



Scale-mixture Birnbaum-Saunders quantile regression models applied to personal accident insurance data

Alan Dasilva¹ · Helton Saulo^{2,3} · Roberto Vila^{2,4} · Suvra Pal³

Received: 9 July 2024 / Revised: 23 November 2024 / Accepted: 26 November 2024 /
Published online: 14 December 2024

© The Author(s) under exclusive licence to Sociedade Brasileira de Matemática Aplicada e Computacional 2024

Abstract

The modeling of personal accident insurance data has been a topic of high relevance in the insurance literature. This type of data often exhibits positive skewness and heavy tails. In this work, we propose a new quantile regression model based on the scale-mixture Birnbaum-Saunders distribution for modeling personal accident insurance data. The maximum likelihood estimates of the model parameters are obtained via the EM algorithm. Two Monte Carlo simulation studies are performed using the R software. The first study aims to analyze the performances of the EM algorithm to obtain the maximum likelihood estimates, and the randomized quantile and generalized Cox-Snell residuals. In the second simulation study, the size and power of the Wald, likelihood ratio, score and gradient tests are evaluated. The two simulation studies are conducted considering different quantiles of interest and sample sizes. Finally, a real insurance data set is analyzed to illustrate the proposed approach.

Keywords Scale-mixture Birnbaum-Saunders distribution · EM algorithm · Hypothesis tests · Monte Carlo simulation · Quantile regression

1 Introduction

Birnbaum-Saunders (BS) regression models have proven to be good alternatives in insurance data modeling; see Leiva (2016) and Naderi et al. (2020). Usually, this kind of data shows positive skewness and heavy tails. Paula et al. (2012), for example, used the BS-Student- t -BS regression model to study how some explanatory variables (covariates) influence the amount of paid money by an insurance policy. The Student- t -BS distribution is a special case of the class of scale-mixture Birnbaum-Saunders (SBS) distributions, which was proposed by Balakrishnan et al. (2009) and is based on the relationship between the class of scale mixtures of normal (SMN) and BS distributions. The main advantage of the SBS distributions over the classical BS distribution (Leiva (2016); Balakrishnan and Kundu (2019)) is their

✉ Helton Saulo
heltonsaulo@gmail.com

¹ Institute of Mathematics and Statistics, University of Sao Paulo, Sao Paulo, Brazil

² Department of Statistics, University of Brasilia, Brasilia, Brazil

³ Department of Mathematics, University of Texas at Arlington, Arlington, TX, USA

⁴ Department of Mathematics and Statistics, McMaster University, Hamilton, Ontario, Canada

capability to make robust estimation of parameters in a way similar to that of SMN models, in addition to facilitating the implementation of an expectation maximization (EM) algorithm; see Balakrishnan et al. (2009). Recently, Lachos et al. (2017) proposed a Bayesian regression model based on the SBS distributions for censored data.

The class of SMN distributions comprises a family of symmetric distributions studied initially by Andrews and Mallows (1974). These distributions have been gaining considerable attention with the publication of some works, as can be seen in Efron and Olshen (1978), West (1987), Lange and Sinsheimer (1993), Gneiting (1997), Taylor and Verbyla (2004), Walker and Gutiérrez-Peña (2007) and Lachos and Labra (2007). This family of distributions has as special cases flexible distributions with heavy-tails, which are often employed for robust estimation; see Lange et al. (1989) and Lucas (1997).

Quantile regression models emerged as a broader proposal in the analysis of the relationship between the response variable and the respective covariates of a regression model. These models were initially studied by Koenker and Bassett (1978) and have since been increasingly used in regression analysis; see Koenker and Hallock (2001), Yu and Moyeed (2001) and Koenker (2004). Quantile regression models are robust alternatives in relation to the regression models for mean because the process of estimating parameters of the quantile regression model is not subject to the influence of outliers as in the case of linear regression. Moreover, these models are capable of showing different effects of covariates on the response along the quantiles of response variable.

In this work, we propose a new parametric quantile regression model for strictly positive data based on a reparameterization of the SBS distributions. We first introduce a reparameterization of the SBS model by inserting a quantile parameter, and then develop the new SBS quantile regression model. In addition to the quantile approach, another advantage in relation to the regression approach developed in Lachos et al. (2017) is the modeling of dependent variable without the need of applying the logarithm transformation, which can cause problems with regard to interpretation and loss in power of the study; see Huang and Qu (2006) for an explanation. We illustrate the proposed methodology by using the insurance data set studied by Paula et al. (2012). The results show that a quantile approach provides a richer characterization of the effects of covariates on the dependent variable.

The rest of this paper is organized as follows. In Sect. 2, we describe the usual SBS distribution and propose a reparameterization of this distribution in terms of a quantile parameter. In this section, we also discuss some mathematical properties of the proposed reparameterized model. In Sect. 3, we introduce the SBS quantile regression model. In this section, we also describe the steps of the EM algorithm (Dempster et al. (1977)) for maximum likelihood (ML) estimation of the model parameters, the standard error calculation, hypothesis tests, model selection criteria and residuals. In Sect. 4, we carry out two Monte Carlo simulation studies. In the first simulation study, we evaluate the performance of the EM algorithm through the calculated biases, mean square errors (MSEs) and coverage probabilities of the asymptotic confidence intervals. We also evaluate the performances of residuals, such as the generalized Cox-Snell –GCS– (Cox and Snell (1968)) and randomized quantile –RQ– residuals (Dunn and Smyth (1996)), for assessing the goodness-of-fit and identifying any departure from model assumptions. In the second simulation study, we assess the performances of formal test procedures such as the likelihood ratio test, score test, Wald test and gradient test. In Sect. 5, we apply the SBS quantile regression model to a real insurance data set. Finally, in Sect. 6, we make some concluding remarks and discuss potential future research problems.

2 Preliminaries

In this section, we first present the SBS distribution. Then, we present the proposed quantile-based SBS distribution, which we call as QSBS distribution. Some properties and special cases of the QSBS distribution are also discussed.

2.1 SBS distribution

Let Y be a random variable (RV) whose stochastic representation can be written as follows:

$$Y = \mu + \sqrt{g(U)} \cdot X, \quad (1)$$

where μ is a location parameter, $X \sim N(0, \sigma^2)$ with $\sigma > 0$, U is a RV distributed independent of X and having cumulative distribution function (CDF) $H(\cdot)$, which is indexed by an extra parameter ν (or extra parameter vector $\mathbf{v}$), and $g(\cdot)$ is a strictly positive function. Note that when $g(U) = 1/U$, the distribution of Y in (1) reduces to the normal/independent distribution, presented in Lange and Sinsheimer (1993). Similarly, when $g(U) = U$ in (1), the distribution of Y reduces to the SMN distribution, studied by Fernández and Steel (1999).

The probability density function (PDF) of Y , a RV that follows a SMN distribution with location and scale parameters $\mu \in \mathbb{R}$ and $\sigma^2 > 0$, respectively, is given by the following Lebesgue-Stieltjes integral:

$$\phi_{\text{SMN}}(y) = \int_{(0, \infty) \cap \text{Supp}(U)} \phi(y; \mu, g(u)\sigma^2) dH(u), \quad y \in \mathbb{R}, \quad (2)$$

where $\text{Supp}(U)$ is the support of U , $\phi(\cdot; \mu, g(u)\sigma^2)$ is the normal PDF with mean μ and variance $g(u)\sigma^2$, and $H(u)$ is the CDF of U . Let us denote $Y \sim \text{SMN}(\mu, \sigma^2; H)$. When $\mu = 0$ and $\sigma^2 = 1$, we denote $Y \sim \text{SMN}(H)$.

According to Balakrishnan et al. (2009), a RV T follows a SBS distribution if it has the following stochastic representation:

$$T = \frac{\beta}{4} \left[\alpha Y + \sqrt{(\alpha Y)^2 + 4} \right]^2,$$

where $Y = \sqrt{g(U)} \cdot Z \sim \text{SMN}(H)$, such that $Z \sim N(0, 1)$ and $\alpha > 0$ and $\beta > 0$ are the shape and scale parameters, respectively, of the BS distribution. In this case, the notation $T \sim \text{SBS}(\alpha, \beta; H)$ is used. When $Y \sim N(0, 1)$, then T follows a classical BS distribution. The PDF of $T \sim \text{SBS}(\alpha, \beta; H)$ can be expressed as

$$f_T(t) \phi_{\text{SMN}}(a(t)) a'(t), \quad t > 0,$$

where $\phi_{\text{SMN}}(\cdot)$ is the PDF given in (2) with $\mu = 0$ and $\sigma^2 = 1$, $a(t) = (\sqrt{t/\beta} - \sqrt{\beta/t})/\alpha$, and $a'(t) = t^{-3/2}(t + \beta)/(2\alpha\beta^{1/2})$ is the derivative of $a(t)$ with respect to t . The CDF of T is given by

$$F_T(t) = \Phi_{\text{SMN}}(a(t)), \quad t > 0,$$

where $\Phi_{\text{SMN}}(\cdot)$ is the CDF of the $\text{SMN}(H)$ distribution. The 100 q -th quantile of $T \sim \text{SBS}(\alpha, \beta; H)$ is given by

$$Q = t_q = \frac{\beta}{4} \left[\alpha y_q + \sqrt{(\alpha y_q)^2 + 4} \right]^2, \quad (3)$$

where y_q is the $q \times 100$ -th quantile of $Y \sim \text{SMN}(H)$.

2.2 Quantile-based SBS distributions

Consider a fixed number $q \in (0, 1)$ and the one-to-one transformation $(\alpha, \beta; H) \mapsto (\alpha, Q; H)$, where Q is the $100q$ -th quantile of $T \sim \text{SBS}(\alpha, \beta; H)$ defined in (3). Then, we obtain the following stochastic representation based on a quantile parameter:

$$T = \frac{Q}{\gamma_\alpha^2} \left[\alpha \sqrt{g(U)} \cdot Z + \sqrt{(\alpha \sqrt{g(U)} \cdot Z)^2 + 4} \right]^2, \quad (4)$$

where $\gamma_\alpha = \alpha y_q + \sqrt{(\alpha y_q)^2 + 4}$, $Z \sim N(0, 1)$, and $g(U) = 1/U$ is such that the distribution of $g(U)$ has a known PDF. The stochastic representation (4) can be used to generate random numbers, to obtain the moments, and also in the implementation of the EM algorithm.

From the stochastic representation (4), we obtain a quantile-based reparameterization of the SBS distribution with CDF and PDF given, respectively, by

$$\begin{aligned} F_T(t) &= \Phi_{\text{SMN}}(a(t)) = \int_{(0, \infty) \cap \text{Supp}(U)} \Phi(a \sqrt{g(u)}(t)) dH(u), \quad t > 0, \\ f_T(t) &= \phi_{\text{SMN}}(a(t)) a'(t) = \int_{(0, \infty) \cap \text{Supp}(U)} \phi(a \sqrt{g(u)}(t)) a' \sqrt{g(u)}(t) dH(u), \quad t > 0, \end{aligned} \quad (5)$$

where $\Phi(\cdot)$ and $\phi(\cdot)$ are the CDF and PDF of the standard normal distribution, respectively, $a_x(t) = (\sqrt{\gamma_{\alpha x}^2 t / 4Q} - \sqrt{4Q / \gamma_{\alpha x}^2 t}) / (\alpha x)$, $a'_x(t) = (\gamma_{\alpha x}^2 / 2 + 2Q/t) / (\alpha x \gamma_{\alpha x} \sqrt{4Qt})$ and $\gamma_\alpha = \alpha y_q + \sqrt{(\alpha y_q)^2 + 4}$. In this case, the notation $T \sim \text{QSBS}(\alpha, Q; H)$ is used. Some properties of the QSBS distribution are presented below.

