On some properties of a new asymmetric-based tobit model

Mário F. Desousa[1], **Helton Saulo**[2], **Víctor Leiva**[3], **Paulo Scalco**[1]

[1]Faculty of Management, Accounting and Economics, Universidade Federal de Goiás, Brazil
[2]Institute of Mathematics and Statistics, Universidade Federal de Goiás, Brazil
[3]Faculty of Engineering and Sciences, Universidad Adolfo Ibáñez, Chile

Abstract

A strong assumption of the standard tobit model is that its errors follow a normal distribution. However, not all applications are well modeled by this distribution. Some efforts have relaxed the normality assumption by considering more flexible distributions such as t and log-alpha-power. Nevertheless, presence of asymmetry cannot be described by these flexible distributions. To overcome this problem, we propose a tobit model with errors following a Birnbaum-Saunders distribution. We discuss some of its properties including maximum likelihood estimators, residuals and global and local influence diagnostic tools. As part of a further extension of this paper, we leave the following issues. The implementation of Monte Carlo simulations to evaluate the performance of both the maximum likelihood estimators and residuals, and the application to real data.

Keywords Birnbaum-Saunders distribution; Diagnostics; Maximum Likelihood Estimation; Tobit models.

Resumo

Um forte pressuposto dos modelos tobit é que os erros seguem uma distribuição normal, mas nem todas as aplicações são modeladas da melhor forma com essa distribuição. Esforços no sentido de relaxar a hipótese de normalidade consideram distribuições mais flexíveis como a t e a log-alpha-power, contudo essas não modelam assimetria. De modo a resolver o problema, propomos um modelo tobit basedo na distribuição Birnbaum-Saunders e discutimos algumas de suas propriedades, incluindo estimadores de máxima verossimilhança, resíduos e ferramentas de diagnóstico de influência global e local. Como parte de uma extensão natural desse artigo, deixamos os seguintes problemas. A implementação de simulações de Monte Carlo para avaliar o desempenho dos estimadores de máxima verossimilhança e dos resíduos, e a aplicação a dados reais.

Keywords Distribuição Birnbaum-Saunders; Diagnósticos; Máxima verossimilhança; Modelos tobit.

JEL Codes: C13,C24,C52

Área ANPEC: Microeconomia, Métodos Quantitativos e Finanças

1 Introduction

The tobit model was proposed by Tobin (1958) for limited dependent (responses) variables and named tobit by Goldberger (1964), because of its similarity with the probit model. According to Amemiya (1984), Tobin (1958) was motivated to develop his model by a case-study where he needed to study the relationship between household expenditure on a durable good and household incomes. The common regression approach with ordinary least squares could not be used because there were many observations where the expenditure was zero, which obviously destroyed the assumption of linearity. To solve the problem, Tobin (1958) proposed a model that could fit the data appropriately formulating a regression model whose response was limited (or censored) to a prefixed limiting value.

Tobit models have been used extensively to describe censored data in econometrics. The censoring occurs when the response of the regression model is not directly observable but its independent variables. Although tobit models were born in economics, they have been applied in biology and engineering as well. The biology research relies most in analyzing the survival time of a patient (Leiva et al., 2007); the engineers use it to model time to failure of various types of materials or machines (Villegas et al., 2011); whereas the sociologists employ it to describe the duration of marriage, residing in a particular region or unemployment, as well the time between births, but the tobit models have also been used to describe inheritance, ratio of unemployed hours to employed hours and expected age of retirement (Amemiya, 1984). Tobit models could also be used in environmental sciences, where censored data also are present (Thorarinsdottir and Gneiting, 2010; Helsel, 2011).

As mentioned, tobit models have a strong normality assumption. Proposals of tobit models that relax the assumption of normality are of great importance in empirical economics, because most of the data available in the real-world cannot be modeled optimally with a normal distribution. Barros et al. (2010) noticed that the asymmetry of the data and its kurtosis can be, and usually are, different from the expected for a normal distribution, so that more flexible models are needed.

The Birnbaum-Saunders (BS) distribution has been widely studied. It is positively skewed, has a failure rate with upside-down bathtub shape, and a close relation with the normal model; see the seminal paper by Birnbaum and Saunders (1969) and the books by Johnson et al. (1995, pp. 651-663) and Leiva (2016). The BS distribution was derived in terms of shape and scale parameters, but the latter is also its median. Thus, the BS distribution can be seen as an analogue to the normal model, but in an asymmetrical setting, where the median can be a better measure of central tendency than the mean. The BS distribution has been applied to biological, economic, engineering and environmental data, which have been conducted by international, transdisciplinary groups of researchers. Some of its recent applications are attributed to Qu and Xie (2011), Ferreira et al. (2012), Li et al. (2012), Saulo et al. (2013), Leiva et al. (2014, 2015a,b), Garcia-Papani et al. (2016) and Marchant et al. (2016a,b).

The main objective of this paper is propose a tobit-BS model by relaxing the assumption of normality and supposing that its errors follow a BS distribution, which according to the best of our knowledge, it is a topic that has not been studied before. The secondary objectives of this paper are: (i) to develop inference for the tobit-BS model based on the maximum likelihood (ML) method; (ii) to derive residuals and global and local influence tools for model checking and diagnostics. As part of a future extension, we leave the implementation of Monte Carlo simulations and an application to real data.

The paper presents, along with this introduction, five sections. Section 2 provides a background of classic tobit models and of the BS and log-BS distributions. Section 3 formulates the tobit-BS model along with inference and estimation based on the ML method. Section 4 derives diagnostics for the proposed model including residuals analysis and local influence tools. Finally, Section 5 discusses some conclusions and future works.

2 Background

2.1 The tobit model

Let $\boldsymbol{Y} = (Y_1, \ldots, Y_m, Y_{m+1}, \ldots, Y_n)^\top$ be a sample of size n, that is, independent (IND) random variables but not independent identically distributed (IID) necessarily. Assume that this sample includes m censored data to the left and $n - m$ observed (complete or uncensored) data. Thus, such censoring scheme can be visualized under a regression setting with a censored response Y^*, which is a (unobserved) latent variable. Hence, the m censored data (unobserved)correspond to the values of Y^* less than or equal to a threshold point y_0 (censoring to the left), so that all of these data take the value y_0. The other $n - m$ data (observed) are related to values of Y^* greater than y_0, which can be described by a linear regression structure of the type $\boldsymbol{x}_i^\top \boldsymbol{\beta}$. All this modeling environment may be formulated by the normal tobit model with censored response to the left as

$$Y_i = \begin{cases} y_o, & \text{if } Y_i^* \ \leq \ y_0, \quad i = 1, \ldots, m; \\ \boldsymbol{x}_i^\top \boldsymbol{\beta} + \varepsilon_i, & \text{if } Y_i^* \ > \ y_0, \quad i = m+1, \ldots, n; \end{cases} \tag{1}$$

where $\varepsilon_i \overset{\text{IID}}{\sim} \mathrm{N}(0, \sigma^2)$ is the model error term, $\boldsymbol{\beta}$ is a vector of regression coefficients corresponding to unknown parameters to be estimated, and $\boldsymbol{x}_i$ is a vector containing the values of explanatory variables (covariates). Observe that y_0 given in (1) is a prefixed limiting value that does the response of the regression model defined in (1) to be limited (or censored), as mentioned by Tobin (1958).

