), response (b), covariate-X (c) and covariate-Z (d) perturbations in the RBS regression model with varying precision based on landrent data.*

5.2. Data set II

5.2.1. Case-study description

Four different light snacks (named B, C, D and E) were compared across 20 weeks with a traditional snack (named A) in an experiment developed in the School of Public Health of the University of São Paulo, Brazil. For the light snacks, the hydrogenated vegetable fat (HVF) was replaced by canola oil under different proportions: B (0% HVF, 22% canola oil), C (17% HVF, 5% canola oil), D (11% HVF, 11% canola oil), E (5% HVF, 17% canola oil), and A (22% HVF, 0% canola oil). A random sample of 15 units of each snack type was analyzed in a laboratory and various variables were measured. Then, a total of 75 units were considered during ten weeks, making 750 units in total during the experiment; see Paula [26]. One of the objectives of that study was to determine the ideal storing time.

5.2.2. Previous studies on these data

These data are available from an `R` package named `ssym` by the command `data(Snacks)`. Snack data were analyzed by Paula [26], who discussed diagnostic methods in double generalized linear models. The author modeled simultaneously the mean and precision of the gamma distribution, which is an asymmetrical distribution, such as the RBS distribution is.

5.2.3. Variables to be modeled

The considered variables are: texture (Y), snack type (S) and quantity of weeks (W). The variable texture is compared across time among the five snack types. Paula [26] observed that the means and CVs seem to be different among the snacks, changing across weeks, with indication of quadratic tendencies for the means.

5.2.4. Estimation and model cheking

Based on the descriptive analysis carried out by Paula [26], we consider the RBS regression model with varying precision. Here, Y_{ijk} denotes the texture corresponding to the unit k of the snack type i in the jth week, with $i = 1(A), 2(B), 3(C), 4(D), 5(E)$, $j = 2, 4, \ldots, 18, 20$, and $k = 1, 2, \ldots, 15$. We fitted 54 different models (named Models $1, \ldots, 54$) with identity, logarithm and square root link functions for g and h. Table 9 presents the AIC and BIC values for each fitted model. The tests for varying precision are significative for all these models. Note that three models present the smallest AIC and BIC values (Models 22, 28 and 34). We select Model 28, because it considers a link function different from the used in the first application. For the RBS regression model with varying precision, the ML estimates of its parameters and estimated asymptotic SEs in parenthesis are shown in Table 10. Model assumptions are verified using residuals in Figure 6 (right), because unusual features are not detected. In addition, the independence assumption also is verified by the QQ-plot with simulated envelope and by the residual plot displayed in Figure 6, from which the case #91 (type A snack) is detected as atypical.

TABLE 9
Summary of the RBS models with varying precision fitted to snack data.

Model	g	Predictor	h	Predictor	AIC	BIC
1	Identity	$S+W$	Identity	S	5967.835	6018.655
2				$S+W$	5963.295	6018.736
3				$S+W+W^2$	5964.303	6024.363
4		$S+W+W^2$		S	5930.878	5986.319
5				$S+W$	5932.753	5992.814
6				$S+W+W^2$	5934.666	5999.347
7	Identity	$S+W$	Logarithm	S	5967.835	6018.655
8				$S+W$	5961.163	6016.604
9				$S+W+W^2$	5962.058	6022.119
10		$S+W+W^2$		S	5930.878	5986.319
11				$S+W$	5932.833	5992.894
12				$S+W+S^2$	5934.689	5999.370
13	Identity	$S+W$	Square root	S	5967.835	6018.655
14				$S+W$	5962.112	6017.553
15				$S+W+W^2$	5962.950	6023.011
16		$S+W+W^2$		S	5930.878	5986.319
17				$S+W$	5932.801	5992.862
18				$S+W+W^2$	5934.696	5999.377
19	Logarithm	$S+W$	Identity	S	5977.846	6028.666
20				$S+W$	5972.214	6027.655
21				$S+W+W^2$	5972.678	6032.739
22		$S+W+W^2$		S	5923.985	5979.425
23				$S+W$	5925.954	5986.014
24				$S+W+W^2$	5927.877	5992.558
25	Logarithm	$S+W$	Logarithm	S	5977.846	6028.666
26				$S+W$	5971.337	6026.777
27				$S+W+W^2$	5972.023	6032.084
28		$S+W+W^2$		S	5923.985	5979.425
29				$S+W$	5925.984	5986.045
30				$S+W+W^2$	5927.823	5992.504
31	Logarithm	$S+W$	Square root	S	5977.846	6028.666
32				$S+W$	5971.567	6027.008
33				$S+W+W^2$	5971.987	6032.048
34		$S+W+W^2$		S	5923.985	5979.425
35				$S+W$	5925.980	5986.041
36				$S+W+W^2$	5927.873	5992.554
37	Square root	$S+W$	Identity	S	5972.507	6023.328
38				$S+W$	5967.308	6022.749
39				$S+W+W^2$	5968.048	6028.109
40		$S+W+W^2$		S	5927.194	5982.635
41				$S+W$	5929.116	5989.177
42				$S+W+W^2$	5931.038	5995.719
43	Square root	$S+W$	Logarithm	S	5972.507	6023.328
44				$S+W$	5965.698	6021.139
45				$S+W+W^2$	5966.457	6026.518
46		$S+W+W^2$		S	5927.194	5982.635
47				$S+W$	5929.184	5989.245
48				$S+W+W^2$	5931.036	5995.717
49	Square root	$S+W$	Square root	S	5972.507	6023.328
50				$S+W$	5966.337	6021.778
51				$S+W+W^2$	5966.950	6027.011
52		$S+W+W^2$		S	5927.194	5982.635
53				$S+W$	5929.160	5989.221
54				$S+W+W^2$	5931.059	5995.740

Table 10

ML estimates, estimated SE, t-and-p-values for the coefficients of the model with snack data.