Proposition 1 If $T \sim \text{QSBS}(\alpha, Q; H)$ then $\lim_{t \rightarrow 0^+} f_T(t) = \lim_{t \rightarrow \infty} f_T(t) = 0$.

Proof Since $f_T(\cdot)$ is a PDF, it is clear that $\lim_{t \rightarrow \infty} f_T(t) = 0$. On the other hand, by (5) we observe that $f_T(t) = \int_{(0, \infty) \cap \text{Supp}(U)} f_{S_u}(t) dH(u)$, where $f_{S_u}(\cdot)$ is the corresponding PDF of a RV S_u having a quantile-based Birnbaum-Saunders (QBS) distribution with parameters $\alpha \sqrt{u}$ and Q , denoted by $S_u \sim \text{QBS}(\alpha \sqrt{g(u)}, Q)$. It is well-known that the QBS distribution is unimodal. Then there is a unique $t_0 = t_0(\alpha \sqrt{g(u)}, Q) > 0$ such that $f_{S_u}(t) \leq \max\{f_{S_u}(t) : t > 0\} = f_{S_u}(t_0) \leq \max\{f_{S_u}(t_0) : u > 0\}$. Applying the bounded convergence theorem, we have $\lim_{t \rightarrow 0^+} f_T(t) = \int_{(0, \infty) \cap \text{Supp}(U)} \lim_{t \rightarrow 0^+} f_{S_u}(t) dH(u) = 0$ because $\lim_{t \rightarrow 0^+} f_{S_u}(t) = 0$. This completes the proof. $\square$

Proposition 2 If $g(\cdot)$ in (1) is chosen such that the function

$$G(t) = \frac{\phi'_{\text{SMN}}(a(t))}{\phi_{\text{SMN}}(a(t))} + \frac{a''(t)}{[a'(t)]^2}, \quad t > 0,$$

has a unique zero $t = t_0$, then the QSBS PDF (5) is unimodal.

Proof Suppose that $g(\cdot)$ is such that $G(t_0) = 0$. A simple calculus shows that $f'_T(t) = f_T(t) a'(t) G(t)$. Then $f'_T(t_0) = 0$ because $G(t_0) = 0$, $f_T(t) > 0$ and $a'(t) > 0$. That is, t_0 is the unique critical point of $f_T(\cdot)$. Furthermore, by Proposition 1, $\lim_{t \rightarrow 0^+} f_T(t) = \lim_{t \rightarrow \infty} f_T(t) = 0$. Consequently, the QSBS PDF is increasing on $(0, t_0)$ and is decreasing on (t_0, ∞) . This proves the unimodality. $\square$

In the proofs of the following propositions, we adopt the following notations: $a_x(t) = a_x(t; \alpha, Q)$ and $a(t) = a(t; \alpha, Q) = a_1(t)$.

Proposition 3 Let $T \sim \text{QSBS}(\alpha, Q; H)$. Then, T allows the following conditional stochastic representation:

$$T|U = u \sim \text{QBS}(\alpha\sqrt{g(u)}, Q).$$

Proof If $U = u$ then $T = a^{-1}(\sqrt{g(u)}Z)$. Thus, the conditional distribution of T given $U = u$ is the same as the distribution of the elements presented as $a^{-1}(\sqrt{g(u)}Z)|U = u$. This implies that $F_T(t|U = u) = \Phi(a(t)/\sqrt{g(u)}) = \Phi(a_{\sqrt{g(u)}}(t))$ because U is independent of Z . Then, the proof follows. $\square$

Let $T \sim \text{QSBS}(\alpha, Q; H)$ and f be a Borel measurable function. Then, by using the law of total expectation, we have

$$\mathbb{E}[f(T)] = \mathbb{E}[\mathbb{E}[f(T)|U]] = \int_{(0, \infty) \cap \text{Supp}(U)} \mathbb{E}[f(T)|U = u] dH(u). \quad (6)$$

By taking $f(t) = \exp(ist)$, $s \in \mathbb{R}$, in (6), where $i = \sqrt{-1}$ is the unit imaginary number, from Proposition 3 and Formula (2.13) of Leiva (2016), we get the following result:

Proposition 4 Let $T \sim \text{QSBS}(\alpha, Q; H)$. Then, the characteristic function of T , denoted by $\phi_T(s) = \mathbb{E}(\exp(isT))$, can be written as

$$\phi_T(s) = \frac{1}{2} \mathbb{E} \left[\left(1 + \frac{1}{\sqrt{1 - 2i(4Q/\gamma_\alpha^2)s^2g(U)}} \right) \exp \left(\frac{1 - \sqrt{1 - 2i(4Q/\gamma_\alpha^2)s^2g(U)}}{\alpha^2g(U)} \right) \right].$$

Taking k consecutive partial derivatives of $\phi_T(s)$ with respect to s , assuming that we can interchange the derivative with the expectation, and then evaluating at $s = 0$ and dividing by i^k , we get the following result:

Proposition 5 Let $T \sim \text{QSBS}(\alpha, Q; H)$ and $g(U)$ be a RV as in (1) with finite moments of all order. Then, the k -th moment of T has the following form:

$$\mathbb{E}[T^k] = \left(\frac{4Q}{\gamma_\alpha^2} \right)^k \sum_{l=0}^k \binom{2k}{2l} \sum_{j=0}^l \binom{l}{j} \omega_{k+j-l} \left(\frac{\alpha}{2} \right)^{2(k+j-l)},$$

where $\omega_r = \mathbb{E}[g^r(U)]$ denotes the r -th moment of $g(U)$.

Another way to obtain the expression of the above proposition is to consider $f(t) = t^k$ in (6) and by using Proposition 3 and Formula (2.18) of Leiva (2016).

The following two results prove that the QSBS distribution belongs to the family of scale and reciprocal invariant distributions:

Proposition 6 Let $T \sim \text{QSBS}(\alpha, Q; H)$. Then, $cT \sim \text{QSBS}(\alpha, cQ; H)$ with $c > 0$.

Proof It is simple to see that $a(t/c; \alpha, Q) = a(t; \alpha, cQ)$. Then, by using (5), we have $\mathbb{P}(cT \leq t) = \Phi_{\text{SMN}}(a(t/c; \alpha, Q)) = \Phi_{\text{SMN}}(a(t; \alpha, cQ))$ because $c > 0$. This guarantees the result. $\square$

Proposition 7 Let $T \sim \text{QSBS}(\alpha, Q; H)$. Then, $1/T \sim \text{QSBS}(\alpha, \gamma_\alpha^4/16Q; H)$.

Proof A simple observation shows that $a(1/t; \alpha, Q) = -a(t; \alpha, \gamma_\alpha^4/16Q)$. Then, by using (5) together with the identity $\phi_{\text{SMN}}(-y) = 1 - \phi_{\text{SMN}}(y)$, we get $\mathbb{P}(1/T \leq t) = 1 - \Phi_{\text{SMN}}(a(1/t; \alpha, Q)) = 1 - \Phi_{\text{SMN}}(-a(t; \alpha, \gamma_\alpha^4/16Q)) = \Phi_{\text{SMN}}(a(t; \alpha, \gamma_\alpha^4/16Q))$. Hence, the proof follows. $\square$

From Proposition 7, we get the following formula for the negative moments of the QSBS distribution:

Proposition 8 Let $T \sim \text{QSBS}(\alpha, Q; H)$ and $g(U)$ be a RV as in (1) with finite moments of all order. Then, the k -th negative moment of T has the following form:

$$\mathbb{E}[T^{-k}] = \left(\frac{\gamma_\alpha^2}{4Q}\right)^k \sum_{l=0}^k \binom{2k}{2l} \sum_{j=0}^l \binom{l}{j} \omega_{k+j-l} \left(\frac{\alpha}{2}\right)^{2(k+j-l)},$$

where $\omega_r = \mathbb{E}[g^r(U)]$ denotes the r -th moment of $g(U)$.

By using Proposition 5, if $T \sim \text{QSBS}(\alpha, Q; H)$, then the expectation, variance and coefficients of variation (CV), skewness (CS) and kurtosis (CK) are given, respectively, by

$$\begin{aligned} \mathbb{E}[T] &= \frac{2Q}{\gamma_\alpha^2} (2 + \omega_1 \alpha^2), \quad \text{Var}[T] = \frac{4Q^2 \alpha^2}{\gamma_\alpha^4} [\omega_1 + (2\omega_2 - \omega_1^2) \alpha^2], \\ \text{CV}[T] &= \frac{\alpha \sqrt{4\omega_1 + (2\omega_2 - \omega_1^2) \alpha^2}}{2 + \omega_1 \alpha^2}, \\ \text{CS}[T] &= \frac{4\alpha [3(\omega_2 - \omega_1^2) + (2\omega_3 - 3\omega_1 \omega_2 + \omega_1^3) \alpha^2 / 2]}{[4\omega_1 + (2\omega_2 - \omega_1^2) \alpha^2]^{3/2}}, \\ \text{CK}[T] &= \frac{16\omega_2 + (32\omega_3 - 48\omega_1 \omega_2 + 24\omega_1^3) \alpha^2 + (8\omega_4 - 16\omega_1 \omega_3 + 12\omega_1^2 \omega_2 - \omega_1^4) \alpha^4}{[4\omega_1 + (2\omega_2 - \omega_1^2) \alpha^2]^2}, \end{aligned}$$

where $\omega_r = \mathbb{E}[g^r(U)]$ denotes the r -th moment of $g(U)$.

2.3 Special cases of the QSBS distribution

In this section, we present some particular cases of the QSBS family of distributions, i.e., QSBS models based on the contaminated-normal, slash, and Student- t distributions.

Let $T \sim \text{QSBS}(\alpha, Q; H)$, where H is the CDF of the variable U , whose PDF (or probability mass function) is defined as h_U . By using (2) and by considering $g(u) = 1/u$, for determined functions h_U , the following formulas (Table 1) for the PDF ϕ_{SMN} are obtained. In Table 1, δ_{ij} denotes the delta de Kronecker, $\mathbb{1}_A$ is the indicator function of an event A , $\phi(\cdot)$ is the standard normal PDF and $\Gamma(\cdot)$ is the complete gamma function. For the rest of this paper, we consider $g(u) = 1/u$.