Note the similarity between the normal probit model and the normal tobit model defined in (1). In the normal probit model, the response is an latent (unobserved) variable described by

$$Y_i^* = \boldsymbol{x}_i^\top \boldsymbol{\beta} + \varepsilon_i, \quad i = 1, \ldots, n, \tag{2}$$

where $\boldsymbol{x}_i$, $\boldsymbol{\beta}$ and ε_i are defined analogously as in (1). As it is not possible to observe the latent variable Y_i^*, the indicator variable

$$Y_i = \begin{cases} 0, & Y_i^* \leq y_0, \quad i = 1, \ldots, m; \\ 1, & Y_i^* > y_0, \quad i = m+1, \ldots, n; \end{cases} \tag{3}$$

is defined, However, instead of using $\boldsymbol{x}^\top\boldsymbol{\beta}$, from (2) and (3), the expected response is formulated as

$$\mathrm{E}[Y|\boldsymbol{x}] = \mathrm{P}(Y=1) = \mathrm{P}(Y^* > y_0) = \mathrm{P}(\boldsymbol{x}^\top\boldsymbol{\beta} + \varepsilon > y_0) = 1 - \Phi(y_0 - \boldsymbol{x}^\top\boldsymbol{\beta}), \tag{4}$$

where Φ is the cumulative distribution function (CDF) of the standard normal distribution. Note that other CDFs might be assumed for Φ in (4), so expanding the covering of the probit model.

Probit and tobit models are the same for the latent variable (Y^*), but models for the observed response (Y) are different. In the tobit model, we know the value of Y^* when $Y^* > y_0$, whereas in the probit model we just know that $Y^* > y_0$, but we do not know its value. Thus, there is more information in the tobit model than in the probit model. Also, the estimates of $\boldsymbol{\beta}$ in the tobit model are more efficient than in the probit model. Moreover, for the censored cases of the probit model, it is not possible to estimate the variance of Y^*. However, in the tobit model, this variance can be estimated from the data. For more details, see Scott (1997, p. 199).

2.2 The Birnbaum-Saunders distribution

If a random variable T follows a BS distribution with shape parameter α and scale parameter σ, we use the notation $T \sim \mathrm{BS}(\alpha, \sigma)$. This distribution can be defined by its CDF given by

$$F_T(t; \alpha, \sigma) = \Phi\left(\frac{1}{\alpha}\left(\sqrt{t/\sigma} - \sqrt{\sigma/t}\right)\right), \quad t > 0, \alpha > 0, \sigma > 0. \tag{5}$$

Then, the probability density function (PDF) of T obtained from (5) is expressed as

$$f_T(t; \alpha, \sigma) = \frac{1}{2\alpha}\left(\sqrt{1/\sigma t} + \sqrt{\sigma/t^{\frac{3}{2}}}\right)\phi\left(\frac{1}{\alpha}\left(\sqrt{t/\sigma} - \sqrt{\sigma/t}\right)\right), \quad t > 0, \alpha > 0, \sigma > 0, \tag{6}$$

where ϕ is the standard normal PDF. Thus, the PDF in (6) can be rewritten as

$$f_T\left(t; \alpha, \sigma\right) = \frac{\exp(\alpha^{-2})}{2\alpha\sqrt{2\pi\sigma}} \exp\left(-\frac{1}{2\alpha^2}\left(\frac{t}{\sigma} + \frac{\sigma}{t}\right)\right) t^{-\frac{3}{2}}\left(t + \sigma\right), \quad t > 0, \alpha > 0, \sigma > 0. \tag{7}$$

Note that the results provided in (5) and (6) may be obtained from the transformation theorem of random variables by using

$$T = \sigma\left(\alpha Z/2 + \sqrt{(\alpha Z/2)^2 + 1}\right)^2, \tag{8}$$

where $Z \sim \mathrm{N}(0, 1)$. Also from (8), it may be verified that a continuous random variable T has a BS distribution with parameters $\alpha > 0$ and $\sigma > 0$, if and only if $Z = (1/\alpha)\big(\sqrt{T/\sigma} - \sqrt{\sigma/T}\big) \sim \mathrm{N}(0, 1)$. Some properties of the BS distribution are presented as follows. If $T \sim \mathrm{BS}(\alpha, \sigma)$, then: (i) $\mathrm{E}(T) = \sigma(1 + \alpha^2/2)$ and $\mathrm{Var}(T) = (\alpha\sigma)^2(1 + 5\alpha^2/4)$; (ii) for $b > 0$, $bT \sim \mathrm{BS}(\alpha, b\sigma)$, which means that the BS distribution is closed under scalar multiplication (proportionality); (iii) $1/T \sim \mathrm{BS}(\alpha, 1/\sigma)$, implying that the BS distribution is closed under reciprocation; (iv) the median of the distribution of T is σ, which can be directly obtained when $q = 0.5$ from its quantile function given by

$$t(q;\alpha,\sigma) = F_T^{-1}(q;\alpha,\sigma) = \sigma\left(\frac{\alpha z(q)}{2} + \sqrt{\left(\frac{\alpha z(q)}{2}\right)^2 + 1}\right)^2, \quad 0 < q < 1,$$

where $z(q)$ is the standard normal quantile function; and (v) the BS distribution is positively skewed as α increases and approximately symmetrical around σ as α goes to zero; see Figure 1(left). Properties of proportionality and reciprocation given above in (ii) and (iii) are easily verified by using once again the mentioned transformation theorem.

The survival function (SF) and hazard rate (HR) of $T \sim \mathrm{BS}(\alpha,\sigma)$ are given respectively by

$$\begin{aligned} S_T(t;\alpha,\sigma) &= \Phi\left(-\frac{1}{\alpha}\left(\sqrt{\frac{t}{\sigma}} - \sqrt{\frac{\sigma}{t}}\right)\right), \\ h_T(t;\alpha,\sigma) &= \frac{\phi\left(\frac{1}{\alpha}\left(\sqrt{\frac{t}{\sigma}} - \sqrt{\frac{\sigma}{t}}\right)\right)\frac{t^{-3/2}(t+\sigma)}{2\alpha\sigma^{1/2}}}{\Phi\left(-\frac{1}{\alpha}\left(\sqrt{\frac{t}{\sigma}} - \sqrt{\frac{\sigma}{t}}\right)\right)}, \quad t > 0. \end{aligned}$$

2.3 The log-Birnbaum-Saunders distribution

When modeling data with the BS distribution, its logarithmic version (log-BS) is needed. A random variable Y has a log-BS distribution with shape ($\alpha > 0$) and location ($\mu \in \mathbb{R}$) parameter, which is denoted by $\text{log-BS}(\alpha,\mu)$, if and only if $Z = (2/\alpha)\sinh((Y-\mu)/2) \sim \mathrm{N}(0,1)$. Then, the CDF of Y is given by

$$F_Y(y;\alpha,\mu) = \Phi\left(\frac{2}{\alpha}\sinh\left(\frac{y-\mu}{2}\right)\right), \quad y \in \mathbb{R}, \mu \in \mathbb{R}, \alpha > 0. \tag{9}$$