Coefficient	Estimate	SE	t-value	p-value
(Intercept)	3.842	0.037	103.734	< 0.001
Type B	-0.182	0.031	-5.921	< 0.001
Type C	-0.078	0.034	-2.295	0.022
Type D	-0.265	0.029	-9.184	< 0.001
Type E	-0.282	0.029	-9.839	< 0.001
Week	0.061	0.006	10.300	< 0.001
Week^2	-0.002	0.000	-8.105	< 0.001
(Intercept)	3.104	0.116	26.728	< 0.001
Type B	0.585	0.164	3.575	< 0.001
Type C	0.123	0.164	0.750	0.453
Type D	1.008	0.164	6.154	< 0.001
Type E	1.043	0.166	6.297	< 0.001

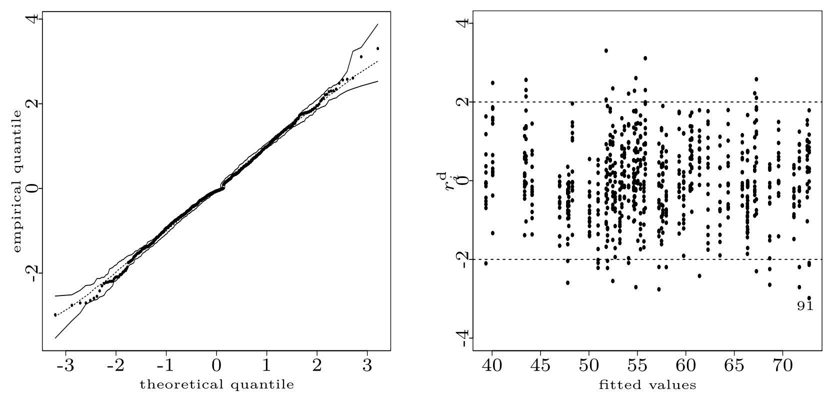

Fig 6. *QQ-plot with simulated envelope for r_i^d (left) and fitted values vs. r_i^d (right) based on snack data and RBS regression model with varying precision.*

5.2.5. Influence local

Diagnostics for the RBS regression model with varying precision for snack data are presented in Figure 7. This Figure displays C_i versus weeks for $\boldsymbol{\theta}$ under case-weight, response and covariate-X perturbation, from where cases #76 (type A), #91 (type A), #136 (type A), #346 (type C), #465 (type D) and #750 (type E) are detected as potentially influential. We investigate the impact on the model inference when they are removed, but their elimination does not change the inference for the mean nor for the precision. Figure 8 presents the predicted mean and precision values for the texture across weeks, for each snack type. From Figure 8 (left), note that the types A and C snacks have the largest mean values across weeks. From Figure 8 (right), note that the types D and E snacks have the largest precision for the texture. An objective of the case-study is to determine the ideal storing time, and looking at Figure 8 (left), it seems to be about 14 weeks.

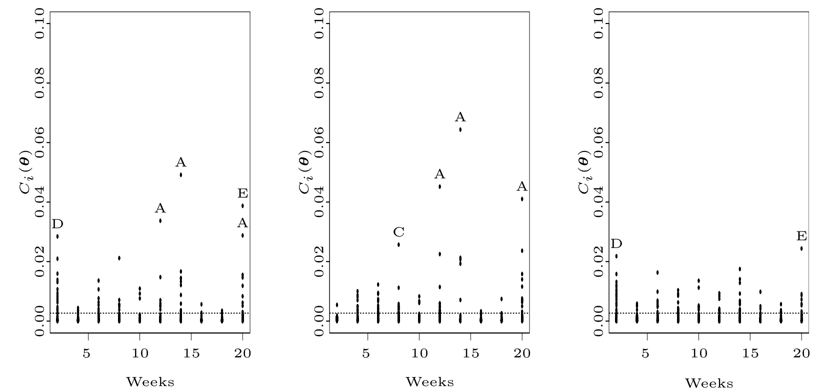

FIG 7. *Plots of C_i vs. weeks for $\boldsymbol{\theta}$ under case-weight (left), response (center), covariate-X (right) perturbations in the RBS regression model with varying precision based on snack data.*

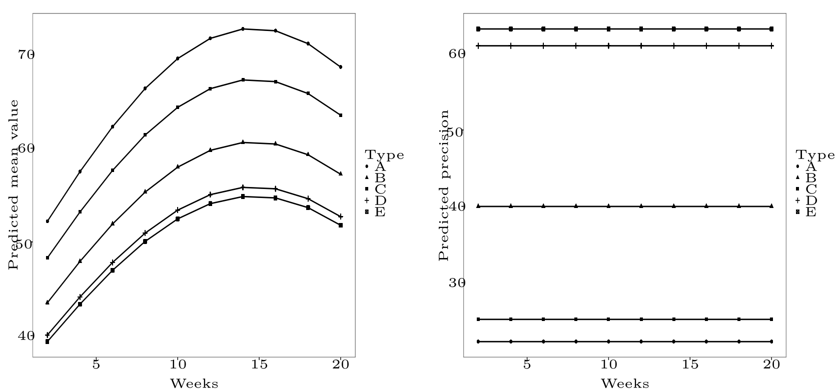

FIG 8. *Predicted mean value of texture (left) and predicted precision of texture (right) for each snack type across weeks.*

6. Concluding remarks

In this paper, we have developed a methodology based on a reparameterized Birnbaum-Saunders regression model with varying precision, generalizing the existing works in the literature on the topic. We have dealt with the issues of estimating the model parameters and of performing hypothesis testing on the corresponding precision parameter. We have considered three classic statistics, such as likelihood ratio, score and Wald, and one recently proposed called gradient, to test varying precision in our model. In addition, we have conducted Monte Carlo simulations to study the finite-sample performance of these tests. The simulation results have indicated that the gradient and score tests are preferred, but the former one has an advantage because it is simpler to calculate. Also, we have presented fours types of residuals when the mean and precision are modeled simultaneously. Again using Monte Carlo simulations, we have evaluated their performances and determined their distributions empirically, which have resulted to follow a normal distribution reasonably. Furthermore, we have developed diagnostic tools for this new class of regression models based on generalized leverage and local influence methods, considering case-weight, covariate and response perturbations. These tools have shown to be useful for detecting influential cases. Moreover, we have carried out applications with two real-world case-studies, which have shown the potential of using the new methodology based on a reparameterized Birnbaum-Saunders regression model with varying precision.

Appendix A: Inference

A.1. Score vector

The elements of the score vector, obtained by differentiation of the log-likelihood function expressed in (2.4) with respect to $\boldsymbol{\theta}$, are given, for $j = 1, \ldots, p$ and $r = 1, \ldots, q$, as

$$\dot{\ell}(\beta_j) = \tfrac{\partial \ell(\boldsymbol{\theta})}{\partial \beta_j} = \sum_{i=1}^{n} [y_i^\star - \mu_i^\star]\, a_i\, x_{ij}, \quad \dot{\ell}(\alpha_r) = \tfrac{\partial \ell(\boldsymbol{\theta})}{\partial \alpha_r} = \sum_{i=1}^{n} [y_i^\star - \delta_i^\star]\, b_i\, z_{ir},$$

where

$$y_i^\star = \tfrac{\delta_i}{\delta_i y_i + y_i + \delta_i \mu_i} + \tfrac{y_i[\delta_i+1]}{4{\mu_i}^2} - \tfrac{\delta_i^2}{4y_i[\delta_i+1]}, \mu_i^\star = \tfrac{1}{2\mu_i}, a_i = \tfrac{\mathrm{d}\mu_i}{\mathrm{d}\eta_i} = \tfrac{1}{\mathrm{d}g(\mu_i)/\mathrm{d}\mu_i},$$

$$y_i^\star = \tfrac{y_i + \mu_i}{\delta_i y_i + y_i + \delta_i \mu_i} - \tfrac{y_i}{4\mu_i} - \tfrac{\delta_i[\delta_i+2]\mu_i}{4[\delta_i+1]^2 y_i}, \delta_i^\star = -\tfrac{\delta_i}{2[\delta_i+1]}, b_i = \tfrac{\mathrm{d}\delta_i}{\mathrm{d}\tau_i} = \tfrac{1}{\mathrm{d}h(\delta_i)/\mathrm{d}\delta_i}.$$