2.3.1 Quantile contaminated-normal BS distribution

Let $T \sim \text{QSBS}(\alpha, Q; H)$ and h_U be as given in Table 1. By using (5) and ϕ_{SMN} in Table 1, it follows that the PDF of T can be expressed as (for $t > 0$)

$$\begin{aligned} f_T(t) &= \left[v \sqrt{\delta} \phi \left(\frac{\sqrt{\delta}}{\alpha \gamma_\alpha} \sqrt{\frac{4Q}{t}} \left(\frac{\gamma_\alpha^2 t}{4Q} - 1 \right) \right) + (1 - v) \phi \left(\frac{1}{\alpha \gamma_\alpha} \sqrt{\frac{4Q}{t}} \left(\frac{\gamma_\alpha^2 t}{4Q} - 1 \right) \right) \right] \\ &\quad \times \frac{1}{\alpha \gamma_\alpha \sqrt{4Qt}} \left(\frac{\gamma_\alpha^2}{2} + \frac{2Q}{t} \right). \end{aligned}$$

Let us denote $T \sim \text{CN-BS}(\alpha, Q, \mathbf{v})$ with $\mathbf{v} = (v, \delta)$. The PDF of the conditional distribution of $U|T = t$ is given by

$$h_{U|T}(u|t) = v p(t, u) \delta_{u\delta} + (1 - v) p(t, u) \delta_{u1},$$

Table 1 Examples of functions h_U and ϕ_{SMN} associated with some QSBS distributions with $g(u) = 1/u$

Distribution	$h_U(u)$	$\phi_{SMN}(y)$	Parameter
Quantile contaminated-normal BS	$[\nu\delta_{u\delta} + (1 - \nu)\delta_{u1}]\mathbb{1}_{(0,1)}(u)$	$[\nu\sqrt{\delta}\phi(\sqrt{\delta}y) + (1 - \nu)\phi(y)]\mathbb{1}_{(-\infty,\infty)}(y)$	$0 < \nu < 1, \ 0 < \delta < 1$
Quantile slash BS	$\nu u^{\nu-1}\mathbb{1}_{(0,1)}(u)$	$[\nu\int_0^1 u^{\nu-1}\phi(y; 0, \frac{1}{u})du]\mathbb{1}_{(-\infty,\infty)}(y)$	$\nu > 0$
Quantile Student- t BS	$\frac{(\frac{\nu}{2})^{\nu/2}}{\Gamma(\frac{\nu}{2})}u^{\nu/2-1}\exp(-\frac{u\nu}{2})\mathbb{1}_{(0,\infty)}(u)$	$\frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\pi}\sqrt{\nu}\Gamma(\frac{\nu}{2})}\left(1 + \frac{y^2}{\nu}\right)^{-\frac{\nu+1}{2}}\mathbb{1}_{(-\infty,\infty)}(y)$	$\nu > 0$

such that we obtain $p(t, u)$ by calculating

$$p(t, u) = \frac{\sqrt{u} \exp\left(-u \frac{2Q}{t\alpha^2\gamma_\alpha^2} \left(\frac{\gamma_\alpha^2 t}{4Q} - 1\right)^2\right)}{v\sqrt{\delta} \exp\left(-\delta \frac{2Q}{t\alpha^2\gamma_\alpha^2} \left(\frac{\gamma_\alpha^2 t}{4Q} - 1\right)^2\right) + (1-v) \exp\left(-\frac{2Q}{t\alpha^2\gamma_\alpha^2} \left(\frac{\gamma_\alpha^2 t}{4Q} - 1\right)^2\right)}.$$

Then, the conditional expectation of $U|T = t$ can be written as

$$\mathbb{E}[U|T = t] = \frac{1 - v + v\delta^{3/2} \exp\left((1-\delta) \frac{2Q}{t\alpha^2\gamma_\alpha^2} \left(\frac{\gamma_\alpha^2 t}{4Q} - 1\right)^2\right)}{1 - v + v\sqrt{\delta} \exp\left((1-\delta) \frac{2Q}{t\alpha^2\gamma_\alpha^2} \left(\frac{\gamma_\alpha^2 t}{4Q} - 1\right)^2\right)}. \quad (7)$$

2.3.2 Quantile slash-BS distribution

Let $T \sim \text{QSBS}(\alpha, Q; H)$ and h_U be as given in Table 1, meaning that $U \sim \text{Beta}(v, 1)$. From (5) and ϕ_{SMN} in Table 1, we can write the PDF of T as

$$f_T(t) = \left[v \int_0^1 u^{v-1} \phi\left(\frac{1}{\alpha\gamma_\alpha} \sqrt{\frac{4Q}{t}} \left(\frac{\gamma_\alpha^2 t}{4Q} - 1\right); 0, \frac{1}{u}\right) du \right] \frac{1}{\alpha\gamma_\alpha \sqrt{4Qt}} \left(\frac{\gamma_\alpha^2}{2} + \frac{2Q}{t}\right),$$

$$t > 0.$$

In this case, the notation $T \sim \text{SL-BS}(\alpha, Q, v)$ is used. Note that

$$U|T = t \sim \text{Gama}\left(\frac{1}{2} + v, \frac{2Q}{t\alpha^2\gamma_\alpha^2} \left(\frac{\gamma_\alpha^2 t}{4Q} - 1\right)^2\right),$$

truncated to the interval $[0, 1]$. We can then obtain the conditional expectation as

$$\mathbb{E}[U|T = t] = \left[\frac{1 + 2v}{\frac{4Q}{t\alpha^2\gamma_\alpha^2} \left(\frac{\gamma_\alpha^2 t}{4Q} - 1\right)^2} \right] \frac{P_1\left(\frac{3}{2} + v, \frac{2Q}{t\alpha^2\gamma_\alpha^2} \left(\frac{\gamma_\alpha^2 t}{4Q} - 1\right)^2\right)}{P_1\left(\frac{1}{2} + v, \frac{2Q}{t\alpha^2\gamma_\alpha^2} \left(\frac{\gamma_\alpha^2 t}{4Q} - 1\right)^2\right)}, \quad (8)$$

where $P_x(a, b)$ is the gamma CDF evaluated at x and with parameters a and b .

2.3.3 Quantile Student- t BS distribution

Let $T \sim \text{QSBS}(\alpha, Q; H)$ and h_U be as given in Table 1, meaning that $U \sim \text{Gamma}(v/2, v/2)$. Then, based on (5) and ϕ_{SMN} in Table 1, we obtain the PDF of T as

$$f_T(t) = \frac{\Gamma(\frac{v+1}{2})}{\sqrt{\pi}\sqrt{v}\Gamma(\frac{v}{2})} \left[1 + \frac{4Q}{vt\alpha^2\gamma_\alpha^2} \left(\frac{\gamma_\alpha^2 t}{4Q} - 1\right)^2 \right]^{-\frac{v+1}{2}} \frac{1}{\alpha\gamma_\alpha \sqrt{4Qt}} \left(\frac{\gamma_\alpha^2}{2} + \frac{2Q}{t}\right), \quad t > 0,$$

with notation $T \sim t_v\text{-BS}(\alpha, Q, v)$. In this case, we have

$$U|T = t \sim \text{Gamma}\left(\frac{v+1}{2}, \frac{v}{2} + \frac{2Q}{t\alpha^2\gamma_\alpha^2} \left(\frac{\gamma_\alpha^2 t}{4Q} - 1\right)^2\right),$$

and

$$\mathbb{E}[U|T = t] = \frac{v + 1}{v + \frac{4Q}{t\alpha^2\gamma_\alpha^2} \left(\frac{\gamma_\alpha^2 t}{4Q} - 1 \right)^2}. \quad (9)$$

Figure 1 displays different shapes of the special cases of the QSBS distribution. We observe that as q increases, the curves become flatter. We also observe that larger the value of α , the heavier the tail of the distribution is.

3 QSBS quantile regression model

3.1 Model and EM algorithm

Let $\mathbf{T} = (T_1, T_2, \dots, T_n)^\top$ be a random sample of size n , where $T_i \sim \text{QSBS}(\alpha, Q_i; H)$, for $i = 1, 2, \dots, n$, and let $\mathbf{t} = (t_1, t_2, \dots, t_n)^\top$ denote the corresponding realization of $\mathbf{T}$. We propose a QSBS quantile regression model with structure for Q_i expressed as

$$h(Q_i) = \eta_i = \mathbf{x}_i^\top \boldsymbol{\beta}, \quad i = 1, 2, \dots, n,$$

where $h: \mathbb{R}^+ \rightarrow \mathbb{R}^+$ is a strictly monotonic, positive and at least twice differentiable link function, $Q_i = h^{-1}(\mathbf{x}_i^\top \boldsymbol{\beta})$, $\mathbf{x}_i^\top = (1, x_{i1}, \dots, x_{ip})$ is the vector of covariates (fixed and known), and $\boldsymbol{\beta} = (\beta_0, \beta_1, \dots, \beta_p)^\top$ is the corresponding vector of regression coefficients.

Under some restrictions on the columns of the covariate matrix $(x_{ij})_{(n+1) \times p}$ with $x_{0j} = 1$, $j = 1, \dots, p$, the following result shows that the proposed QSBS regression model is identifiable. This result is fundamental because it excludes the possibility of having multiple estimators of the unknown parameters which best fit the observed data for the parametric model. If no restrictions are imposed on the covariate matrix, the model loses the property of identifiability. The lack of this property may be eliminated under certain technical restrictions (see, e.g. Rothenberg 1971) or through reparametrization.

Proposition 9 *If the n -dimensional vectors $\mathbf{v}_0 = (1, 1, \dots, 1)^\top$, $\mathbf{v}_1 = (x_{11}, x_{21}, \dots, x_{n1})^\top$, $\mathbf{v}_2 = (x_{12}, x_{22}, \dots, x_{n2})^\top$, ..., $\mathbf{v}_p = (x_{1p}, x_{2p}, \dots, x_{np})^\top$ are linearly independent, then the function*

$$(\alpha, \boldsymbol{\beta})^\top \mapsto F_{T_i}(t) \equiv F_{T_i}(t; \alpha, \boldsymbol{\beta}), \quad \forall t > 0, \quad i = 1, 2, \dots, n,$$

is one-one.

Proof Let $(\alpha, \boldsymbol{\beta})^\top$ and $(\alpha', \boldsymbol{\beta}')^\top$ be two parameter vectors such that $F_{T_i}(t; \alpha, \boldsymbol{\beta}) = F_{T_i}(t; \alpha', \boldsymbol{\beta}')$, $i = 1, 2, \dots, n$. In what follows we verify that $(\alpha, \boldsymbol{\beta})^\top = (\alpha', \boldsymbol{\beta}')^\top$.