Consequently, from (9), the PDF of Y is obtained as

$$f_Y(y;\alpha,\mu) = \frac{1}{\alpha\sqrt{2\pi}}\cosh\left(\frac{y-\mu}{2}\right)\exp\left(-\frac{2}{\alpha^2}\sinh^2\left(\frac{y-\mu}{2}\right)\right), \quad y \in \mathbb{R}, \mu \in \mathbb{R}, \alpha > 0, \tag{10}$$

whereas that the logarithm of the PDF given in (10) is expressed as

$$\log(f_Y(y;\alpha,\mu)) = -\log(2) - \frac{\log(2\pi)}{2} + \log\left(\frac{2}{\alpha}\cosh\left(\frac{y-\mu}{2}\right)\right) - \frac{2}{\alpha^2}\sinh^2\left(\frac{y-\mu}{2}\right), \quad y \in \mathbb{R}. \tag{11}$$

Some properties of the log-BS distribution are presented as follows. If $Y \sim \text{log-BS}(\alpha,\mu)$, then: (i) $T = \exp(Y) \sim \mathrm{BS}(\alpha,\sigma)$, which means that the log-BS PDF given in (10) can be obtained from the standard normal PDF as in (9) or from the BS PDF defined in (7); (ii) $\mathrm{E}(Y) = \mu$; (iii) there is no closed form for the variance of Y, but based upon an asymptotic approximation for the moment generating function of the log-BS distribution, it follows that, if $\alpha \to 0$, then $\mathrm{Var}(T) = \alpha^2 - \alpha^4/4$, whereas that, if $\alpha \to \infty$, then $\mathrm{Var}(T) = 4(\log^2(\sqrt{2}\alpha) + 2 - 2\log(\sqrt{2}\alpha))$; (iv) if $X = \pm Y + d$, then $X \sim \text{log-BS}(\alpha, \pm\mu + d)$; and (v) the log-BS distribution is symmetric around μ, unimodal for $\alpha \leq 2$ and bimodal for $\alpha > 2$; see Figure 1(right).

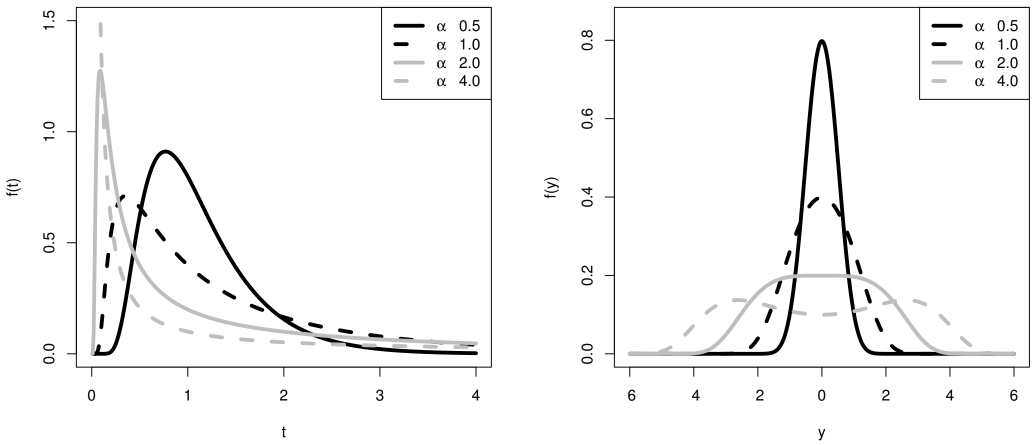

Figure 1: PDF of (left) BS$(\alpha, 1)$ and (right) log-BS$(\alpha, 0)$ distributions for the indicated value of α.

3 The tobit-BS model

3.1 Formulation

Consider the BS regression model

$$T_i = \exp(\boldsymbol{x}_i^\top \boldsymbol{\beta})\delta_i, \quad i = 1, \ldots, n, \tag{12}$$

originally proposed by (Rieck and Nedelman, 1991), where T_i is the response, $\boldsymbol{x}_i$ and $\boldsymbol{\beta}$ are analogously defined as in (2), and $\delta_i \sim \mathrm{BS}(\alpha, 1)$ is the model error. By applying logarithm in (12), we obtain

$$Y_i = \boldsymbol{x}_i^\top \boldsymbol{\beta} + \varepsilon_i, \quad i = 1, \ldots, n, \tag{13}$$

where $Y_i = \log(T_i)$ is the log-response i, $\boldsymbol{x}_i$ and $\boldsymbol{\beta}$ are similar to (1), and $\varepsilon_i = \log(\delta_i) \sim \text{log-BS}(\alpha, 0)$ is the error term of the model. Then, based on (1) and (13), we propose a tobit-BS model as

$$Y_i = \begin{cases} y_0, & \text{if} \quad Y_i^* \leq y_0, \quad i = 1, \ldots, m, \\ \boldsymbol{x}_i^\top \boldsymbol{\beta} + \varepsilon_i, & \text{if} \quad Y_i^* > y_0, \quad i = m+1, \ldots, n, \end{cases} \tag{14}$$

where $Y_i^* = \log(T_i^*)$, and $\boldsymbol{\beta}$, $\boldsymbol{x}_i$ and ε_i are given in (13).

3.2 Estimation

We estimate the parameters of the tobit-BS model defined in (14) with the ML method, in which case the log-likelihood function for $\boldsymbol{\theta} = (\alpha, \boldsymbol{\beta}^\top)^\top$ obtained from (11) takes the form

$$l(\boldsymbol{\theta}) = -(n-m)\log(2)-(n-m)\frac{\log(2\pi)}{2}+\sum_{i=1}^{m}(\log\Phi(\zeta_{i2}^{\mathrm{c}}))+\sum_{i=m+1}^{n}\left(\log(\zeta_{i1})-\zeta_{i2}^{2}/2\right), \quad (15)$$

where

$$\zeta_{i2}^{\mathrm{c}} = \tfrac{2}{\alpha}\sinh\left(\tfrac{y_0-\boldsymbol{x}_i^\top\boldsymbol{\beta}}{2}\right), \quad \zeta_{i1} = \tfrac{2}{\alpha}\cosh\left(\tfrac{y_i-\boldsymbol{x}_i^\top\boldsymbol{\beta}}{2}\right), \quad \zeta_{i2} = \tfrac{2}{\alpha}\sinh\left(\tfrac{y_i-\boldsymbol{x}_i^\top\boldsymbol{\beta}}{2}\right). \quad (16)$$