Therefore, we can write the $[p+q]\times 1$ score vector $\dot{\boldsymbol{\ell}}(\boldsymbol{\theta})$ in the form $[\dot{\boldsymbol{\ell}}(\boldsymbol{\beta})^\top \dot{\boldsymbol{\ell}}(\boldsymbol{\alpha})^\top]^\top$, with $\dot{\boldsymbol{\ell}}(\boldsymbol{\beta}) = \boldsymbol{X}^\top \boldsymbol{A}[\boldsymbol{y}^\star - \boldsymbol{\mu}^\star]$ and $\dot{\boldsymbol{\ell}}(\boldsymbol{\alpha}) = \boldsymbol{Z}^\top \boldsymbol{B}[\boldsymbol{y}^\star - \boldsymbol{\delta}^\star]$, where $\boldsymbol{A} = [a_i \delta_{ij}^n]$ and $\boldsymbol{B} = [b_i \delta_{ij}^n]$, with δ_{ij}^n being the Kronecker delta. Thus, $\boldsymbol{A}$ and $\boldsymbol{B}$ are two $n \times n$ diagonal matrices.

A.2. Hessian matrix

The corresponding Hessian matrix is given by

$$\ddot{\boldsymbol{\ell}}(\boldsymbol{\theta}) = \begin{bmatrix} \ddot{\boldsymbol{\ell}}(\boldsymbol{\beta}) & \ddot{\boldsymbol{\ell}}(\boldsymbol{\beta\alpha}) \\ \ddot{\boldsymbol{\ell}}(\boldsymbol{\alpha\beta}) & \ddot{\boldsymbol{\ell}}(\boldsymbol{\alpha}) \end{bmatrix} = \begin{bmatrix} \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}^\top} & \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\beta} \partial \boldsymbol{\alpha}^\top} \\ \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\alpha} \partial \boldsymbol{\beta}^\top} & \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\alpha} \partial \boldsymbol{\alpha}^\top} \end{bmatrix} = \begin{bmatrix} \boldsymbol{X}^\top \boldsymbol{C} \boldsymbol{X} & \boldsymbol{X}^\top \boldsymbol{M} \boldsymbol{Z} \\ \boldsymbol{Z}^\top \boldsymbol{M} \boldsymbol{X} & \boldsymbol{Z}^\top \boldsymbol{W} \boldsymbol{Z} \end{bmatrix}. \tag{A.1}$$

The elements of the first block of the matrix $\ddot{\boldsymbol{\ell}}(\boldsymbol{\theta})$, $\ddot{\boldsymbol{\ell}}(\boldsymbol{\beta})$ say, are obtained from the derivative $\partial^2 \ell(\boldsymbol{\theta})/\partial \beta_j \partial \beta_k = \sum_{i=1}^n c_i\, x_{ij}\, x_{ik}$, which may be written in matrix form as $\ddot{\boldsymbol{\ell}}(\boldsymbol{\beta}) = \boldsymbol{X}^\top \boldsymbol{C} \boldsymbol{X}$, where $\boldsymbol{C} = c_i\, \delta_{ij}^n$ and $c_i = d_{\mu^2}^{(i)}\, (a_i)^2 + d_{\mu}^{(i)}\, a_i^{'}\, a_i$, with

$$d_\mu^{(i)} \quad = \quad \frac{\partial \ell(\theta)}{\partial \mu_i} = \overbrace{\tfrac{\delta_i}{\delta_i y_i + y_i + \delta_i \mu_i} + \tfrac{y_i[\delta_i+1]}{4{\mu_i}^2} - \tfrac{\delta_i^2}{4y_i[\delta_i+1]}}^{y_i^\star} - \overbrace{\tfrac{1}{2\mu_i}}^{\mu_i^\star} = y_i^\star - \mu_i^\star,$$

$$d^{(i)}_{\mu^2} = \frac{\partial^2 \ell(\theta)}{\partial \mu_i^2} = \frac{1}{2\mu_i^2} - \frac{\delta_i^2}{[\delta_i y_i + y_i + \delta_i \mu_i]^2} - \frac{y_i[\delta_i+1]}{2{\mu_i}^3}, \quad a_i^{'} = -\frac{\mathrm{d}^2 g(\mu_i)/\mathrm{d}\mu_i^2}{[\mathrm{d}g(\mu_i)/\mathrm{d}\mu_i]^2}.$$

The elements of the second block of the matrix $\ddot{\boldsymbol{\ell}}(\boldsymbol{\theta})$, $\ddot{\boldsymbol{\ell}}(\boldsymbol{\beta\alpha})$ say, are obtained from the derivative $\partial \ell(\boldsymbol{\theta})/\partial \beta_j \partial \alpha_r = \sum_{i=1}^n m_i\, x_{ij}\, z_{ir}$, where $m_i = d^{(i)}_{\mu\delta}\, a_i\, b_i$, with

$$d^{(i)}_{\mu\delta} = \frac{\partial^2 \ell(\theta)}{\partial \mu_i \partial \delta_i} = \frac{y_i}{[\delta_i y_i + y_i + \delta_i \mu_i]^2} + \frac{y_i}{4\mu_i^2} - \frac{\delta_i[\delta_i+2]}{4[\delta_i+1]^2 y_i}.$$

In matrix form, we have $\ddot{\boldsymbol{\ell}}(\boldsymbol{\beta\alpha}) = \boldsymbol{X}^\top \boldsymbol{M} \boldsymbol{Z}$, with $\boldsymbol{M} = m_i\, \delta^n_{ij}$. Consequently, we obtain the third block of the matrix $\ddot{\boldsymbol{\ell}}(\boldsymbol{\theta})$, $\ddot{\boldsymbol{\ell}}(\boldsymbol{\alpha\beta})$ say, since this is the transpose of the matrix $\ddot{\boldsymbol{\ell}}(\boldsymbol{\beta\alpha})$. The elements of the last block of $\ddot{\boldsymbol{\ell}}(\boldsymbol{\theta})$, $\ddot{\boldsymbol{\ell}}(\boldsymbol{\alpha})$ say, are obtained from the derivative $\partial^2 \ell(\boldsymbol{\theta})/\partial \alpha_l \partial \alpha_s = \sum_{i=1}^n w_i z_{il} z_{is}$, where $w_i = d^{(i)}_{\delta^2} (b_i)^2 + d^{(i)}_{\delta}\, b_i^{'}\, b_i$, with