Indeed, from (5), we have

$$F_{T_i}(t; \alpha, \boldsymbol{\beta}) = \Phi_{\text{SMN}}(a(t; \alpha, Q_i)) = \Phi_{\text{SMN}}(a(t; \alpha', Q'_i)) = F_{T_i}(t; \alpha', \boldsymbol{\beta}'), \quad \forall t > 0,$$

with $Q_i = h^{-1}(\mathbf{x}_i^\top \boldsymbol{\beta})$, $Q'_i = h^{-1}(\mathbf{x}_i^\top \boldsymbol{\beta}')$, $a(t; \alpha, Q_i) = (\sqrt{t/(4Q_i/\gamma_\alpha^2)} - \sqrt{(4Q_i/\gamma_\alpha^2)/t})/\alpha$ and $a(t; \alpha', Q'_i) = (\sqrt{t/(4Q'_i/\gamma_{\alpha'}^2)} - \sqrt{(4Q'_i/\gamma_{\alpha'}^2)/t})/\alpha'$, $i = 1, 2, \dots, n$. As $\Phi_{\text{SMN}}(\cdot)$ is a strictly increasing function, it is one-one. Hence,

$$\frac{1}{\alpha} \left(\sqrt{\frac{t}{(4Q_i/\gamma_\alpha^2)}} - \sqrt{\frac{(4Q_i/\gamma_\alpha^2)}{t}} \right) = \frac{1}{\alpha'} \left(\sqrt{\frac{t}{(4Q'_i/\gamma_{\alpha'}^2)}} - \sqrt{\frac{(4Q'_i/\gamma_{\alpha'}^2)}{t}} \right), \quad \forall t > 0. \quad (10)$$

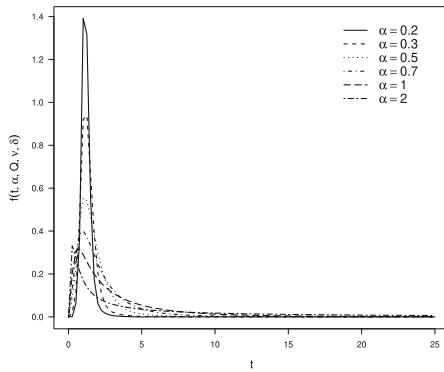
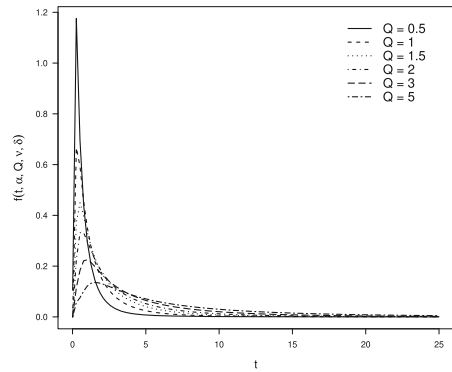
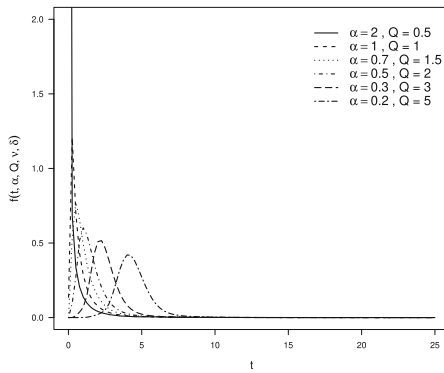
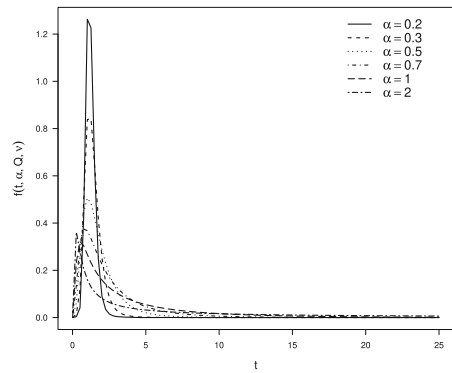
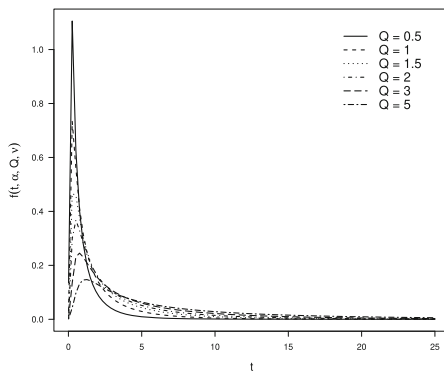
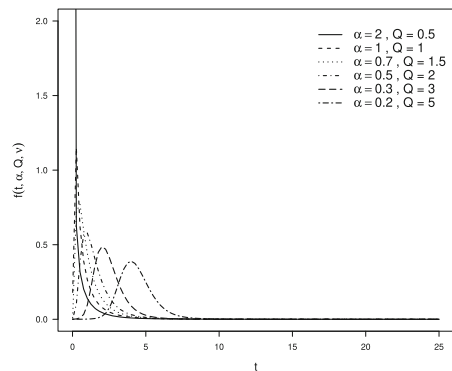
(a) CN-BS($\alpha, 1, (0.2, 0.2)$), $q = 0.25$ (b) CN-BS($1, Q, (0.2, 0.2)$), $q = 0.50$ (c) CN-BS($\alpha, Q, (0.2, 0.2)$), $q = 0.75$ (d) SL-BS($\alpha, 1, 2$), $q = 0.25$ (e) SL-BS($1, Q, 2$), $q = 0.50$ (f) SL-BS($\alpha, Q, 2$), $q = 0.75$

Fig. 1 PDFs for some QSBS family members considering different values of α , Q and $q = 0.25, 0.5, 0.75$, and for the SL-BS and CN-BS models, with $\nu = 2, 2, (0.2, 0.2)$, respectively

By taking $t = 4Q'_i/\gamma_{\alpha'}^2$ in the above identity, we get

$$\frac{1}{\alpha} \left(\sqrt{\frac{(4Q'_i/\gamma_{\alpha'}^2)}{(4Q_i/\gamma_{\alpha}^2)}} - \sqrt{\frac{(4Q_i/\gamma_{\alpha}^2)}{(4Q'_i/\gamma_{\alpha'}^2)}} \right) = 0 \iff 4Q_i/\gamma_{\alpha}^2 = 4Q'_i/\gamma_{\alpha'}^2, \quad i = 1, 2, \dots, n.$$

Substituting the identity $4Q_i/\gamma_{\alpha}^2 = 4Q'_i/\gamma_{\alpha'}^2$ into (10) and using the fact that $a(t; \alpha, Q_i)$ and $a(t; \alpha', Q'_i)$ are positive, we have $\alpha = \alpha'$, and therefore, $Q_i = Q'_i$. Consequently, $h(Q_i) = \mathbf{x}_i^{\top} \boldsymbol{\beta} = \mathbf{x}_i^{\top} \boldsymbol{\beta}' = h(Q'_i)$, $i = 1, 2, \dots, n$, or equivalently,

$$\beta_0 + \sum_{j=1}^p x_{ij} \beta_j = \beta'_0 + \sum_{j=1}^p x_{ij} \beta'_j, \quad i = 1, 2, \dots, n.$$

The above identity can be written in vector form as follows

$$(\beta_0 - \beta'_0) \mathbf{v}_0 + (\beta_1 - \beta'_1) \mathbf{v}_1 + (\beta_2 - \beta'_2) \mathbf{v}_2 + (\beta_p - \beta'_p) \mathbf{v}_p = \mathbf{0},$$

where $\mathbf{0} = (0, 0, \dots, 0)^{\top}$ is the n -dimensional zero vector. Since the vectors $\mathbf{v}_0, \mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_p$ are linearly independent, from the last identity we obtain $\beta_j = \beta'_j$, for $j = 0, 1, 2, \dots, p$. This completes the proof. $\square$

We can implement the EM algorithm (Dempster et al. (1977)) to obtain the ML estimate of $\boldsymbol{\theta} = (\alpha, \boldsymbol{\beta})^{\top}$. The parameter $\mathbf{v}$ that indexes the PDF of U , denoted by $h_U(\cdot)$, will be estimated using the profile log-likelihood. Specifically, let $\mathbf{t} = (t_1, \dots, t_n)^{\top}$ and $\mathbf{u} = (u_1, \dots, u_n)^{\top}$ be the observed and missing data, respectively, with $\mathbf{T} = (T_1, T_2, \dots, T_n)^{\top}$ and $\mathbf{U} = (U_1, U_2, \dots, U_n)^{\top}$ being their corresponding random vectors. Thus, the complete data vector is written as $\mathbf{y}_c = (\mathbf{t}^{\top}, \mathbf{u}^{\top})^{\top}$. From Proposition 3,

$$\begin{aligned} T_i | U_i = u_i &\stackrel{\text{ind.}}{\sim} \text{QBS}(\alpha_u, Q_i), \\ U_i &\stackrel{\text{ind.}}{\sim} h_U(u_i), \quad i = 1, 2, \dots, n, \end{aligned}$$

where QBS denotes the BS distribution reparameterized by the quantile proposed by Sánchez et al. (2020) and $\alpha_u = \sqrt{g(u)}\alpha$. The complete data log-likelihood function for the QSBS quantile regression model associated with $\mathbf{y}_c = (\mathbf{t}^{\top}, \mathbf{u}^{\top})^{\top}$ is given by

$$\begin{aligned} \ell_c(\boldsymbol{\theta}; \mathbf{t}_c) &\propto -n \log(\alpha \gamma_{\alpha}) - \frac{1}{2} \sum_{i=1}^n \log(Q_i) + \sum_{i=1}^n \log \left(\frac{\gamma_{\alpha}^2}{2} + \frac{2Q_i}{t_i} \right) \\ &\quad - \frac{2}{(\alpha \gamma_{\alpha})^2} \sum_{i=1}^n \frac{1}{g(u_i)} \frac{Q_i}{t_i} \left(\frac{\gamma_{\alpha}^2 t_i}{4Q_i} - 1 \right)^2. \end{aligned} \quad (11)$$

By taking the expected value of the complete data log-likelihood function in (11), conditional on $\mathbf{U} = \mathbf{u}$, and letting $\hat{u}_i = E[1/g(U_i) | T_i = t_i; \hat{\boldsymbol{\theta}}]$, $i = 1, 2, \dots, n$, which we obtain from (7), (8) and (9), we have

$$\begin{aligned} \mathcal{Q}(\boldsymbol{\theta}; \hat{\boldsymbol{\theta}}) &\propto -n \log(\alpha \gamma_{\alpha}) - \frac{1}{2} \sum_{i=1}^n \log(Q_i) + \sum_{i=1}^n \log(\gamma_{\alpha}^2 t_i + 4Q_i) \\ &\quad - \frac{2}{(\alpha \gamma_{\alpha})^2} \sum_{i=1}^n \hat{u}_i \frac{Q_i}{t_i} \left(\frac{\gamma_{\alpha}^2 t_i}{4Q_i} - 1 \right)^2. \end{aligned} \quad (12)$$

Then, the steps to obtain the ML estimates of $\theta = (\alpha, \beta)^\top$ via the EM algorithm are given by

E-Step. Compute $\hat{u}_i$, given $\theta = \hat{\theta}$, for $i = 1, 2, \dots, n$.

M-Step. Update $\hat{\theta}$ by maximizing $\mathcal{Q}(\theta; \hat{\theta})$ with respect to θ .

The maximization in M-Step can be performed by taking the derivative of Equation (12) with respect to $\theta = (\alpha, \beta_0, \dots, \beta_p)$ and then equating it to zero, thereby providing the likelihood equations. The partial derivatives of (12) with respect to θ are given by

$$\begin{aligned} \frac{\partial \mathcal{Q}(\theta; \hat{\theta})}{\partial \alpha} &= -n \left(\frac{\gamma_\alpha + \alpha \gamma'_\alpha}{\alpha \gamma_\alpha} \right) + 2\gamma_\alpha \gamma'_\alpha \sum_{i=1}^n \frac{t_i}{\gamma_\alpha^2 t_i + 4Q_i} + \frac{1}{\alpha^3} \sum_{i=1}^n \hat{u}_i \left[\frac{\gamma_\alpha^2 t_i}{4Q_i} + \frac{4Q_i}{\gamma_\alpha^2 t_i} - 2 \right] \\ &\quad - \frac{1}{\alpha^2} \sum_{i=1}^n \hat{u}_i \left[\frac{\gamma_\alpha \gamma'_\alpha t_i}{4Q_i} - \frac{4Q_i \gamma'_\alpha}{\gamma_\alpha^3 t_i} \right], \end{aligned} \quad (13)$$

$$\frac{\partial \mathcal{Q}(\theta; \hat{\theta})}{\partial \beta_j} = \sum_{i=1}^n \left\{ \left[-\frac{1}{2Q_i} + \frac{4}{\gamma_\alpha^2 t_i + 4Q_i} - \frac{\hat{u}_i}{\alpha^2} \left(-\frac{\gamma_\alpha^2 t_i}{8Q_i^2} + \frac{2}{\gamma_\alpha^2 t_i} \right) \right] \left(\frac{\partial Q_i}{\partial \eta_i} \right) \left(\frac{\partial \eta_i}{\partial \beta_j} \right) \right\}, \quad (14)$$

for $j = 0, 1, \dots, p$, where $\gamma'_\alpha = y_q + \alpha y_q^2 / \sqrt{(\alpha y_q)^2 + 4}$ is the first derivative of γ_α with respect to α . The system of equations are solved using the BFGS quasi-Newton method; see Jamshidian and Jennrich (1997); McLachlan and Krishnan (2007). The convergence of the EM algorithm can be established using the following stopping criterion: $|\mathcal{Q}(\theta^{(r+1)}; \hat{\theta}^{(r)}) - \mathcal{Q}(\theta^{(r)}; \hat{\theta}^{(r)})| < \varepsilon$, where ε is a prespecified tolerance level; see McLachlan and Krishnan (2007). Starting values are required to initiate the EM procedure, namely, $\hat{\alpha}^{(0)}, \hat{\beta}_0^{(0)}, \dots, \hat{\beta}_p^{(0)}$. These can be obtained from the work of Rieck and Nedelman (1991).