To obtain the ML estimators, it is necessary to maximize the log-likelihood given in (15). The score vector $\dot{\boldsymbol{\ell}} = \partial\ell(\boldsymbol{\theta})/\partial\boldsymbol{\theta} = (\dot{\ell}_\alpha, \dot{\boldsymbol{\ell}}_{\boldsymbol{\beta}}^\top)^\top$, containing the first partial derivatives of (15), where

$$\dot{\ell}_\alpha = \begin{cases} -\frac{\lambda(\zeta_{i2}^{\mathrm{c}})\zeta_{i2}^{\mathrm{c}}}{\alpha}, & i=1,\ldots,m; \\ \frac{\zeta_{i2}^2-1}{\alpha}, & i=m+1,\ldots,n; \end{cases} \quad \dot{\boldsymbol{\ell}}_{\boldsymbol{\beta}} = \begin{cases} -\frac{x_{ij}\cosh(\delta)\lambda(\zeta_{i2}^{\mathrm{c}})}{\alpha}, & i=1,\ldots,m; \\ \frac{x_{ij}\sinh(2\delta_i)}{\alpha^2}-\frac{x_{ij}\tanh(\delta_i)}{2}, & i=m+1,\ldots,n; \end{cases} \quad (17)$$

where $\delta_i = y_0 - x_i^\top\beta/2$, if $i = 1,\ldots,m$, and $\delta_i = y_i - x_i^\top\boldsymbol{\beta}/2$, if $i = m+1,\ldots,n$, and $\lambda(\zeta_{i2}^{\mathrm{c}}) = \phi(\zeta_{i2}^{\mathrm{c}})/\Phi(\zeta_{i2}^{\mathrm{c}})$. The ML estimator of $\boldsymbol{\theta}$ is obtained equating (17) to zero. However, the system of equations defined by these two equations does not have an analytic solution. Leiva et al. (2007) suggested the use of the Broyden-Fletcher-Goldfarb-Shanno algorithm, using as a starting values for the numerical procedure $\widehat{\alpha}^2 = 4(\sinh((y_i - \boldsymbol{x}_i^\top\widehat{\boldsymbol{\beta}})/2))^2/(n-m)$ and $\widehat{\boldsymbol{\beta}} = (\boldsymbol{X}^\top\boldsymbol{X})^{-1}\boldsymbol{X}^\top\boldsymbol{y}$, where $\boldsymbol{X}$ is a matrix composed by the rows $\boldsymbol{x}_i$. Recall that the log-BS distribution can be bimodal if $\alpha > 2$. It implicates that the log-likelihood function has more than one maximum value. However, Rieck and Nedelman (1991) and Leiva (2016) argued that $\alpha > 2$ is unusual in practice, which means that the maximum point is often unique.

3.3 Inference

Assuming that the regularity conditions defined in Cox and Hinkley (1974) are satisfied, the ML estimators $\widehat{\alpha}$ and $\widehat{\boldsymbol{\beta}}$ are consistent and follow a multivariate normal joint asymptotic distribution. This distribution has an asymptotic mean vector with elements α and $\boldsymbol{\beta}$ and have an asymptotic covariance matrix equal to $\mathcal{J}(\boldsymbol{\theta})^{-1}$, which can be approximated by the expected Fisher information matrix. Therefore, as $n \to \infty$, we have that

$$\sqrt{n}(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \quad \overset{\mathrm{d}}{\to} \quad \mathrm{N}_{p+1}\left(\mathbf{0}_{p+1}, \mathcal{J}(\boldsymbol{\theta})^{-1}\right), \quad (18)$$

where $\mathcal{J}(\boldsymbol{\theta}) = \lim_{n\to\infty}(1/n)\mathcal{I}(\boldsymbol{\theta})$, with $\mathcal{I}(\boldsymbol{\theta})$ being the expected Fisher information matrix. In addition, $\overset{\mathrm{d}}{\to}$ means convergence in distribution to and $\mathbf{0}_{p+1}$ is a $p \times 1$ vector of zeros. Note that $\widehat{\mathcal{I}}(\boldsymbol{\theta})^{-1}$ is a consistent estimator of the asymptotic variance-covariance matrix of $\widehat{\boldsymbol{\theta}}$, $\mathcal{J}(\boldsymbol{\theta})^{-1}$ say. In practice, one may approximate the expected Fisher information matrix by its observed version (Efron and Hinkley, 1978), whereas the elements of the diagonal of the inverse of the observed information matrix can be used to approximate the corresponding standard errors (SEs). The observed Fisher information matrix is obtained from the Hessian matrix, which contains the second partial derivatives of (15) and given by

$$\ddot{\boldsymbol{\ell}} = \begin{pmatrix} \mathrm{tr}(\boldsymbol{G}) & \boldsymbol{k}^\top \boldsymbol{X} \\ \boldsymbol{X}^\top \boldsymbol{k} & \boldsymbol{X}^\top \boldsymbol{V} \boldsymbol{X} \end{pmatrix}, \tag{19}$$

where $\boldsymbol{V} = \mathrm{diag}\{v_1(\boldsymbol{\theta}), v_2(\boldsymbol{\theta}), v_3(\boldsymbol{\theta}), \ldots, v_n(\boldsymbol{\theta})\}$, $\boldsymbol{k} = (k_1(\boldsymbol{\theta}), k_2(\boldsymbol{\theta}), k_3(\boldsymbol{\theta}), \ldots, k_n(\boldsymbol{\theta}))^\top$ and $\boldsymbol{G} = \mathrm{diag}\{g_1(\boldsymbol{\theta}), g_2(\boldsymbol{\theta}), g_3(\boldsymbol{\theta}), \ldots, g_n(\boldsymbol{\theta})\}$, with

$$v_i(\boldsymbol{\theta}) = \begin{cases} \frac{\sinh(\delta_i)\lambda(\zeta_{i2}^{\mathrm{c}})}{2\alpha} - \frac{\zeta_{i1}^{\mathrm{c}}\zeta_{i2}^{\mathrm{c}}\cosh(\delta_i)\lambda(\zeta_{i2}^{\mathrm{c}})}{2\alpha} - \frac{\cosh^2(\delta_i)\lambda^2(\zeta_{i2}^{\mathrm{c}})}{\alpha^2}, & i = 1, \ldots, m; \\ \frac{1}{4}(\mathrm{sech}(\delta_i))^2) - \frac{1}{\alpha^2}\cosh(2\delta_i), & i = m+1, \ldots, n; \end{cases} \tag{20}$$

$$k_i(\boldsymbol{\theta}) = \begin{cases} -\frac{1}{2\alpha}\lambda(\zeta_{i2}^{\mathrm{c}})(\zeta_{i2}^{\mathrm{c}})^2\zeta_{i1}^{\mathrm{c}} + \frac{\lambda(\zeta_{i2}^{\mathrm{c}})\cosh(\delta_i)}{\alpha^2} - \frac{\zeta_{i2}^{\mathrm{c}}}{\alpha^2}\lambda^2(\zeta_{i2}^{\mathrm{c}})\cosh(\delta_i), & i = 1, \ldots, m; \\ -\frac{2}{\alpha^3}\sinh(2\delta_i), & i = m+1, \ldots, n; \end{cases} \tag{21}$$

$$g_i(\boldsymbol{\theta}) = \begin{cases} -\frac{\zeta_{i2}^{\mathrm{c}}}{\alpha^2}(((\zeta_{i2}^{\mathrm{c}})^2\lambda(\zeta_{i2}^{\mathrm{c}}) + \zeta_{i2}^{\mathrm{c}}\lambda^2(\zeta_{i2}^{\mathrm{c}})) - 2\lambda(\zeta_{i2}^{\mathrm{c}})), & i = 1, \ldots, m; \\ \frac{1}{\alpha^2} - \frac{3\zeta_{i2}}{\alpha^2}, & i = m+1, \ldots, n. \end{cases} \tag{22}$$

and ζ_{i1}, ζ_{i2} and ζ_{i2}^{c} presented in (16), whereas $\zeta_{i1}^{\mathrm{c}} = (2/\alpha)\cosh((y_0 - \boldsymbol{x}_i^\top\boldsymbol{\beta})/2)$, for $i = 1, \ldots, m$. Asymptotic inference for the tobit-BS model parameters can be made by using the results given in (18) and (19).