$$\begin{aligned}
d^{(i)}_{\delta} &= \frac{\partial \ell(\theta)}{\partial \delta_i} = \overbrace{\frac{y_i+\mu_i}{\delta_i y_i + y_i + \delta_i \mu_i} - \frac{y_i}{4\mu_i} - \frac{\delta_i[\delta_i+2]\mu_i}{4[\delta_i+1]^2 y_i}}^{y_i^\star} + \overbrace{\frac{\delta_i}{2[\delta_i+1]}}^{-\delta_i^\star} = y_i^\star - \delta_i^\star,\\
d^{(i)}_{\delta^2} &= \frac{\partial^2 \ell(\theta)}{\partial \delta_i^2} = \frac{1}{2[\delta_i+1]^2} - \frac{[y_i+\mu_i]^2}{[\delta_i y_i + y_i + \delta_i \mu_i]^2} - \frac{\mu_i}{2[\delta_i+1]^3 y_i}, \quad b_i^{'} = -\frac{\mathrm{d}^2 h(\delta_i)/\mathrm{d}\delta_i^2}{[\mathrm{d}h(\delta_i)/\mathrm{d}\delta_i]^2},
\end{aligned}$$

whereas its matrix form can be expressed as $\ddot{\boldsymbol{\ell}}(\boldsymbol{\alpha}) = \boldsymbol{Z}^\top \boldsymbol{W} \boldsymbol{Z}$, with $\boldsymbol{W} = w_i\, \delta^n_{ij}$. The inverse Hessian matrix is given by

$$\ddot{\boldsymbol{\ell}}(\boldsymbol{\theta})^{-1} = \left[\begin{array}{cc} \ddot{\boldsymbol{\ell}}(\boldsymbol{\beta})^{-1} & \ddot{\boldsymbol{\ell}}(\boldsymbol{\beta\alpha})^{-1} \\ \ddot{\boldsymbol{\ell}}(\boldsymbol{\alpha\beta})^{-1} & \ddot{\boldsymbol{\ell}}(\boldsymbol{\alpha})^{-1} \end{array}\right],$$

where $\ddot{\boldsymbol{\ell}}(\boldsymbol{\beta})^{-1} = [\boldsymbol{X}^\top \boldsymbol{W}_1 \boldsymbol{X}]^{-1}$, $\ddot{\boldsymbol{\ell}}(\boldsymbol{\beta\alpha})^{-1} = -[\boldsymbol{X}^\top \boldsymbol{W}_1 \boldsymbol{X}]^{-1} \boldsymbol{X}^\top \boldsymbol{M} \boldsymbol{Z} [\boldsymbol{Z}^\top \boldsymbol{W} \boldsymbol{Z}]^{-1}$, $\ddot{\boldsymbol{\ell}}(\boldsymbol{\alpha\beta})^{-1} = -[[\boldsymbol{X}^\top \boldsymbol{W}_1 \boldsymbol{X}]^{-1} \boldsymbol{X}^\top \boldsymbol{M} \boldsymbol{Z} [\boldsymbol{Z}^\top \boldsymbol{W} \boldsymbol{Z}]^{-1}]^\top$, and $\ddot{\boldsymbol{\ell}}(\boldsymbol{\alpha})^{-1} = [\boldsymbol{Z}^\top \boldsymbol{W}_2 \boldsymbol{Z}]^{-1}$, with $\boldsymbol{W}_1 = \boldsymbol{C} - \boldsymbol{M} \boldsymbol{Z} [\boldsymbol{Z}^\top \boldsymbol{W} \boldsymbol{Z}]^{-1} \boldsymbol{Z}^\top \boldsymbol{M}$ and $\boldsymbol{W}_2 = \boldsymbol{W} - \boldsymbol{M} \boldsymbol{X} [\boldsymbol{X}^\top \boldsymbol{C} \boldsymbol{X}]^{-1} \boldsymbol{X}^\top \boldsymbol{M}$. The corresponding Fisher information matrix is given by

$$\boldsymbol{i}(\boldsymbol{\theta}) = \begin{bmatrix} \boldsymbol{i}(\boldsymbol{\beta}) & \boldsymbol{i}(\boldsymbol{\beta\alpha}) \\ \boldsymbol{i}(\boldsymbol{\alpha\beta}) & \boldsymbol{i}(\boldsymbol{\alpha}) \end{bmatrix} = \left[\begin{array}{cc} \boldsymbol{X}^\top \boldsymbol{V} \boldsymbol{X} & \boldsymbol{X}^\top \boldsymbol{S} \boldsymbol{Z} \\ \boldsymbol{Z}^\top \boldsymbol{S} \boldsymbol{X} & \boldsymbol{Z}^\top \boldsymbol{U} \boldsymbol{Z} \end{array}\right], \tag{A.2}$$

where $\boldsymbol{V} = \mathrm{diag}\{\mathcal{E}_{c_1}, \ldots, \mathcal{E}_{c_n}\}$, $\boldsymbol{S} = \mathrm{diag}\{\mathcal{E}_{m_1}, \ldots, \mathcal{E}_{m_n}\}$, $\boldsymbol{U} = \mathrm{diag}\{\mathcal{E}_{w_1}, \ldots, \mathcal{E}_{w_n}\}$, with

$$\mathcal{E}_{c_i} = \left[\frac{\delta_i}{2\mu_i^2} + \frac{\delta_i^2}{\{\delta_i+1\}^2}\, I(\boldsymbol{\theta})\right] a_i^2, \quad \mathcal{E}_{m_i} = \left[\frac{1}{2\mu_i\{\delta_i+1\}} + \frac{\delta_i \mu_i}{\{\delta_i+1\}^3}\, I(\boldsymbol{\theta})\right] a_i\, b_i,$$

$$\mathcal{E}_{w_i} = \left[\frac{\delta_i^2 + 3\delta_i + 1}{2\delta_i^2\{\delta_i+1\}^2} + \frac{\mu_i^2}{\{\delta_i+1\}^4}\, I(\boldsymbol{\theta})\right] b_i^2,$$

and

$$I(\boldsymbol{\theta}) = \int_0^\infty \frac{\sqrt{\delta_i+1}\exp(\delta_i/2)}{4\sqrt{\pi \mu_i} y_i^{3/2}} \left[y_i + \frac{\delta_i \mu_i}{\delta_i+1}\right]^{-1} \exp\left(-\frac{\delta_i}{4}\left[\frac{\{\delta_i+1\} y_i}{\delta_i \mu_i} + \frac{\delta_i \mu_i}{\{\delta_i+1\} y_i}\right]\right) \mathrm{d}y_i. \tag{A.3}$$