A disadvantage of the EM algorithm in relation to Newton-type methods is that we cannot obtain the estimates of the standard errors directly through the Fisher information matrix. There are several approaches proposed to obtain the standard errors of ML estimators; see, for example, Baker (1992), Oakes (1999) and Louis (1982), among others. The approach developed by Louis (1982) is based on the missing information principle. In this method, the score function of the incomplete data log-likelihood function is related to the conditional expectation of the complete data log-likelihood function as follows: $S_o(\mathbf{y}; \theta) = E[\partial \ell_c(\theta; \mathbf{t}) / \partial \theta]$, where $S_o(\mathbf{y}; \theta) = \partial \ell_o(\theta; \mathbf{t}) / \partial \theta$ and $S_c(\mathbf{y}; \theta) = \partial \ell_c(\theta; \mathbf{t}) / \partial \theta$ are the score functions of the incomplete data and complete data, respectively. Meilijson (1989) presents a definition of the empirical information matrix as follows:

$$\mathbf{I}_e(\theta; \mathbf{t}) = \sum_{i=1}^n s(t_i; \theta) s^\top(t_i; \theta) - \frac{1}{n} S(\mathbf{t}; \theta) S^\top(\mathbf{t}; \theta),$$

such that $S(\mathbf{t}; \theta) = \sum_{i=1}^n s(t_i; \theta)$ and $s(t_i; \theta)$ is called the empirical score function of the i -th observation, $i = 1, 2, \dots, n$. Replacing θ with its respective ML estimates $\hat{\theta}$, we obtain $S(\mathbf{t}; \hat{\theta}) = 0$. Then, the empirical information matrix is given by

$$\mathbf{I}_e(\hat{\theta}; \mathbf{t}) = \sum_{i=1}^n s(t_i; \hat{\theta})^{(r)} s^\top(t_i; \hat{\theta})^{(r)},$$

where $s(t_i; \hat{\theta})^{(r)}$ is the empirical score function obtained in the r -th iteration using the EM algorithm, that is,

$$s(t_i; \hat{\theta})^{(r)} = E[s(t_i, u_i^{(r)}; \theta^{(r)}) | t_i],$$

where $u_i^{(r)}$ is the latent variable obtained in the r -th iteration with conditional distribution $h_{U|T}(u_i | t_i)$ and the partial derivatives of the complete data log-likelihood function in relation to α and β are given, respectively, by (13) and (14).

By the method of Louis (1982), an approximation to the observed information matrix observed at the r -th iteration can be obtained as $I_e(\hat{\theta}; t)^{(r)} = \sum_{i=1}^n s(t_i; \hat{\theta})^{(r)} s^\top(t_i; \hat{\theta})^{(r)}$. Therefore, an approximation of the variance-covariance matrix is obtained by inverting the empirical Fisher information matrix $I_e(\hat{\theta}; t)^{(r)}$ after convergence.

We estimate the extra parameter ν (or extra parameter vector $\mathbf{v}$) by using the profile log-likelihood approach. First, we compute the ML estimate of θ by using the EM algorithm, for each $\nu \in \{a, \dots, b\}$, where a and b are predefined limits, then the final estimate of ν is the one that maximizes the log-likelihood function and the associated estimate of θ is then the final one; see Saulo et al. (2021). The extra parameter is estimated through the profile log-likelihood approach because it provides robustness to outlying observations under Student- t models. According to Lucas (1997), the robustness only holds if the degree of freedom is fixed rather than directly estimated in the maximization of the log-likelihood function. Moreover, some difficulties in computing the extra parameter can be found in other models.

3.2 Residual analysis

Residuals are important tools to assess goodness-of-fit and departures from the assumptions of the postulated model. In particular, the generalized Cox-Snell (GCS) and randomized quantile (RQ) residuals are widely used for these purposes; see Lee and Wang (2003), Dasilva et al. (2020) and Saulo et al. (2022). The GCS and RQ residuals are respectively given by

$$\hat{r}_i^{\text{GCS}} = -\log[1 - \hat{\Phi}_{\text{SMN}}(a(t))]$$

and

$$\hat{r}_i^{\text{RQ}} = \Phi^{-1}[\Phi_{\text{SMN}}(a(t))],$$

for $i = 1, 2, \dots, n$, where $\hat{\Phi}_{\text{SMN}}(a(t))$ is the CDF of $T \sim \text{QSBS}(\hat{\alpha}, \hat{\beta}, \nu)$ fitted to the data, and $\Phi(\cdot)^{-1}$ is the standard normal quantile function. These residuals asymptotically follow a standard exponential distribution and a standard normal distribution, respectively, under correct model specification. For both residuals, the distributional assumption can be verified using graphical techniques, hypothesis tests and descriptive statistics.

3.3 Hypothesis testing

We consider here the Wald, score, likelihood ratio and gradient statistical tests for the QSBS quantile regression model; see Terrell (2002) and Saulo et al. (2022, 2021). Consider θ to be a p -dimensional vector of parameters that index a QSBS quantile regression model. Assume our interest lies in testing the hypothesis $\mathcal{H}_0 : \theta_1 = \theta_1^{(0)}$ against $\mathcal{H}_1 : \theta_1 \neq \theta_1^{(0)}$, where $\theta = (\theta_1^\top, \theta_2^\top)^\top$, θ_1 is a $r \times 1$ vector of parameters of interest and θ_2 is a $(p - r) \times 1$ vector of nuisance parameters. Then, the Wald (W), score (R), likelihood ratio (LR) and gradient (T) statistics are given, respectively, by

$$\begin{aligned}
S_W &= (\hat{\boldsymbol{\theta}} - \tilde{\boldsymbol{\theta}})^\top \mathcal{J}(\hat{\boldsymbol{\theta}})(\hat{\boldsymbol{\theta}} - \tilde{\boldsymbol{\theta}}), \\
S_R &= S(\tilde{\boldsymbol{\theta}})^\top \mathcal{J}(\tilde{\boldsymbol{\theta}})^{-1} S(\tilde{\boldsymbol{\theta}}), \\
S_{LR} &= -2[\ell(\tilde{\boldsymbol{\theta}}) - \ell(\hat{\boldsymbol{\theta}})], \\
S_T &= S(\tilde{\boldsymbol{\theta}})^\top (\hat{\boldsymbol{\theta}} - \tilde{\boldsymbol{\theta}}),
\end{aligned}$$

where $\ell(\cdot)$ is the log-likelihood function, $S(\cdot)$ is the score function, $\mathcal{J}(\cdot)$ is the Fisher information matrix, and $\hat{\boldsymbol{\theta}} = (\hat{\boldsymbol{\theta}}_1^\top, \hat{\boldsymbol{\theta}}_2^\top)^\top$ and $\tilde{\boldsymbol{\theta}} = (\boldsymbol{\theta}_1^{(0)\top}, \tilde{\boldsymbol{\theta}}_2^\top)^\top$ are the unrestricted and restricted ML estimators of $\boldsymbol{\theta}$, respectively. In particular, we use the EM algorithm to obtain the ML estimates. The log-likelihood function in the S_{LR} statistic is replaced by the expected value of the complete data log-likelihood function. Moreover, for the S_W , S_R and S_T statistics, the score function and Fisher information matrix are approximated using their respective empirical versions presented in Sect. 3.1. In regular cases, we have, under $\mathcal{H}_0$ and as $n \rightarrow \infty$, the Wald, score, likelihood ratio and gradient statistical tests converging in distribution to χ_r^2 . We then reject $\mathcal{H}_0$ at nominal level α if the test statistic is larger than $\chi_{1-\alpha, r}^2$, the upper α quantile of the χ_r^2 distribution.

4 Monte Carlo simulation

In this section, the results from two Monte Carlo simulation studies are presented. In the first study, we evaluate the performances of the EM algorithm for ML estimation and residuals. In the second study, we evaluate the performances of the aforementioned statistical tests. The simulations were performed using the R software; see R Core Team (2020). The simulation scenario considers sample size $n = \{50, 100, 200\}$ and quantiles $q \in \{0.10, 0.25, 0.50, 0.75, 0.90\}$ (or subset), with a logarithmic link function for Q_i , and 5,000 Monte Carlo replications.

4.1 ML estimation and model selection

In this study, the data generation model includes two covariates and is given by

$$\log(Q_i) = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i}, \quad i = 1, \dots, n, \quad (15)$$

where x_1 and x_2 are covariates obtained from a uniform distribution in the interval (0,1), $\beta_0 = 2.5$, $\beta_1 = 3$ and $\beta_2 = 0.9$. Moreover, the simulation scenario considers shape parameters $\alpha = 0.2, 0.5, 1.0$ and extra parameters $\boldsymbol{v} = (0.1, 0.3)$ (CN-BS), $\nu = 4$ (SL-BS) and $\nu = 11$ (t_ν -BS). The response observations, $t_1, \dots, t_n$, are then generated by using (4) and (15).

We evaluate the performance of the EM algorithm for obtaining the ML estimates using the bias, mean square error (MSE), average width (AW) and coverage probability (CP) of 95% asymptotic confidence interval. The Monte Carlo estimates of these quantities are given, respectively, by

$$\begin{aligned}
\text{Bias}(\hat{\varphi}) &= \frac{1}{M} \sum_{r=1}^M (\hat{\varphi}^{(r)} - \varphi), \\
\text{MSE}(\hat{\varphi}) &= \frac{1}{M} \sum_{r=1}^M (\hat{\varphi}^{(r)} - \varphi)^2,
\end{aligned}$$

$$\text{CP}(\widehat{\varphi}) = \frac{1}{M} \sum_{r=1}^M I(\varphi \in [\widehat{\varphi}_L^{(r)}, \widehat{\varphi}_U^{(r)}]),$$

$$\text{AW}(\widehat{\varphi}) = \frac{1}{M} \sum_{r=1}^M |\widehat{\varphi}_U^{(r)} - \widehat{\varphi}_L^{(r)}|,$$

where M is the number of Monte Carlo replications, φ is the true parameter, $\widehat{\varphi}^{(r)}$ is the r -th ML estimate of φ , and $I(\cdot)$ is an indicator function of φ belonging to the r -th asymptotic interval $[\widehat{\varphi}_L^{(r)}, \widehat{\varphi}_U^{(r)}]$ with $\widehat{\varphi}_L^{(r)}$ and $\widehat{\varphi}_U^{(r)}$ denoting the lower and upper bounds, respectively.