4 Diagnostic analysis

4.1 Residual analysis

The objective of the residual analysis is to assess if the errors hold the distributional, heteroscedasticity and autocorrelation assumptions, as well the presence of atypical data. In classic regression models, Pearson and studentized residuals are often used. However, in tobit models, in general, even under normality, these type of residuals are inadequate Barros et al. (2010). For the tobit-BS model, we propose a deviance component (DC) and martingale-type (MT) residuals. These residuals are often used in generalized linear models and survival analysis, respectively.

In general, the DC residual is defined as (McCullagh and Nelder, 1989)

$$r_i^{\mathrm{DC}} = \mathrm{sign}(2\delta_i)\sqrt{2(\ell_i(\widehat{\boldsymbol{\theta}}_s) - \ell_i(\widehat{\boldsymbol{\theta}})}, \quad i = 1, \ldots, n,$$

where $\ell_i(\boldsymbol{\theta})$ represents the log-likelihood function for the case i, $\widehat{\boldsymbol{\theta}}_s$ is the ML estimate of $\boldsymbol{\theta}$ for the saturated model (a model with n parameters), $\boldsymbol{\theta}$ is the ML estimate of the model of interest, and $\widehat{\mu}_i$ is the expected value of Y_i, $\mathrm{E}(Y_i)$ say. Davison and Gigli (1989) specified the DC residual for censored data as

$$r_i^{\mathrm{DC}} = \mathrm{sign}(2\delta_i)\sqrt{-2\log(\widehat{S}_Y(y_i; \boldsymbol{\theta}))}, \quad i = 1, \ldots, n \tag{23}$$

where $\widehat{S}_Y$ is the corresponding SF of the model and the sign function is defined as $\text{sign}(x) = \{-1, 0, 1\}$, for $x < 0$, $x = 0$ and $x > 0$, respectively. For the tobit-BS model and from (23), the DC residual using is described as

$$r_i^{\text{DC}} = \begin{cases} \text{sign}(2\widehat{\delta}_i)\sqrt{-2\log\left(\Phi(-\frac{2}{\widehat{\alpha}}\sinh(\widehat{\delta}_i))\right)}, & i = 1, \ldots, m; \\ \text{sign}(2\widehat{\delta}_i)\sqrt{2\log\left(\cosh(\widehat{\delta}_i)\right) + \frac{2}{\widehat{\alpha}^2}(\sinh(\widehat{\delta}_i))^2}, & i = m+1, \ldots, n. \end{cases} \tag{24}$$

The MT residual in defined as

$$r_i^{\text{MT}} = \text{sign}(r_i^{\text{M}})\sqrt{-2(r_i^{\text{M}} + \eta_i \log(\eta_i - r_i^{\text{M}}))}, \quad i = 1, \ldots, n, \tag{25}$$

where r_i^{M} is the martingale residual given by

$$r_i^{\text{M}} = \eta_i + \log\left(\widehat{S}_Y(y_i)\right), \quad i = 1, \ldots, n, \tag{26}$$

with $\widehat{S}_Y$ being the ML estimate for the SF of Y (Klein and Moeschberger, 1997; Ortega et al., 2003). For the tobit-BS model and from (26), the martingale residual takes the form

$$r_i^{\text{M}} = \eta_i + \log(\Phi(\tfrac{2}{\widehat{\alpha}}\sinh(\widehat{\delta}_i))), \quad i = 1, \ldots, n, \tag{27}$$

where $\eta_i = 0$ indicates that the case i is censored, whereas $\eta_i = 1$ indicates that the case i is uncensored. Barros et al. (2010) noted that the MT residual does not represent the deviance component of the tobit model, instead it merely is a transformation of the martingale residual that produces symmetrically distributed residuals around zero. For more details, the interested reader can see Therneau et al. (1990), Leiva et al. (2007) and Barros et al. (2010).

4.2 Global influence

A global influence method commonly used in statistical modeling is the Cook distance, which allows us to assess the effect of each case on the estimated parameters. The methodology proposed by Cook (1977) and Cook and Weisberg (1982) suggests the deletion of each observation and the evaluation of the log-likelihood function without the case i. The generalized Cook distance (GCD) for a standard tobit model was given by Barros et al. (2010) and it takes the form

$$\text{GCD}_i(\boldsymbol{\theta}) = \frac{1}{p+1}(\widehat{\boldsymbol{\theta}} - \widehat{\boldsymbol{\theta}}_{(i)})\ddot{\boldsymbol{\ell}}^{-1}(\widehat{\boldsymbol{\theta}} - \widehat{\boldsymbol{\theta}}_{(i)}), \quad i = 1, \ldots, n, \tag{28}$$

where p is the number of model coefficients and $\widehat{\boldsymbol{\theta}}_i$ is the ML estimate of $\boldsymbol{\theta}$ without the case i. In order to facilitate the calculations, a first order approximation $\widehat{\boldsymbol{\theta}} - \widehat{\boldsymbol{\theta}}_i \approx \ddot{\boldsymbol{\ell}}_{(i)}^{-1}\dot{\boldsymbol{\ell}}_{(i)}$ is used in (28) and it becomes

$$\text{GCD}_i(\boldsymbol{\theta}) = \frac{1}{p+1}(\dot{\boldsymbol{\ell}}_{(i)}^{\top}\ddot{\boldsymbol{\ell}}_{(i)}^{-1}(-\ddot{\boldsymbol{\ell}})\ddot{\boldsymbol{\ell}}_{(i)}^{-1}\dot{\boldsymbol{\ell}}_{(i)}), \quad i = 1, \ldots, n, \tag{29}$$

where $\dot{\boldsymbol{\ell}}_{(i)}$ and $\ddot{\boldsymbol{\ell}}_{(i)}$ are the score vector and the Hessian matrix from the tobit-BS model defined in (17) and (19), respectively. Without the use of the approximation described in (29), we must calculate the GCD eliminating each case. However, with this approximation, we need to calculate the GCD just once. Usually, the diagnostics analysis relies on the vector β of coefficients. Then, in that case, (28) becomes

$$\mathrm{GCD}_i(\boldsymbol{\beta}) = \frac{1}{p}(\widehat{\boldsymbol{\beta}} - \widehat{\boldsymbol{\beta}}_{(i)})\ddot{\boldsymbol{\ell}}^{-1}(\widehat{\boldsymbol{\beta}} - \widehat{\boldsymbol{\beta}}_{(i)}), \quad i = 1, \ldots, n. \tag{30}$$

To determine whether the case i is potentially influential on $\boldsymbol{\beta}$, we use the same benchmark used by Zhu and Zhang (2004) and Barros et al. (2010), which is $2/n$. If the value of the GCD for the case i is greater than $2/n$, the case i is potentially influential on the estimated vector of parameters.