Integral defined in (A.3) can be calculated numerically using the `integrate` function of the `R` software. The corresponding inverse expected Fisher information matrix is given by

$$\boldsymbol{i}^{-1}(\boldsymbol{\theta}) = \begin{bmatrix} [\boldsymbol{X}^\top \boldsymbol{W}_3 \boldsymbol{X}]^{-1} & -[\boldsymbol{X}^\top \boldsymbol{W}_3 \boldsymbol{X}]^{-1}\boldsymbol{X}^\top \boldsymbol{S}\boldsymbol{Z}[\boldsymbol{Z}^\top \boldsymbol{U}\boldsymbol{Z}]^{-1} \\ -[[\boldsymbol{X}^\top \boldsymbol{W}_3 \boldsymbol{X}]^{-1}\boldsymbol{X}^\top \boldsymbol{S}\boldsymbol{Z}[\boldsymbol{Z}^\top \boldsymbol{U}\boldsymbol{Z}]^{-1}]^\top & [\boldsymbol{Z}^\top \boldsymbol{W}_4 \boldsymbol{Z}]^{-1} \end{bmatrix},$$

where $\boldsymbol{W}_3 = \boldsymbol{V} - \boldsymbol{S}\boldsymbol{Z}[\boldsymbol{Z}^\top \boldsymbol{U}\boldsymbol{Z}]^{-1}\boldsymbol{Z}^\top \boldsymbol{S}$ and $\boldsymbol{W}_4 = \boldsymbol{U} - \boldsymbol{S}\boldsymbol{X}[\boldsymbol{X}^\top \boldsymbol{V}\boldsymbol{X}]^{-1}\boldsymbol{X}^\top \boldsymbol{S}$.

A.3. Score residual

Considering a known regression structure for the precision parameter, we obtain the Fisher scoring iterative procedure for $\boldsymbol{\beta}$ by

$$\boldsymbol{\beta}^{(m+1)} = \boldsymbol{\beta}^{(m)} + [\boldsymbol{X}^\top \boldsymbol{V}^{(m)} \boldsymbol{X}]^{-1}\boldsymbol{X}^\top \boldsymbol{A}^{(m)}[\boldsymbol{y}^\star - \boldsymbol{\mu}^{\star(m)}],$$

rewritten as a weighted LS estimate by $\boldsymbol{\beta}^{(m+1)} = [\boldsymbol{X}^\top \boldsymbol{V}^{(m)} \boldsymbol{X}]^{-1}\boldsymbol{X}^\top \boldsymbol{V}^{(m)}\boldsymbol{z}_1^{\star(m)}$, where $\boldsymbol{z}_1^{\star(m)} = \boldsymbol{\eta}^{(m)} + [\boldsymbol{V}^{(m)}]^{-1}\boldsymbol{A}^{(m)}[\boldsymbol{y}^\star - \boldsymbol{\mu}^{\star(m)}]$. With the convergence of the iterative procedure, we have $\widehat{\boldsymbol{\beta}} = [\boldsymbol{X}^\top \widehat{\boldsymbol{V}}\boldsymbol{X}]^{-1}\boldsymbol{X}^\top \widehat{\boldsymbol{V}}\widehat{\boldsymbol{z}}_1^\star$, for $\widehat{\boldsymbol{z}}_1^\star = \widehat{\boldsymbol{\eta}} + \widehat{\boldsymbol{V}}^{-1}\widehat{\boldsymbol{A}}[\boldsymbol{y}^\star - \widehat{\boldsymbol{\mu}}^\star]$, which can be interpreted as the LS solution of the linear regression of $\widehat{\boldsymbol{z}}_1^\star$ on $\boldsymbol{X}$ with weight matrix $\widehat{\boldsymbol{V}}$. By using this result, we propose a residual based on the solution of weighted LS of $\widehat{\boldsymbol{z}}_1^\star$ against $\boldsymbol{X}$, which we call score residual and define as $\boldsymbol{r} = \widehat{\boldsymbol{V}}^{1/2}[\widehat{\boldsymbol{z}}_1^\star - \widehat{\boldsymbol{\eta}}] = \widehat{\boldsymbol{V}}^{-1/2}\widehat{\boldsymbol{A}}[\boldsymbol{y}^\star - \widehat{\boldsymbol{\mu}}^\star]$.

Appendix B: Perturbation matrices

B.1. Case-weight perturbation

Under this scheme, which is the most used to evaluate LI in a model, we wish to determine if the contributions of the cases with different weights impact the ML estimate of $\boldsymbol{\theta}$. Here, the perturbed log-likelihood function is $\ell(\boldsymbol{\theta}|\boldsymbol{\omega}) = \sum_{i=1}^n \ell_i(\boldsymbol{\theta}|\omega_i) = \sum_{i=1}^n \omega_i \ell_i(\boldsymbol{\theta})$. Then, considering its derivative with respect to $\boldsymbol{\omega}^\top$, we obtain $\boldsymbol{\Delta}$ as in (3.4), where the elements of $\boldsymbol{\Delta}(\boldsymbol{\beta})$ and $\boldsymbol{\Delta}(\boldsymbol{\alpha})$ are respectively given by $\boldsymbol{\Delta}(\boldsymbol{\beta})_{ji} = d_\mu^{(i)}\, a_i\, x_{ij}$ and $\boldsymbol{\Delta}(\boldsymbol{\alpha})_{ri} = d_\delta^{(i)}\, b_i\, z_{ir}$, for $i = 1, \ldots, n$, $j = 1, \ldots, p$ and $r = 1, \ldots, q$, which must be evaluated at $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}$. In matrix form, we have to $\boldsymbol{\Delta}(\boldsymbol{\beta}) = \boldsymbol{X}^\top\, a_i\, d_\mu^{(i)} \delta_{ij}^n$ and $\boldsymbol{\Delta}(\boldsymbol{\alpha}) = \boldsymbol{Z}^\top\, b_i\, d_\delta^{(i)} \delta_{ij}^n$.

B.2. Response perturbation

We additively perturb the response in the RBS model as $Y_{i\omega} = Y_i + \omega_i S_{Y_i}$, where S_{Y_i} can be the SD of Y. Here, the perturbed log-likelihood function is $\ell(\boldsymbol{\theta}|\boldsymbol{\omega}) = \sum_{i=1}^n \ell_i(\boldsymbol{\theta}|\omega_i)$, with $\boldsymbol{\omega}_0 = \boldsymbol{0}_{n\times 1}$. Then, we obtain $\boldsymbol{\Delta}$ as in (3.4), where $\boldsymbol{\Delta}(\boldsymbol{\beta})_{ji} = d_{y\mu}^{(i)}\, a_i\, x_{ij}\, S_{Y_i}$ and $\boldsymbol{\Delta}(\boldsymbol{\alpha})_{ri} = d_{y\delta}^{(i)}\, b_i\, z_{ir}\, S_{Y_i}$, for $i = 1, \ldots, n$, $j = 1, \ldots, p$ and $r = 1, \ldots, q$, which must be evaluated at $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}$. In matrix form, $\mathbf{\Delta}(\boldsymbol{\beta}) = \boldsymbol{X}^\top a_i\, d_{y\mu}^{(i)}\, S_{Y_i}\, \delta_{ij}^n$ and $\mathbf{\Delta}(\boldsymbol{\alpha}) = \boldsymbol{Z}^\top b_i\, d_{y\delta}^{(i)}\, S_{Y_i}\, \delta_{ij}^n$. Relating GL and LI under this scheme, which allows us to compare them, with $\widehat{\mathbf{\Sigma}}_{\boldsymbol{y}} = \widehat{S}_{y_i}\, \delta_{ij}^n$, we have that (3.3) can be expressed as