In addition to analyzing the performance of the EM algorithm, this Monte Carlo simulation study also investigates the empirical distribution of the GCS and RQ residuals, which are assessed by the empirical mean (MN), median (MD), standard deviation (SD), coefficient of skewness (CS) and coefficient of (excess) kurtosis (CK).

Tables 2 and 3 report the empirical bias, MSE, AW and CP for the quantile regression models based on the SL-BS distributions (due to space limitations we do not present the CN-BS and t_v -BS results). A look at the results in Tables 2 and 3 allows us to conclude that as the sample size increases the bias, MSE and AW both decreases, as one would expect. Moreover, these quantities increase as α increases. Finally, the CP approaches the 95% nominal level as the sample size increases.

Tables 4 and 5 report the empirical bias, MSE, AW, and CP for the quantile regression models based on the Weibull distribution reparametrized in terms of a quantile parameter, such that this reparameterization has α and Q parameters, where α is the shape parameter and Q is the parameter related to the quantile; see (Sánchez et al. 2021). The simulated data was generated considering the SL-BS distribution for the same parameter values used in Tables 2 and 3. From Tables 4 and 5, we conclude that as the sample size increases the bias, MSE and AW both decrease. Nevertheless, the $\hat{\alpha}$ empirical bias and MSE are very high, and the PC is not close to the 95% nominal level in any of the parameters, considering all the sample sizes analyzed. Finally, the results associated with the SL-BS distributions (Tables 2 and 3) are better compared to those of the Weibull quantile regression model, as expected.

Tables 6 and 7 present the empirical MN, MD, SD, CS and CK. Note that these values are expected to be 1, 0.69, 1, 2 and 6, respectively, for the GCS residuals, and 0, 0, 1, 0 and 0, respectively, for the RQ residuals. From Tables 6 and 7, note that for the CN-BS and t_v -BS cases, as the sample size increases the values of the empirical MN, MD, SD, CS and CK approaches these values of the reference EXP(1) and N(0,1) distributions. Therefore, the GCS and RQ residuals conform well with the reference distributions for the CN-BS and t_v -BS cases. Nevertheless, in the results of the SL-BS model the considered residuals do not conform well with the reference distributions.

4.2 Hypothesis tests

We now present Monte Carlo simulation studies to evaluate the performances of the Wald, score, likelihood ratio and gradient tests. We considered two measures: null rejection rate (size) and non-null rejection rate (power). The data generating model has three covariates and is given by

$$\log(Q_i) = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i}, \quad i = 1, 2, \dots, n,$$

where the values of the coefficients β_j , $j = 0, 1, \dots, 3$, not fixed in H_0 are all equal to 1. The covariate values were obtained as U(0,1) random draws. The interest lies in testing

Table 2 Empirical bias, MSE, AW and CP from simulated data for the indicated ML estimates of the SL-BS quantile regression model ($\nu = 4$)

α	q	$n = 50$				$n = 100$				$n = 200$				
		Bias	MSE	CP	AW	Bias	MSE	CP	AW	Bias	MSE	CP	AW	
0.2	0.1	$\hat{\alpha}$	-0.0065	0.0005	0.9260	0.0886	-0.0035	0.0002	0.9366	0.0605	-0.0016	0.0001	0.9472	0.0421
		$\hat{\beta}_0$	0.0078	0.0099	0.9512	0.4135	0.0045	0.0039	0.9480	0.2485	0.0023	0.0022	0.9524	0.1893
		$\hat{\beta}_1$	0.0024	0.0111	0.9510	0.4325	-0.0005	0.0058	0.9506	0.3071	0.0004	0.0034	0.9524	0.2307
		$\hat{\beta}_2$	0.0006	0.0167	0.9578	0.5383	0.0010	0.0067	0.9520	0.3292	-0.0001	0.0031	0.9536	0.2215
	0.25	$\hat{\alpha}$	-0.0063	0.0005	0.9536	0.0884	-0.0035	0.0002	0.9502	0.0605	-0.0016	0.0001	0.9460	0.0421
		$\hat{\beta}_0$	0.0035	0.0067	0.9258	0.3497	0.0041	0.0051	0.9338	0.2374	0.0007	0.0018	0.9452	0.1831
		$\hat{\beta}_1$	0.0010	0.0105	0.9478	0.5102	-0.0014	0.0067	0.9472	0.3071	-0.0003	0.0033	0.9448	0.2467
		$\hat{\beta}_2$	0.0008	0.0131	0.9516	0.4362	-0.0015	0.0061	0.9462	0.3292	0.0008	0.0031	0.9504	0.2223
	0.5	$\hat{\alpha}$	-0.0072	0.0005	0.9542	0.0886	-0.0033	0.0002	0.9498	0.0605	-0.0019	0.0001	0.9530	0.0421
		$\hat{\beta}_0$	-0.0013	0.0067	0.9198	0.3935	-0.0006	0.0037	0.9382	0.2331	-0.0003	0.0023	0.9422	0.1803
		$\hat{\beta}_1$	0.0035	0.0118	0.9456	0.4325	0.0016	0.0067	0.9500	0.3071	-0.0001	0.0037	0.9506	0.2467
		$\hat{\beta}_2$	-0.0001	0.0117	0.9468	0.5383	0.0000	0.0068	0.9516	0.3292	0.0008	0.0033	0.9472	0.2223
	0.75	$\hat{\alpha}$	-0.0071	0.0005	0.9470	0.0886	-0.0036	0.0002	0.9522	0.0605	-0.0017	0.0001	0.9540	0.0421
		$\hat{\beta}_0$	-0.0041	0.0107	0.9212	0.3993	-0.0038	0.0033	0.9316	0.2377	-0.0013	0.0018	0.9418	0.1832
		$\hat{\beta}_1$	-0.0017	0.0176	0.9496	0.4325	0.0003	0.0064	0.9470	0.3071	-0.0002	0.0031	0.9522	0.2467
		$\hat{\beta}_2$	-0.0020	0.0140	0.9498	0.5383	0.0005	0.0067	0.9486	0.3292	-0.0006	0.0031	0.9554	0.2223
	0.9	$\hat{\alpha}$	-0.0065	0.0005	0.9260	0.0886	-0.0035	0.0002	0.9366	0.0605	-0.0016	0.0001	0.9474	0.0421
		$\hat{\beta}_0$	-0.0109	0.0099	0.9502	0.4138	-0.0057	0.0039	0.9470	0.2490	-0.0017	0.0023	0.9492	0.1904
		$\hat{\beta}_1$	0.0024	0.0111	0.9510	0.4325	-0.0006	0.0058	0.9506	0.3071	0.0005	0.0039	0.9528	0.2467
		$\hat{\beta}_2$	0.0006	0.0167	0.9578	0.5383	0.0010	0.0067	0.9520	0.3292	-0.0012	0.0031	0.9542	0.2223

Table 2 continued

α	q	$n = 50$				$n = 100$				$n = 200$			
		Bias	MSE	CP	AW	Bias	MSE	CP	AW	Bias	MSE	CP	AW
0.5	0.1	$\hat{\alpha}$	0.0030	0.9264	0.2234	-0.0091	0.0014	0.9374	0.1522	-0.0040	0.0007	0.9478	0.1057
		$\hat{\beta}_0$	0.0582	0.9534	1.0184	0.0110	0.0227	0.9502	0.6065	0.0069	0.0134	0.9510	0.4628
		$\hat{\beta}_1$	0.0661	0.9550	1.0671	-0.0010	0.0342	0.9514	0.7523	0.0015	0.0228	0.9540	0.6019
		$\hat{\beta}_2$	0.0013	0.9986	1.3303	0.0026	0.0396	0.9530	0.8063	-0.0030	0.0182	0.9560	0.5424
		$\hat{\alpha}$	-0.0180	0.0031	0.2234	-0.0096	0.0015	0.9514	0.1522	-0.0044	0.0007	0.9504	0.1057
0.5	0.25	$\hat{\beta}_0$	0.0158	0.0421	0.9857	0.0074	0.0228	0.9346	0.5812	0.0039	0.0108	0.9392	0.4468
		$\hat{\beta}_1$	-0.0068	0.0647	1.0671	0.0012	0.0327	0.9464	0.7523	-0.0007	0.0178	0.9506	0.6019
		$\hat{\beta}_2$	0.0034	0.0734	1.3303	0.0001	0.0382	0.9494	0.8063	0.0012	0.0193	0.9484	0.5424
		$\hat{\alpha}$	-0.0183	0.0031	0.2234	-0.0083	0.0015	0.9508	0.1522	-0.0043	0.0007	0.9552	0.1057
		$\hat{\beta}_0$	-0.0013	0.0459	0.9717	0.0004	0.0282	0.9356	0.5705	0.0004	0.0119	0.9354	0.4399
0.75	0.5	$\hat{\beta}_1$	-0.0006	0.0813	1.0671	-0.0009	0.0383	0.9530	0.7523	-0.0004	0.0208	0.9520	0.6019
		$\hat{\beta}_2$	-0.0015	0.0649	1.3303	0.0002	0.0427	0.9548	0.8063	-0.0001	0.0180	0.9446	0.5424
		$\hat{\alpha}$	-0.0182	0.0030	0.2234	-0.0087	0.0014	0.9508	0.1522	-0.0049	0.0007	0.9516	0.1057
		$\hat{\beta}_0$	-0.0122	0.0671	0.9860	-0.0075	0.0264	0.9378	0.5820	-0.0048	0.0108	0.9432	0.4470
		$\hat{\beta}_1$	0.0009	0.0893	1.0671	0.0006	0.0456	0.9514	0.7523	-0.0015	0.0164	0.9522	0.6019
0.9	0.9	$\hat{\beta}_2$	-0.0047	0.0767	1.3303	0.0030	0.0353	0.9466	0.8063	0.0029	0.0190	0.9484	0.5424
		$\hat{\alpha}$	-0.0168	0.0030	0.2234	-0.0091	0.0014	0.9374	0.1522	-0.0040	0.0007	0.9478	0.1057
		$\hat{\beta}_0$	-0.0274	0.0583	1.0189	-0.0144	0.0231	0.9456	0.6078	-0.0044	0.0134	0.9496	0.4632
		$\hat{\beta}_1$	0.0061	0.0656	1.0671	-0.0010	0.0342	0.9514	0.7523	0.0015	0.0228	0.9540	0.6019
		$\hat{\beta}_2$	0.0013	0.0986	1.3303	0.0026	0.0396	0.9530	0.8063	-0.0030	0.0182	0.9560	0.5424

Table 3 Empirical bias, MSE, AW and CP from simulated data for the indicated ML estimates of the SL-BS quantile regression model ($\nu = 4$), (continued)