4.3 Local Influence

One way to assess the effect produced by one case on the ML estimates is the deletion of each of them from the data set and then evaluate if it exercises influence on the estimates or not. This approach is known as global influence. However, the local influence method relies on a geometric differentiation, taking the curvature of the plane of the log-likelihood function. This method does not require any deletion. Often the differential comparison is made before and after a perturbation. There are many ways to conduct a local influence analysis. We use case-weight, response and covariate perturbation schemes.

Recalling that $\boldsymbol{\theta} = (\alpha, \boldsymbol{\beta}^\top)^\top$ is the vector of parameter, let $\ell(\boldsymbol{\theta}|\boldsymbol{\omega})$ be the log-likelihood function of the model defined in (14) perturbed by $\boldsymbol{\omega}$, where $\boldsymbol{\omega}$ is a subset of $\Omega \in \mathbb{R}^n$. One way to evaluate the influence of a perturbation over the estimates of $\boldsymbol{\theta}$ is to use the likelihood displacement (LD), which is given by

$$\mathrm{LD}(\boldsymbol{\omega}) = 2(\ell(\widehat{\boldsymbol{\theta}}) - \ell(\widehat{\boldsymbol{\theta}}_{\boldsymbol{\omega}})), \tag{31}$$

where $\widehat{\boldsymbol{\theta}}_\omega$ is the ML estimate of $\boldsymbol{\theta}$ based on $\ell(\boldsymbol{\theta}|\boldsymbol{\omega})$. To detect what case exercises influence on $\mathrm{LD}(\boldsymbol{\omega})$ defined in (31), we use the normal curvature in the direction of the vector $\boldsymbol{l}$, with $||\boldsymbol{l}|| = 1$. The normal curvature given by Barros et al. (2010) can be expressed as

$$C_l(\boldsymbol{\theta}) = 2|\boldsymbol{l}^\top \boldsymbol{\Delta}^\top \ddot{\boldsymbol{\ell}}^{-1} \boldsymbol{\Delta} \boldsymbol{l}|, \tag{32}$$

where $\boldsymbol{\Delta}$ is a matrix of perturbations and $\ddot{\boldsymbol{\ell}}$ is the Hessian matrix given in (19). The matrix $\boldsymbol{\Delta}$ has elements

$$\partial^2 \ell(\theta|\boldsymbol{\omega})/\partial\theta_j \partial\omega_i, \quad j = 1, \ldots, p+1, i = 1, \ldots, n,$$

and must be evaluated at $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}$ and $\boldsymbol{\omega} = \boldsymbol{\omega}_0$, where $\boldsymbol{\omega}_0$ is a non-perturbation vector. To determine what case are influential under small perturbations, Barros et al. (2010) proposed an index plot based on the eigenvector of $\boldsymbol{l}_{\max}$, which can be constructed by using the maximum eigenvalue of

$$\boldsymbol{B}(\boldsymbol{\theta}) = |\boldsymbol{\Delta}^\top \ddot{\boldsymbol{\ell}}^{-1} \boldsymbol{\Delta}|, \tag{33}$$

evaluated at $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}$. On the one hand, if the interest relies just on the vector of parameter $\boldsymbol{\beta}$, (32) becomes

$$C_l(\boldsymbol{\theta}) = 2|\boldsymbol{l}^\top \boldsymbol{\Delta}^\top (\ddot{\boldsymbol{\ell}}^{-1} - \boldsymbol{B}_1) \boldsymbol{\Delta} \boldsymbol{l}|. \tag{34}$$

Note that (34) removes α from the analysis. Thus, the influence detection is only made on $\boldsymbol{\beta}$, where $\boldsymbol{B_1}$ under a tobit-BS model takes the form

$$\boldsymbol{B_1} = \begin{pmatrix} \mathrm{tr}(\boldsymbol{G})^{-1} & 0 \\ 0 & 0 \end{pmatrix}. \tag{35}$$

On the other hand, if the interest is just on α, (32) becomes

$$C_l(\boldsymbol{\theta}) = 2\left|\boldsymbol{l}^\top \boldsymbol{\Delta}^\top (\ddot{\boldsymbol{\ell}}^{-1} - \boldsymbol{B_2})\boldsymbol{\Delta}\boldsymbol{l}\right|, \tag{36}$$

where $\boldsymbol{B}_2$ for the tobit-BS model is given by

$$\boldsymbol{B_2} = \begin{pmatrix} 0 & 0 \\ 0 & (\boldsymbol{X}^\top \boldsymbol{V}\boldsymbol{X})^{-1} \end{pmatrix}, \tag{37}$$

with $\mathrm{tr}(\boldsymbol{G})$ and $(\boldsymbol{X}^\top \boldsymbol{V}\boldsymbol{X})^{-1}$ being obtained from the Hessian matrix of the tobit-BS model given in (19).

The maximum normal curvature vector, denoted by $\boldsymbol{l}_{\max}$, is an important direction to assess the local influence on $\widehat{\boldsymbol{\theta}}$, however it is not the only one. The vector $\boldsymbol{l}_i = \boldsymbol{e}_{in}$ denotes the direction of the case i, where $\boldsymbol{e}_{in}$ is the canonical basis of $\mathbb{R}^n$. It assumes zero for every case except the case i, which assumes the value one. The normal curvature of each observation is given by $C_i(\theta) = 2|b_{ii}|$, where b_{ii} represents the ith element of the matrix defined in (33) for each observation. A observation is potentially influential on $\boldsymbol{\theta}$ if $C_i(\widehat{\boldsymbol{\theta}}) > 2C(\widehat{\boldsymbol{\theta}})$, where $C(\widehat{\boldsymbol{\theta}})$ is the mean of the C_i's for $i = 1, \ldots, n$; see Lesaffre and Verbeke (1998). Below we present three perturbation schemes and each of their corresponding perturbation matrix $\boldsymbol{\Delta}$

Case-weight Perturbation The case-weight perturbation scheme allows us to evaluate the cases that, under different weights, affect the ML estimates of $\boldsymbol{\theta}$. Suppose that the log-likelihood function is given by $\ell(\boldsymbol{\theta}|\boldsymbol{\omega}) = \sum_{=1}^{n} \boldsymbol{\omega}_i \ell_i$, with ℓ_i given in (15) and $\boldsymbol{\omega}_i \in [0, 1]$ is a perturbation vector. Considering the partial derivative with respect to $\boldsymbol{\omega}^\top$, we obtain

$$\frac{\partial \ell(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \boldsymbol{\omega}^\top} = \sum_{i=1}^{m} \ell_i(\boldsymbol{\theta})\boldsymbol{e}_{in}^\top,$$

with $\boldsymbol{e}_{in}^\top$ denoting a $n \times 1$ vector.