$$\begin{aligned}\boldsymbol{F}(\boldsymbol{\theta}) = \widehat{\mathbf{\Sigma}}_{\boldsymbol{y}}[\widehat{\boldsymbol{L}}\widehat{\boldsymbol{B}}\boldsymbol{Z}[[\boldsymbol{Z}^\top\widehat{\boldsymbol{W}}_2\boldsymbol{Z}]^{-1}\boldsymbol{Z}^\top\widehat{\boldsymbol{B}}\widehat{\boldsymbol{L}} \\ - [\boldsymbol{Z}^\top\widehat{\boldsymbol{W}}\boldsymbol{Z}]^{-1}\boldsymbol{Z}^\top\widehat{\boldsymbol{M}}\widehat{\boldsymbol{A}}^{-1}\widehat{\mathrm{GL}}(\boldsymbol{\beta})] - \widehat{\boldsymbol{E}}\,\widehat{\mathrm{GL}}(\boldsymbol{\theta})]\widehat{\mathbf{\Sigma}}_{\boldsymbol{y}}.\end{aligned}$$

B.3. Covariate perturbation

Here, we additively perturb a continuous covariate-X replacing x_l by $x_l + \omega S_{X_l}$, where S_{X_l} is the SD of X_l. Here, the linear predictor component i is $\eta_i(\omega) = \beta_1 + \cdots + \beta_l[x_{il} + \omega_i S_{X_l}] + \cdots + \beta_p x_{ip}$ and $\boldsymbol{\omega}_0 = \mathbf{0}_{n\times 1}$, so that now $\ell(\boldsymbol{\theta}|\boldsymbol{\omega}) = \sum_{i=1}^n \ell_i(\mu_i(\omega_i), \delta_i)$. Then, we obtain $\mathbf{\Delta}$ as in (3.4) whose elements are $\mathbf{\Delta}(\boldsymbol{\alpha})_{ri} = \beta_l\, m_i\, z_{ir}\, S_{X_l}$ and

$$\mathbf{\Delta}(\boldsymbol{\beta})_{ji} = \begin{cases} \beta_l\, c_i\, x_{ij}\, S_{X_l}, & \text{for } j \neq l; \\ \beta_l\, c_i\, x_{il}\, S_{X_l} + d_\mu^{(i)}\, a_i\, S_{X_l}, & \text{for } j = l; \end{cases}$$

for $i = 1, \ldots, n$, $j = 1, \ldots, p$ and $r = 1, \ldots, q$, which must be evaluated at $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}$. In matrix form, $\mathbf{\Delta}(\boldsymbol{\beta}) = \beta_l\, S_{X_l}\, \boldsymbol{X}^\top c_i \delta_{ij}^n + S_{X_l}\, \mathbf{1}_{n\times p}^{(k)\top}\, d_\mu^{(i)}\, a_i\, \delta_{ij}^n$ and $\mathbf{\Delta}(\boldsymbol{\alpha}) = \beta_l\, S_{X_l}\, \boldsymbol{Z}^\top m_i\, \delta_{ij}^n$. Now, we additively perturb a continuous covariate-Z substituting z_k by $z_k + \omega S_{Z_k}$, where S_{Z_k} is the SD of Z_k. Here, the linear predictor component i is $\tau_i(\omega) = \alpha_1 + \cdots + \alpha_k[z_{ik} + \omega_i S_{Z_k}] + \cdots + \alpha_q z_{iq}$ and $\boldsymbol{\omega}_0 = \mathbf{0}_{n\times 1}$, so that now $\ell(\boldsymbol{\theta}|\boldsymbol{\omega}) = \sum_{i=1}^n \ell_i(\mu_i, \delta_i(\omega_i))$, $\mathbf{\Delta}(\boldsymbol{\beta})_{ji} = \alpha_k\, m_i\, x_{ij}\, S_{Z_k}$ and

$$\mathbf{\Delta}(\boldsymbol{\alpha})_{ri} = \begin{cases} \alpha_k\, w_i\, z_{ir}\, S_{Z_k}, & \text{for } r \neq k; \\ \alpha_k\, w_i\, z_{ik}\, S_{Z_k} + d_\delta^{(i)}\, b_i\, S_{Z_k}, & \text{for } r = k; \end{cases}$$

which must be evaluated at $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}$. In matrix form, $\mathbf{\Delta}(\boldsymbol{\alpha}) = \alpha_k S_{Z_k} \boldsymbol{Z}^\top w_i \delta_{ij}^n + S_{X_l} \mathbf{1}_{n\times q}^{(l)\top} d_\delta^{(i)} b_i \delta_{ij}^n$ and $\mathbf{\Delta}(\boldsymbol{\beta}) = \alpha_k\, S_{Z_k}\, \boldsymbol{X}^\top m_i\, \delta_{ij}^n$.

B.4. Joint covariate perturbation

A further perturbation scheme that we consider is when some perturbed covariate in X is also present in Z. For example, $x_{il} = z_{ik}$. Then, $\tau_i(\omega) = \alpha_1 + \cdots + \alpha_k[x_{il} + \omega_i S_{X_l}] + \cdots + \alpha_q z_{iq}$ and $\boldsymbol{\omega}_0 = \mathbf{0}_{n\times 1}$. Thus, now $\ell(\boldsymbol{\theta}|\boldsymbol{\omega}) = \sum_{i=1}^n \ell_i(\mu_i(\omega_i), \delta_i(\omega_i))$ and $\mathbf{\Delta}$ in (3.4) has elements

$$\mathbf{\Delta}(\boldsymbol{\beta})_{ji} = \begin{cases} \beta_l\, c_i\, x_{ij}\, S_{X_l} + \alpha_k\, m_i\, x_{ij}\, S_{X_l}, & \text{for } j \neq l; \\ \beta_l\, c_i\, x_{il}\, S_{X_l} + \alpha_k\, m_i\, x_{il}\, S_{X_l} + d_\mu^{(i)}\, a_i\, S_{X_l}, & \text{for } j = l; \end{cases}$$

$$\mathbf{\Delta}(\boldsymbol{\alpha})_{ri} = \begin{cases} \alpha_k\, w_i\, z_{ir}\, S_{X_l} + \beta_l\, m_i\, z_{ir}\, S_{X_l}, & \text{for } r \neq k; \\ \alpha_k\, w_i\, x_{il}\, S_{X_l} + \beta_l\, m_i\, x_{il}\, S_{X_l} + d_\delta^{(i)}\, b_i\, S_{X_l}, & \text{for } r = k. \end{cases}$$