α	q	$n = 50$				$n = 100$				$n = 200$			
		Bias	MSE	CP	AW	Bias	MSE	CP	AW	Bias	MSE	CP	AW
1	0.1	$\hat{\alpha}$	-0.0359	0.0122	0.9266	0.4534	0.0058	0.9374	0.3073	-0.0087	0.0028	0.9476	0.2125
		$\hat{\beta}_0$	0.0312	0.1863	0.9600	1.9111	0.0729	0.9556	1.1191	0.0125	0.0434	0.9530	0.8438
		$\hat{\beta}_1$	0.0093	0.2666	0.9648	2.2700	0.1113	0.9592	1.4024	0.0038	0.0742	0.9560	1.1056
		$\hat{\beta}_2$	0.0163	0.2748	0.9620	2.3184	0.1290	0.9594	1.5024	-0.0053	0.0592	0.9576	0.9961
	0.25	$\hat{\alpha}$	-0.0414	0.0127	0.9616	0.4536	0.0061	0.9580	0.3073	-0.0103	0.0029	0.9530	0.2125
		$\hat{\beta}_0$	0.0278	0.1608	0.9196	1.8955	0.0731	0.9284	1.0815	0.0081	0.0361	0.9442	0.8200
		$\hat{\beta}_1$	0.0031	0.2456	0.9554	2.0457	0.1498	0.9566	1.4024	-0.0028	0.0620	0.9524	1.1056
		$\hat{\beta}_2$	0.0041	0.2423	0.9606	2.5636	0.1439	0.9580	1.5024	0.0017	0.0641	0.9536	0.9961
	0.5	$\hat{\alpha}$	-0.0372	0.0129	0.9680	0.4536	0.0057	0.9580	0.3073	-0.0086	0.0029	0.9536	0.2125
		$\hat{\beta}_0$	-0.0057	0.1361	0.9200	1.8674	0.0736	0.9388	1.0600	0.0025	0.0385	0.9394	0.8063
		$\hat{\beta}_1$	-0.0088	0.2754	0.9582	2.0457	0.1190	0.9532	1.4024	-0.0064	0.0595	0.9524	1.1056
		$\hat{\beta}_2$	0.0192	0.2526	0.9618	2.5636	0.1142	0.9554	1.5024	0.0013	0.0621	0.9528	0.9961
	0.75	$\hat{\alpha}$	-0.0398	0.0130	0.9634	0.4536	0.0059	0.9544	0.3073	-0.0091	0.0029	0.9528	0.2125
		$\hat{\beta}_0$	-0.0289	0.1368	0.9158	1.8952	0.0708	0.9372	1.0829	-0.0096	0.0337	0.9446	0.8206
		$\hat{\beta}_1$	-0.0062	0.2156	0.9482	2.0457	0.1237	0.9532	1.4024	-0.0001	0.0592	0.9492	1.1056
		$\hat{\beta}_2$	0.0062	0.2486	0.9608	2.5636	0.1315	0.9524	1.5024	0.0053	0.0575	0.9542	0.9961
	0.9	$\hat{\alpha}$	-0.0370	0.0122	0.9304	0.4543	0.0058	0.9374	0.3073	-0.0087	0.0028	0.9476	0.2125
		$\hat{\beta}_0$	-0.0506	0.1411	0.9506	1.6096	0.0742	0.9510	1.1214	-0.0090	0.0434	0.9526	0.8447
		$\hat{\beta}_1$	-0.0029	0.2486	0.9574	2.1531	0.1113	0.9592	1.4024	0.0038	0.0742	0.9560	1.1056
		$\hat{\beta}_2$	0.0104	0.2601	0.9586	2.1956	0.1290	0.9594	1.5024	-0.0053	0.0592	0.9576	0.9961

Table 4 Empirical bias, MSE, AW and CP from simulated SL-BS data for the indicated ML estimates of the Weibull quantile regression model

α	q	$n = 50$				$n = 100$				$n = 200$			
		Bias	MSE	CP	AW	Bias	MSE	CP	AW	Bias	MSE	CP	AW
0.2	0.1	$\hat{\alpha}$	4.5834	21.4565	0.0000	1.9762	4.3689	19.3173	0.0000	1.3057	4.2296	18.0191	0.0000
		$\hat{\beta}_0$	-0.0768	0.0190	0.8568	0.3898	-0.0954	0.0166	0.7250	0.2848	-0.1069	0.0164	0.5344
		$\hat{\beta}_1$	-0.0013	0.0218	0.8624	0.4411	0.0003	0.0100	0.8696	0.2978	0.0013	0.0075	0.8454
		$\hat{\beta}_2$	-0.0029	0.0226	0.8674	0.4642	-0.0004	0.0116	0.8596	0.3192	-0.0016	0.0059	0.8526
0.25	0.1	$\hat{\alpha}$	4.5895	21.5119	0.0000	1.9779	4.3699	19.3263	0.0000	1.3055	4.2285	18.0094	0.0000
		$\hat{\beta}_0$	-0.0063	0.0143	0.8874	0.3917	-0.0115	0.0060	0.8998	0.2514	-0.0171	0.0043	0.8694
		$\hat{\beta}_1$	0.0056	0.0182	0.8530	0.3996	-0.0001	0.0101	0.8672	0.2977	0.0014	0.0075	0.8500
		$\hat{\beta}_2$	0.0016	0.0258	0.8692	0.4959	0.0001	0.0115	0.8572	0.3190	-0.0016	0.0059	0.8544
0.5	0.1	$\hat{\alpha}$	4.5905	21.5190	0.0000	1.9771	4.3677	19.3055	0.0000	1.3045	4.2288	18.0119	0.0000
		$\hat{\beta}_0$	0.02888	0.0145	0.8686	0.3689	0.0315	0.0065	0.8488	0.2328	0.0312	0.0047	0.8246
		$\hat{\beta}_1$	0.0056	0.0184	0.8532	0.3996	-0.0002	0.0101	0.8636	0.2978	0.0009	0.0075	0.8488
		$\hat{\beta}_2$	0.0018	0.0260	0.8694	0.4959	-0.0005	0.0116	0.8638	0.3192	-0.0017	0.0059	0.8530
0.75	0.1	$\hat{\alpha}$	4.5869	21.4873	0.0000	1.9753	4.3679	19.3079	0.0000	1.3049	4.2293	18.0160	0.0000
		$\hat{\beta}_0$	0.0242	0.0146	0.8642	0.3620	0.0328	0.0068	0.8346	0.2267	0.0368	0.0053	0.7920
		$\hat{\beta}_1$	0.0059	0.0183	0.8546	0.3998	-0.0000	0.0100	0.8680	0.2979	0.0009	0.0075	0.8490
		$\hat{\beta}_2$	0.0018	0.0258	0.8682	0.4962	-0.0004	0.0116	0.8620	0.3192	-0.0017	0.0059	0.8498
0.9	0.1	$\hat{\alpha}$	4.5920	21.5282	0.0000	1.9760	4.3684	19.3143	0.0000	1.3048	4.2300	18.0229	0.0000
		$\hat{\beta}_0$	-0.0044	0.0148	0.8522	0.3585	0.0065	0.0064	0.8416	0.2274	0.0135	0.0045	0.8346
		$\hat{\beta}_1$	0.0048	0.0270	0.8654	0.4993	0.0001	0.0100	0.8644	0.2978	0.0011	0.0075	0.8484
		$\hat{\beta}_2$	-0.0014	0.0215	0.8498	0.4321	-0.0002	0.0116	0.8574	0.3192	-0.0014	0.0058	0.8570

Table 4 continued

α	q	$n = 50$				$n = 100$				$n = 200$			
		Bias	MSE	CP	AW	Bias	MSE	CP	AW	Bias	MSE	CP	AW
0.5	0.1	$\hat{\alpha}$	1.4867	2.2762	0.0000	0.8230	2.0023	0.0000	0.5470	1.3539	1.8497	0.0000	0.3716
		$\hat{\beta}_0$	-0.1789	0.1230	0.8658	1.0413	0.0806	0.7640	0.6826	-0.2298	0.0770	0.5930	0.5213
		$\hat{\beta}_1$	0.0133	0.0996	0.8650	0.9591	0.0525	0.8802	0.7120	0.0032	0.0377	0.8642	0.5850
		$\hat{\beta}_2$	0.0035	0.1417	0.8768	1.1905	0.0607	0.8738	0.7632	-0.0038	0.0298	0.8736	0.5271
		$\hat{\alpha}$	1.4870	2.2770	0.0000	0.8223	2.0009	0.0000	0.5469	1.3537	1.8491	0.0000	0.3715
0.5	0.25	$\hat{\beta}_0$	0.0001	0.0784	0.9012	0.9407	0.0310	0.9146	0.6020	-0.0200	0.0203	0.8926	0.4677
		$\hat{\beta}_1$	0.0138	0.0995	0.8664	0.9592	0.0527	0.8796	0.7121	0.0019	0.0377	0.8656	0.5850
		$\hat{\beta}_2$	0.0035	0.1420	0.8752	1.1906	0.0608	0.8700	0.7633	-0.0040	0.0295	0.8756	0.5272
		$\hat{\alpha}$	1.4867	2.2764	0.0000	0.8221	2.0004	0.0000	0.5467	1.3535	1.8484	0.0000	0.3715
		$\hat{\beta}_0$	0.0716	0.0796	0.8774	0.8861	0.0349	0.8608	0.5569	0.0785	0.0251	0.8370	0.4381
0.75	0.5	$\hat{\beta}_1$	0.0137	0.0992	0.8648	0.9595	0.0524	0.8790	0.7124	0.0020	0.0376	0.8666	0.5851
		$\hat{\beta}_2$	0.0043	0.1412	0.8770	1.1909	0.0606	0.8748	0.7635	-0.0042	0.0297	0.8766	0.5272
		$\hat{\alpha}$	1.4864	2.2750	0.0000	0.8220	2.0014	0.0000	0.5468	1.3537	1.8490	0.0000	0.3715
		$\hat{\beta}_0$	0.0481	0.0791	0.8704	0.8688	0.0349	0.8504	0.5420	0.0760	0.0254	0.8250	0.4285
		$\hat{\beta}_1$	0.0132	0.0992	0.8636	0.9593	0.0525	0.8788	0.7122	0.0024	0.0377	0.8654	0.5850
0.9	0.9	$\hat{\beta}_2$	0.0040	0.1418	0.8768	1.1905	0.0609	0.8772	0.7634	-0.0043	0.0297	0.8696	0.5272
		$\hat{\alpha}$	1.4869	2.2771	0.0000	0.8223	2.0010	0.0000	0.5468	1.3536	1.8485	0.0000	0.3715
		$\hat{\beta}_0$	-0.0279	0.0816	0.8620	0.8715	0.0328	0.8602	0.5440	0.0162	0.0214	0.8604	0.4298
		$\hat{\beta}_1$	0.0138	0.0994	0.8634	0.9592	0.0525	0.8778	0.7122	0.0015	0.0377	0.8680	0.5850
		$\hat{\beta}_2$	0.0046	0.1415	0.8786	1.1904	0.0609	0.8716	0.7633	-0.0041	0.0297	0.8720	0.5273

Table 5 Empirical bias, MSE, AW and CP from simulated SL-BS data for the indicated ML estimates of the Weibull quantile regression model. (continued)