After evaluating $\boldsymbol{\theta}$ at $\widehat{\boldsymbol{\theta}}$ we obtain a perturbation matrix

$$\boldsymbol{\Delta} = \sum_{i=1}^{n} \boldsymbol{h}_i \boldsymbol{e}_{in}^\top, \tag{38}$$

where $\boldsymbol{h}_i$ is given by

$$\boldsymbol{h}_i = \begin{pmatrix} \frac{\partial \ell_i(\boldsymbol{\theta})}{\partial \alpha} \\ \frac{\partial \ell_i(\boldsymbol{\theta})}{\partial \beta} \end{pmatrix}.$$

From the log-likelihood function defined in (15) we obtain a explicit expression for $\boldsymbol{h}_i$, which is

$$
\begin{aligned}
\dot{\boldsymbol{\ell}}_{\alpha_i}(\boldsymbol{\theta}|\boldsymbol{\omega}) &= a_i = \begin{cases} -\frac{\lambda(\zeta^{\mathrm{c}}_{i2})\zeta^{\mathrm{c}}_{i2}}{\alpha} & i = 1, \ldots, m, \\ -\frac{1}{\alpha}(1+\zeta^2_{i2}) & i = m+1, \ldots, n, \end{cases} \\
\dot{\boldsymbol{\ell}}_{\boldsymbol{\beta}_i}(\boldsymbol{\theta}|\boldsymbol{\omega}) &= b_i = \begin{cases} -\frac{\lambda(\zeta^{\mathrm{c}}_{i2})\zeta^{\mathrm{c}}_{i1}}{2} & i = 1, \ldots, m, \\ \frac{1}{2}\left(\zeta_{i1}\zeta_{i2} - \frac{\zeta_{i2}}{\zeta_{i1}}\right) & i = m+1, \ldots, n. \end{cases}
\end{aligned}
$$

The case-weight perturbation matrix given by $\boldsymbol{\Delta}$ in (38) can be decomposed in $\boldsymbol{\Delta}_\alpha = (a_1, \ldots, a_n)$ and $\boldsymbol{\Delta}_\beta = \boldsymbol{X}^\top \operatorname{diag}\{b_1, \ldots, b_n\}$.

Response perturbation There are different scenarios that one can consider for the response perturbation. We consider here a response perturbation scheme with an additive perturbation, which is defined by

$$Y_{i\boldsymbol{\omega}} = Y_i + \boldsymbol{\omega}_i S_Y,\ i = m+1, \ldots, n,$$

where S_Y is a scale component that can be the standard deviation of the response. Let us consider that the log-likelihood function of the tobit-BS model is given by $\ell(\boldsymbol{\theta}|\boldsymbol{\omega}) = \sum_{i=1}^n \ell_i(\boldsymbol{\theta}|\boldsymbol{\omega})$, with

$$\ell_i(\boldsymbol{\theta}) = \begin{cases} \sum_{i=1}^m (\log \Phi(\zeta^{\mathrm{c}}_{i2\omega1})), & i = 1, \ldots, m; \\ -\log(2) - \frac{\log(2\pi)}{2} + \sum_{i=m+1}^n \left(\log(\zeta_{i1\omega1}) - \frac{1}{2}(\zeta^2_{i2\omega1})\right), & i = m+1, \ldots, n; \end{cases} \tag{39}$$

where $\zeta_{i1\omega1}, \zeta_{i2\omega1}, \zeta^{\mathrm{c}}_{i2\omega1}$ are as defined in (16) after changing Y for $Y_{i\omega}$. The process of obtaining the perturbation matrix in this scheme is composed by two steps, first the derivative of (39) with respect to $\boldsymbol{\omega}^\top$ is taken, i.e.,

$$\frac{\partial \ell_i(\boldsymbol{\theta}|\boldsymbol{\omega}_i)}{\partial \boldsymbol{\omega}^\top} = \begin{cases} \mathbf{0}_n^\top, & i = 1, \ldots, m, \\ \frac{S_Y}{2}\left(\frac{\zeta_{i2\omega1}}{\zeta_{i1\omega1}}\right) - \zeta_{i2\omega1}\zeta_{i1\omega1}, & i = m+1, \ldots, n. \end{cases} \tag{40}$$

Then, the partial derivative of (40) with respect to θ is computed and evaluated at $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}$, to obtain the response perturbation matrix $\boldsymbol{\Delta}$ with elements

$$\boldsymbol{\Delta}_\alpha = (c_1, \ldots, c_m), \quad \boldsymbol{\Delta}_{\boldsymbol{\beta}} = \boldsymbol{X}^\top \operatorname{diag}\{d_1, \ldots, d_m\},$$

where

$$c_i = \frac{S_Y \zeta_{i1}\zeta_{i2}}{\alpha}, \quad d_i = S_Y\left(\frac{1}{\alpha^2}\cosh(2\delta_i) - \frac{1}{4}\left(\operatorname{sech}(\delta_i)\right)^2\right).$$

Note that the response perturbation scheme, in a tobit model, makes sense only for the non-censored part of the data. This occurs because the censored part of the data is either unobservable or below the threshold y_0, otherwise, the case i receives the same value y_0. Then, there is no perturbation in this part of the data.

Covariate Perturbation As in the response scheme there are several ways to insert covariate perturbation. Here, we insert an additive perturbation that takes the form

$$\boldsymbol{x}_{it\boldsymbol{\omega}} = \boldsymbol{x}_{it} + s_X \boldsymbol{\omega}_i, \quad i = 1, \ldots, n,$$

where S_X can be the standard deviation of the correspond covariate of the model. Considering the log-likelihood function for the tobit-BS model to be $\ell(\boldsymbol{\theta}|\boldsymbol{\omega}) = \sum_{i=1}^{n} \ell_i(\boldsymbol{\theta}|\boldsymbol{\omega})$ from (15), we obtain

$$\ell_i(\boldsymbol{\theta}|\boldsymbol{\omega}) = \begin{cases} \sum_{i=1}^{m} (\log \Phi(\zeta^{\mathrm{c}}_{i2\boldsymbol{\omega}2})) & i = 1, \ldots, m; \\ -\log(2) - \frac{\log(2\pi)}{2} + \sum_{i=m+1}^{n} \left(\log(\zeta_{i1\boldsymbol{\omega}2}) - \frac{1}{2}(\zeta^2_{i2\boldsymbol{\omega}2})\right) & i = m+1, \ldots, n. \end{cases} \tag{41}$$

In order to obtain the perturbation matrix of (41) we take the derivative with respect to $\boldsymbol{\theta}$ and then with respect to $\boldsymbol{\omega}$, and then we evaluate at $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}$ to obtain

$$\begin{aligned} \boldsymbol{\Delta}_\alpha &= \begin{cases} -\frac{S_x \boldsymbol{\beta}_t}{\alpha} \left(\frac{\zeta^{\mathrm{c}}_{i1} \lambda(\zeta^{\mathrm{c}}_{i2})}{2} - \zeta^{\mathrm{c}}_{i2} \left(\frac{2\lambda(\zeta^{\mathrm{c}}_{i2}) \sinh(2\delta_i)}{\alpha^2} + \frac{\lambda^2 \zeta^{\mathrm{c}}_{i2}}{2}\right)\right), & i = 1, \ldots, m; \\ -\frac{2}{\alpha^3} S_x \boldsymbol{\beta}_t \sinh(2\delta_i), & i = m+1, \ldots, n \end{cases} \\ \boldsymbol{\Delta}_{\boldsymbol{\beta}} &= \begin{cases} -\frac{S_x \boldsymbol{\beta}_t x_{ij}}{4} \left(\frac{\lambda(\zeta^{\mathrm{c}}_{i2})(\zeta^{\mathrm{c}}_{i1})^2 \zeta^{\mathrm{c}}_{i2}}{\Phi(\zeta^{\mathrm{c}}_{i2})} + (\zeta^{\mathrm{c}}_{i2})^2 \lambda^2(\zeta^{\mathrm{c}}_{i2}) - \lambda(\zeta^{\mathrm{c}}_{i2}) \zeta^{\mathrm{c}}_{i2}\right) & \\ -\frac{S_x \zeta^{\mathrm{c}}_{i2} \lambda(\zeta^{\mathrm{c}}_{i2})}{2}, & i = 1, \ldots, m; \\ S_x \boldsymbol{\beta}_t x_{ij} \left(\frac{1}{4} (\operatorname{sech}(\delta_i))^2 - \frac{1}{\alpha^2} \cosh(2\delta_i)\right), & i = m+1, \ldots, n, \end{cases} \end{aligned}$$