In matrix form, $\mathbf{\Delta}(\boldsymbol{\beta}) = S_{X_l}[\boldsymbol{X}^\top\{\beta_l c_i \delta_{ij}^n + \alpha_k m_i \delta_{ij}^n\} + \mathbf{1}_{n\times p}^{(k)\top}\, a_i\, d_\mu^{(i)} \delta_{ij}^n]$ and $\mathbf{\Delta}(\boldsymbol{\alpha}) = S_{X_l}[\boldsymbol{Z}^\top\{\alpha_k w_i \delta_{ij}^n + \beta_l m_i \delta_{ij}^n\} + \mathbf{1}_{n\times q}^{(l)\top} b_i d_\delta^{(i)} \delta_{ij}^n]$. As mentioned, the matrices $\mathbf{\Delta}(\boldsymbol{\beta})$ and $\mathbf{\Delta}(\boldsymbol{\alpha})$ must be evaluated at $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}$.

Acknowledgements

The authors thank the Editors and referees for their constructive comments on an earlier version of this manuscript which resulted in this improved version. This study was partially supported by a CNPq and FACEPE grants from Brazil and by FONDECYT 1160868 grant from Chile.

References

[1] Bhatti, C. (2010). The Birnbaum-Saunders autoregressive conditional duration model. *Mathematics and Computers in Simulation*, **80** 2062–2078. doi:10.1016/j.matcom.2010.01.011. MR2665320

[2] Birnbaum, Z. W. and Saunders, S. C. (1969). A new family of life distributions. *Journal of Applied Probability*, **6** 319–327. doi:10.2307/3212003. MR0253493

[3] Cook, R. D. and Weisberg, S. (1983). Diagnostics for heteroscedasticity in regression. *Biometrika*, **70** 1–10. doi:10.2307/2335938. MR0742970

[4] Cox, D. and Hinkley, D. (1974). *Theoretical Statistics*. Chapman and Hall, London, UK. MR0370837

[5] Cysneiros, F., Paula, G., and Galea, M. (2007). Heteroscedastic symmetrical linear models. *Statistical and Probability Letters*, **77** 1084–1090. doi:10.1016/j.spl.2007.01.012. MR2395064

[6] Dunn, P. and Smyth, G. (1996). Randomized quantile residuals. *Journal of Computational and Graphical Statistics*, **5** 236–244. doi:10.2307/1390802

[7] Ferrari, S., Espinheira, P., and Cribari-Neto, F. (2011). Diagnostic tools in beta regression with varying dispersion. *Statistica Neerlandica*, **65** 337–351. doi:10.1111/j.1467-9574.2011.00488.x. MR2857878

[8] Ferreira, M., Gomes, M. I., and Leiva, V. (2012). On an extreme value version of the Birnbaum-Saunders distribution. *REVSTAT Statistical Journal*, **10** 181–210. https://www.ine.pt/revstat/pdf/rs120202.pdf. MR2960731

[9] Galea, M., Leiva, V., and Paula, G. (2004). Influence diagnostics in log-Birnbaum-Saunders regression models. *Journal of Applied Statistics*, **31** 1049–1064. doi:10.1080/0266476042000280409. MR2100116

[10] Garcia-Papani, F., Uribe-Opazo, M., Leiva, V., and Aykroyd, R. (2016). Birnbaum-Saunders spatial modelling and diagnostics applied to agricultural engineering data. *Stochastic Environmental Research and Risk Assessment*, pages in press available at doi:10.1007/s00477-015-1204-4.

[11] Jin, X. and Kawczak, J. (2003). Birnbaum-Saunders and lognormal kernel estimators for modelling durations in high frequency financial data. *Annals of Economics and Finance*, **4** 103–124. http://aeconf.com/Articles/May2003/aef040106.pdf.

[12] Johnson, N., Kotz, S., and Balakrishnan, N. (1995). *Continuous Univariate Distributions*, volume 2. Wiley, New York, US. MR1326603

[13] Leiva, V. (2016). *The Birnbaum-Saunders Distribution*. Academic Press, New York, US. MR3430824

[14] Leiva, V., Marchant, C., Ruggeri, F., and Saulo, H. (2015a). A criterion for environmental assessment using Birnbaum-Saunders attribute control charts. *Environmetrics*, **26** 463–476. doi:10.1002/env.2349. MR3415566

[15] Leiva, V., Rojas, E., Galea, M., and Sanhueza, A. (2014a). Diagnostics in Birnbaum-Saunders accelerated life models with an application to fatigue data. *Applied Stochastic Models in Business and Industry*, **30** 115–131. doi:10.1002/asmb.1944. MR3191346

[16] Leiva, V., Santos-Neto, M., Cysneiros, F. J. A., and Barros, M. (2014b). Birnbaum-Saunders statistical modelling: A new approach. *Statistical Modelling*, **14** 21–48. doi:10.1177/1471082X13494532. MR3179546

[17] Leiva, V., Santos-Neto, M., Cysneiros, F. J. A., and Barros, M. (2016). A methodology for stochastic inventory models based on a zero-adjusted Birnbaum-Saunders distribution. *Applied Stochastic Models in Business and Industry*, **32** 74–89. doi:10.1002/asmb.2124. MR3460890

[18] Leiva, V., Saulo, H., Leão, J., and Marchant, C. (2014c). A family of autoregressive conditional duration models applied to financial data. *Computational Statistics and Data Analysis*, **79** 175–191. doi:10.1016/j.csda.2014.05.016. MR3227995

[19] Leiva, V., Tejo, M., Guiraud, P., Schmachtenberg, O., Orio, P., and Marmolejo, F. (2015b). Modeling neural activity with cumulative damage distributions. *Biological Cybernetics*, **109** 421–433. doi:10.1007/s00422-015-0651-9. MR3397465

[20] Li, A., Chen, Z., and Xie, F. (2012). Diagnostic analysis for heterogeneous log-Birnbaum-Saunders regression models. *Statistical and Probability Letters*, **89** 1690–1698. doi:10.1016/j.spl.2012.05.021. MR2951006

[21] Lin, J., Zhu, L., and Xie, F. (2009). Heteroscedasticity diagnostics for t linear regression models. *Metrika*, **70** 59–77. doi:10.1007/s00184-008-0179-2. MR2506500

[22] Marchant, C., Leiva, V., and Cysneiros, F. (2016a). A multivariate log-linear model for Birnbaum-Saunders distributions. *IEEE Transactions on Reliability*, **65** 816–827 doi:10.1109/TR.2015.2499964.