where $\boldsymbol{\Delta}_{\boldsymbol{\beta}}$ is a matrix $p \times n$ with elements $\Delta_{\boldsymbol{\beta}_{ij}}$ and $j \neq t$, for $j = t$, $\Delta_{\boldsymbol{\beta}_{ij}}$ becomes

$$\boldsymbol{\Delta}_{\boldsymbol{\beta}} = \begin{cases} -\frac{S_x \boldsymbol{\beta}_t x_{ij}}{4} \left(\frac{\lambda(\zeta^{\mathrm{c}}_{i2})(\zeta^{\mathrm{c}}_{i1})^2 \zeta^{\mathrm{c}}_{i2}}{\Phi(\zeta^{\mathrm{c}}_{i2})} + (\zeta^{\mathrm{c}}_{i2})^2 \lambda^2(\zeta^{\mathrm{c}}_{i2}) - \lambda(\zeta^{\mathrm{c}}_{i2}) \zeta^{\mathrm{c}}_{i2}\right) - \frac{S_x \zeta^{\mathrm{c}}_{i2} \lambda(\zeta^{\mathrm{c}}_{i2})}{2} & \\ -\frac{S_x x_{ij}}{\alpha} \left(\cosh(2\delta_i) \lambda(\zeta^{\mathrm{c}}_{i2})\right), & i = 1, \ldots, m; \\ S_x \boldsymbol{\beta}_t x_{ij} \left(\frac{1}{4} (\operatorname{sech}(\delta_i))^2 - \frac{1}{\alpha^2} \cosh(2\delta_i)\right) & \\ +S_x \left(\frac{1}{\alpha^2} \sinh(2\delta_i) - \frac{1}{2} \tanh(\delta_i)\right), & i = m+1, \ldots, n. \end{cases}$$

5 Concluding remarks

It is widely known that the normal distribution does not model perfectly a considerable part of real data sets. In this paper, we have proposed an extension of the standard tobit model, the tobit-BS model, by assuming a Birnbaum-Saunders distribution. We have discussed maximum likelihood estimation of the model parameters. We have also provided global and local influence tools for the tobit-BS model. In special, as a global influence tool we have derived the generalized Cook distance, and for local influence procedures we have derived the normal curvatures under different perturbation schemes. Research regarding the implementation of Monte Carlo simulations and the application to real data is currently in progress and we hope to report some findings in a future paper.

References

Amemiya, T. (1984). Tobit models: A survey. *Journal of Econometrics*, 24:3–61.

Barros, M., Galea, M., Gonzalez, M., and Leiva, V. (2010). Influence diagnostics in the Tobit censored response model. *Statistical Methods and Applications*, 19:379–397.

Birnbaum, Z. W. and Saunders, S. C. (1969). A new family of life distributions. *Journal of Applied Probability*, 6:319–327.

Cook, R. D. (1977). Detection of influential observation in linear regression. *Technometrics*, 19:15–18.

Cook, R. D. and Weisberg, S. (1982). *Residuals and Influence in Regression*. Chapman and Hall, London, UK.

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

Davison, A. and Gigli, A. (1989). Deviance residuals and normal scores plots. *Biometrika*, 76:211–221.

Efron, B. and Hinkley, D. (1978). Assessing the accuracy of the maximum likelihood estimator: Observed vs. expected Fisher information. *Biometrika*, 65:457–487.

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.

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 `http://dx.doi.org/10.1007/s00477--015--1204--4`.

Goldberger, A. (1964). *Econometric Theory*. Wiley, New York, US.

Helsel, D. R. (2011). *Statistics for Censored Environmental Data Using Minitab and R*. Wiley, New York, US.

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

Klein, J. and Moeschberger, M. (1997). *Survival Analysis: Techniques for Censored and Truncated Data*. Springer, New York, US.

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

Leiva, V., Barros, M., Paula, G., and Galea, M. (2007). Influence diagnostics in log-Birnbaum-Saunders regression models with censored data. *Computational Statistics and Data Analysis*, 51:5694–5707.

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.

Leiva, V., Rojas, E., Galea, M., and Sanhueza, A. (2014). Diagnostics in Birnbaum-Saunders accelerated life models with an application to fatigue data. *Applied Stochastic Models in Business and Industry*, 30:115–131.

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.

Lesaffre, E. and Verbeke, G. (1998). Local influence in linear mixed models. *Biometrics*, 54:570–582.

Li, A., Chen, Z., and Xie, F. (2012). Diagnostic analysis for heterogeneous log-Birnbaum-Saunders regression models. *Statistical and Probability Letters*, 89:1690–1698.

Marchant, C., Leiva, V., and Cysneiros, F. (2016a). A multivariate log-linear model for Birnbaum-Saunders distributions. *IEEE Transactions on Reliability*, page in press available at `http://dx.doi.org/10.1109/TR.2015.2499964`.

Marchant, C., Leiva, V., Cysneiros, F., and Vivanco, J. (2016b). Diagnostics in multivariate generalized Birnbaum-Saunders regression models. *Journal of Applied Statistics*, page in press available at `http://dx.doi.org/10.1080/02664763.2016.1148671`.

McCullagh, P. and Nelder, J. (1989). *Generalized Linear Models*. Chapman and Hall, London, UK.

Ortega, E., Bolfarine, H., and Paula, G. (2003). Influence diagnostics in generalized log-gamma regression models. *Computational Statistics and Data Analysis*, 42:165–186.

Qu, H. and Xie, F. (2011). Diagnostics analysis for log-Birnbaum-Saunders regression models with censored data. *Statistica Neerlandica*, 65:1–21.

Rieck, J. and Nedelman, J. (1991). A log-linear model for the Birnbaum-Saunders distribution. *Technometrics*, 3:51–60.

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.

Scott, J. (1997). *Regression Models of Categorical and Limited Dependent Variables*. Sage, California, US.

Therneau, T., Grambsch, P., and Fleming, T. (1990). Martingale-based residuals for survival models. *Biometrika*, 77:147–160.

Thorarinsdottir, T. L. and Gneiting, T. (2010). Probabilistic forecasts of wind speed: Ensemble model output statistics by using heteroscedastic censored regression. *Journal of the Royal Statistical Society A*, 173:371–388.

Tobin, J. (1958). Estimation of relationships for limited dependent variables. *Econometrica*, 26:24–36.

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

Zhu, H. and Zhang, H. (2004). A diagnostic procedure based on local influence. *Biometrika*, 91:579–589.