[23] Marchant, C., Leiva, V., Cysneiros, F., and Vivanco, J. (2016b). Diagnostics in multivariate generalized Birnbaum-Saunders regression models. *Journal of Applied Statistics*, **43** 2829–2849 doi:10.1080/02664763.2016.1148671. MR3546117

[24] Owen, W. (2006). A new three-parameter extension to the Birnbaum-Saunders distribution. *IEEE Transactions on Reliability*, **55** 475–479. doi:10.1109/TR.2006.879646.

[25] Owen, W. and Padgett, W. (2000). A Birnbaum-Saunderss accelerated life model. *IEEE Transactions on Reliability*, **49** 224–229. doi:10.1109/24.877342.

[26] Paula, G. A. (2013). On diagnostics in double generalized linear models. *Computational Statistics and Data Analysis*, **68** 44–51. doi:10.1016/j.csda.2013.06.008. MR3103761

[27] Paula, G. A., Leiva, V., Barros, M., and Liu, S. (2012). Robust statistical modeling using the Birnbaum-Saunders-t distribution applied to

insurance. *Applied Stochastic Models in Business and Industry*, **28** 16–34. doi:10.1002/asmb.887. MR2898899

[28] Qu, H. and Xie, F. (2011). Diagnostics analysis for log-Birnbaum-Saunders regression models with censored data. *Statistica Neerlandica*, **65** 1–21. doi:10.1111/j.1467-9574.2010.00467.x. MR2798680

[29] R-Team (2016). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. https://cran.r-project.org/doc/manuals/r-release/fullrefman.pdf.

[30] Rieck, J. and Nedelman, J. (1991). A log-linear model for the Birnbaum-Saunders distribution. *Technometrics*, **3** 51–60. http://www.jstor.org/stable/1269007.

[31] Rocha, A. V. and Simas, A. B. (2011). Influence diagnostic in a general class of beta regression models. *TEST*, **20** 95–119. doi:10.1007/s11749-010-0189-z. MR2806312

[32] Rojas, F., Leiva, V., Wanke, P., and Marchant, C. (2015). Optimization of contribution margins in food services by modeling independent component demand. *Revista Colombiana de Estadística*, **38** 1–30. doi:10.15446/rce.v38n1.48799. MR3374167

[33] Santos-Neto, M., Cysneiros, F., Leiva, V., and Barros, M. (2016). *RBS: Reparameterized Birnbaum-Saunders regression model*. R package version 0.0.1 https://github.com/santosneto/RBS.

[34] Santos-Neto, M., Cysneiros, F. J. A., Leiva, V., and Ahmed, S. (2012). On new parameterizations of the Birnbaum-Saunders distribution. *Pakistan Journal of Statistics*, **28** 1–26. MR2931825

[35] Santos-Neto, M., Cysneiros, F. J. A., Leiva, V., and Barros, M. (2014). On new parameterizations of the Birnbaum-Saunders distribution and its moments, estimation and application. *REVSTAT Statistical Journal*, **12** 247–272. https://www.ine.pt/revstat/pdf/rs140303.pdf. MR3301849

[36] Saulo, H., Leiva, V., Ziegelmann, F. A., and Marchant, C. (2013). A nonparametric method for estimating asymmetric densities based on skewed Birnbaum-Saunders distributions applied to environmental data. *Stochastic Environmental Research and Risk Assessment*, **27** 1479–1491. doi:10.1007/s00477-012-0684-8

[37] Saumard, A. (2013). Optimal model selection in heteroscedastic regression using piecewise polynomial functions. *Electronic Journal of Statistics*, **7** 1184–1223. doi:10.1214/13-EJS803. MR3056072

[38] Simas, A. B., Barreto-Souza, W., and Rocha, A. V. (2010). Improved estimators for a general class of beta regression models. *Computational Statistics and Data Analysis*, **54** 348–366. doi:10.1016/j.csda.2009.08.017. MR2756431

[39] Smyth, G. (1989). Generalized linear models with varying dispersion. *Journal of the Royal Statistical Society B*, **51** 47–60. http://www.jstor.org/stable/2345840. MR0984992

[40] Smyth, G. and Verbyla, A. (1999). Adjusted likelihood methods for modelling dispersion in generalized linear models. *Envi-*

ronmetrics, **10** 695–709. doi:10.1002/(SICI)1099-095X(199911/12)10:6 <695::AID-ENV385>3.0.CO;2-M.

[41] Stasinopoulos, D. and Rigby, R. (2007). Generalized additive models for location, scale and shape (GAMLSS). *Journal of Statistical Software*, **23** 1–46. doi:10.18637/jss.v023.i07.

[42] Taylor, J. and Verbyla, A. (2004). Joint modeling of location and scale parameters of the t distribution. *Statistical Modelling*, **4** 91–112. doi:10.1191/1471082X04st068oa. MR2109340

[43] Van Keilegom, I. and Wang, L. (2010). Semiparametric modeling and estimation of heteroscedasticity in regression analysis of cross-sectional data. *Electronic Journal of Statistics*, **4** 133–160. doi:10.1214/09-EJS547. MR2645480

[44] Vanegas, L., Rondon, L., and Cysneiros, F. (2012). Diagnostic procedures in Birnbaum-Saunders nonlinear regression models. *Computational Statistics and Data Analysis*, **56** 1662–1680. doi:10.1016/j.csda.2011.10.008. MR2892367

[45] Venezuela, M. K. and Artes, R. (2014). Estimating equations and diagnostic techniques applied to zero-inflated models for panel data. *Electronic Journal of Statistics*, **8** 1641–1660. doi:10.1214/14-EJS936. MR3263133

[46] Villegas, C., Paula, G., and Leiva, V. (2011). Birnbaum-Saunders mixed models for censored reliability data analysis. *IEEE Transactions on Reliability*, **60** 748–758. doi:10.1109/TR.2011.2170251.

[47] Wanke, P. and Leiva, V. (2015). Exploring the potential use of the Birnbaum-Saunders distribution in inventory management. *Mathematical Problems in Engineering*, Article ID 827246:1–9. doi:10.1155/2015/827246. MR3421264

[48] Weisberg, S. (2005). *Applied Linear Regression*. Wiley, New York, US. doi:10.1002/0471704091. MR2112740

[49] Wu, L., Zhang, Z., and Xu, D. (2012). Variable selection for joint mean and dispersion models of the lognormal distribution. *Hacettepe Journal of Mathematics and Statistics*, **41** 307–320. `http://www.hjms.hacettepe.edu.tr/uploads/b6d36a93-bbda-4e78-a4a7-940cc4d2e31f.pdf`. MR3012185

[50] Xie, F. and Wei, B. (2007). Diagnostics analysis for log-Birnbaum-Saunders regression models. *Computational Statistics and Data Analysis*, **51** 4692–4706. doi:10.1016/j.csda.2006.08.030. MR2361364