α	q	$n = 50$				$n = 100$				$n = 200$						
		Bias		MSE	CP	AW	Bias		MSE	CP	AW	Bias		MSE	CP	AW
		$\hat{\alpha}$	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\alpha}$	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\alpha}$	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\alpha}$	$\hat{\beta}_0$	$\hat{\beta}_1$
1	0.1	$\hat{\alpha}$	0.0828	0.0214	0.9188	0.4528	0.0456	0.0087	0.9248	0.3048	0.0248	0.0038	0.9258	0.2094		
		$\hat{\beta}_0$	-0.2728	0.3477	0.8944	1.9055	-0.3095	0.2037	0.8358	1.2435	-0.3374	0.1793	0.7104	0.9451		
		$\hat{\beta}_1$	0.0228	0.2995	0.8834	1.7494	-0.0009	0.1525	0.9020	1.2906	0.0014	0.1049	0.8948	1.0548		
		$\hat{\beta}_2$	0.0097	0.4345	0.8898	2.1711	-0.0005	0.1751	0.8968	1.3833	-0.0070	0.0833	0.8980	0.9505		
	0.25	$\hat{\alpha}$	0.0831	0.0215	0.9196	0.4526	0.0456	0.0087	0.9212	0.3047	0.0249	0.0038	0.9282	0.2094		
		$\hat{\beta}_0$	0.0584	0.2461	0.9096	1.7178	0.0474	0.0940	0.9274	1.0937	0.0370	0.0582	0.9138	0.8455		
		$\hat{\beta}_1$	0.0241	0.2993	0.8832	1.7487	-0.0001	0.1524	0.9016	1.2905	0.0021	0.1048	0.8908	1.0549		
		$\hat{\beta}_2$	0.0084	0.4372	0.8906	2.1701	0.0003	0.1756	0.8974	1.3834	-0.0065	0.0834	0.8972	0.9505		
	0.5	$\hat{\alpha}$	0.0833	0.0215	0.9176	0.4524	0.0454	0.0087	0.9224	0.3046	0.0249	0.0038	0.9248	0.2094		
		$\hat{\beta}_0$	0.1363	0.2502	0.8862	1.6157	0.1503	0.1092	0.8728	1.0096	0.1538	0.0773	0.8498	0.7903		
		$\hat{\beta}_1$	0.0245	0.3003	0.8832	1.7487	0.0001	0.1525	0.8994	1.2909	0.0025	0.1045	0.8926	1.0551		
		$\hat{\beta}_2$	0.0095	0.4341	0.8908	2.1703	-0.0005	0.1745	0.8960	1.3838	-0.0060	0.0832	0.8966	0.9507		
	0.75	$\hat{\alpha}$	0.0834	0.0215	0.9204	0.4524	0.0454	0.0087	0.9230	0.3046	0.0251	0.0038	0.9274	0.2094		
		$\hat{\beta}_0$	0.0412	0.2359	0.8852	1.5837	0.0741	0.0942	0.8898	0.9824	0.0886	0.0627	0.8744	0.7724		
		$\hat{\beta}_1$	0.0235	0.2986	0.8826	1.7487	0.0002	0.1522	0.8992	1.2908	0.0030	0.1041	0.8944	1.0547		
		$\hat{\beta}_2$	0.0095	0.4342	0.8894	2.1700	-0.0020	0.1757	0.8958	1.3835	-0.0075	0.0833	0.8960	0.9502		
	0.9	$\hat{\alpha}$	0.0831	0.0215	0.9196	0.4524	0.0456	0.0087	0.9246	0.3047	0.0250	0.0038	0.9282	0.2094		
		$\hat{\beta}_0$	-0.0917	0.2507	0.8738	1.5907	-0.0466	0.0956	0.8798	0.9865	-0.0229	0.0581	0.8904	0.7751		
		$\hat{\beta}_1$	0.0226	0.2996	0.8808	1.7493	0.0007	0.1524	0.9002	1.2906	0.0023	0.1048	0.8934	1.0546		
		$\hat{\beta}_2$	0.0082	0.4332	0.8902	2.1711	-0.0005	0.1754	0.8986	1.3835	-0.0073	0.0831	0.9012	0.9503		

Table 6 Summary statistics for the GCS residuals

α	q	$n = 50$				$n = 100$				$n = 200$			
		CN-BS	SL-BS	t_0 -BS	t_0 -BS	CN-BS	SL-BS	t_0 -BS	t_0 -BS	CN-BS	SL-BS	t_0 -BS	t_0 -BS
0.2	0.25	MIN	1.0007	0.2249	1.0002	1.0003	3.4373	1.0000	1.0000	1.0002	2.7172	1.0001	1.0001
		MD	0.6980	0.0000	0.6964	0.6956	0.7111	0.6952	0.6952	0.6941	0.2370	0.6940	0.6940
		SD	1.0008	0.7546	1.0007	0.9994	4.4814	0.9993	0.9993	1.0000	3.8889	0.9998	0.9998
		CS	1.6378	4.0803	1.6428	1.7848	1.0108	1.7883	1.7883	1.8767	1.2936	1.8774	1.8774
		CK	3.1251	17.7753	3.1549	4.1353	-0.4425	4.1577	4.1577	4.8594	0.4671	4.8577	4.8577
	0.5	MIN	1.0004	6.9641	0.9999	1.0003	1.1698	1.0000	1.0000	1.0002	2.8319	1.0001	1.0001
		MD	0.6974	4.5335	0.6968	0.6956	0.0004	0.6952	0.6952	0.6941	0.5567	0.6940	0.6940
		SD	0.9997	6.6111	0.9995	0.9994	2.5135	0.9993	0.9993	1.0000	3.8723	0.9999	0.9999
		CS	1.6322	0.5019	1.6373	1.7848	2.3886	1.7883	1.7883	1.8767	1.2494	1.8774	1.8774
		CK	3.0959	-1.2344	3.1217	4.1354	4.9818	4.1578	4.1578	4.8594	0.3308	4.8576	4.8576
0.5	0.75	MIN	1.0004	3.4180	0.9999	1.0003	5.4234	1.0000	1.0000	1.0002	13.7272	1.0001	1.0001
		MD	0.6974	1.1057	0.6968	0.6956	3.1935	0.6952	0.6952	0.6941	14.9827	0.6940	0.6940
		SD	0.9997	4.3039	0.9995	0.9994	5.6753	0.9993	0.9993	1.0000	6.0437	0.9999	0.9999
		CS	1.6322	0.9856	1.6373	1.7849	0.5817	1.7883	1.7883	1.8768	-0.5336	1.8774	1.8774
		CK	3.0961	-0.4622	3.1218	4.1356	-1.1559	4.1579	4.1579	4.8596	-0.6473	4.8578	4.8578
	0.25	MIN	1.0002	6.2318	0.9996	1.0001	5.0268	0.9998	0.9998	1.0001	1.0149	1.0000	1.0000
		MD	0.6976	5.4538	0.6970	0.6957	3.8996	0.6953	0.6953	0.6941	0.2530	0.6939	0.6939
		SD	0.9978	4.8210	0.9976	0.9984	4.3076	0.9983	0.9983	0.9995	1.6239	0.9993	0.9993
		CS	1.6229	0.4127	1.6271	1.7793	0.6738	1.7821	1.7821	1.8734	2.3450	1.8738	1.8738
		CK	3.0526	-1.0874	3.0718	4.1048	-0.6312	4.1224	4.1224	4.8365	6.2482	4.8344	4.8344
0.5	0.5	MIN	1.0001	3.4666	0.9996	1.0001	3.4666	0.9998	0.9998	1.0001	3.2460	1.0000	1.0000
		MD	0.6976	2.4009	0.6970	0.6957	2.4009	0.6953	0.6953	0.6941	2.0450	0.6939	0.6939
		SD	0.9978	3.3556	0.9976	0.9984	3.3556	0.9983	0.9983	0.9995	3.3414	0.9993	0.9993
		CS	1.6229	0.9513	1.6271	1.7793	0.9513	1.7821	1.7821	1.8734	1.1899	1.8738	1.8738
		CK	3.0526	0.0021	3.0717	4.1049	0.0021	4.1224	4.1224	4.8366	0.6885	4.8344	4.8344

Table 6 continued

α	q	$n = 50$				$n = 100$				$n = 200$			
		$n = 50$		t_{ν} -BS		$n = 100$		t_{ν} -BS		$n = 200$		t_{ν} -BS	
		CN-BS	SL-BS	SL-BS	t_{ν} -BS	CN-BS	SL-BS	SL-BS	t_{ν} -BS	CN-BS	SL-BS	SL-BS	t_{ν} -BS
1	0.75	MN	1.0001	1.3331	0.9996	1.0001	0.7800	0.9998	1.0001	2.4466	1.0000	1.0000	
		MD	0.6976	0.4540	0.6970	0.6957	0.2002	0.6953	0.6941	1.0708	0.6939	0.6939	
		SD	0.9978	1.9673	0.9976	0.9984	1.3305	0.9983	0.9995	3.0264	0.9993	0.9993	
		CS	1.6229	1.9690	1.6271	1.7794	2.6832	1.7822	1.8734	1.4214	1.8738	1.8738	
		CK	3.0527	3.8322	3.0718	4.1049	8.5732	4.1224	4.8366	1.2562	4.8345	4.8345	
	0.25	MN	0.9993	1.7884	0.9989	0.9997	0.8007	0.9994	0.9997	2.6806	0.9997	0.9997	
		MD	0.6986	1.0537	0.6979	0.6961	0.3817	0.6957	0.6942	1.7059	0.6941	0.6941	
		SD	0.9921	2.0649	0.9918	0.9956	1.1435	0.9953	0.9978	2.7691	0.9976	0.9976	
		CS	1.5961	1.7839	1.5976	1.7637	2.6123	1.7642	1.8639	1.6326	1.8633	1.8633	
		CK	2.9246	3.3013	2.9240	4.0165	8.8216	4.0183	4.7733	2.5941	4.7663	4.7663	
0.5	0.5	MN	0.9993	1.3375	0.9989	0.9997	1.4969	0.9994	0.9997	0.6229	0.9997	0.9997	
		MD	0.6986	0.7459	0.6979	0.6961	0.8467	0.6957	0.6942	0.2700	0.6941	0.6941	
		SD	0.9921	1.6592	0.9918	0.9956	1.8269	0.9953	0.9978	0.9235	0.9976	0.9976	
		CS	1.5961	1.9805	1.5976	1.7637	2.0812	1.7642	1.8639	2.7777	1.8633	1.8633	
		CK	2.9246	4.4033	2.9239	4.0165	5.0290	4.0183	4.7733	10.5620	4.7663	4.7663	
	0.75	MN	0.9993	1.9122	0.9989	0.9997	3.4851	0.9994	0.9997	0.9887	0.9997	0.9997	
		MD	0.6986	1.0797	0.6979	0.6961	2.3979	0.6957	0.6942	0.4844	0.6941	0.6941	
		SD	0.9921	2.2636	0.9918	0.9956	3.2461	0.9953	0.9978	1.3787	0.9976	0.9976	
		CS	1.5961	1.7622	1.5976	1.7637	1.2772	1.7642	1.8639	2.6184	1.8633	1.8633	
		CK	2.9247	3.0441	2.9240	4.0165	1.1415	4.0184	4.7733	8.8776	4.7663	4.7663	

Table 7 Summary statistics for the RQ residuals

α	q	$n = 50$				$n = 100$				$n = 200$			
		CN-BS	SL-BS	t_U -BS	t_U -BS	CN-BS	SL-BS	t_U -BS	t_U -BS	CN-BS	SL-BS	t_U -BS	t_U -BS
0.2	0.25	MN	3.6621	-0.0003	-0.0003	-0.0004	-0.0222	-0.0001	-0.0002	-0.0002	0.5812	-0.0002	-0.0002
		MD	4.4962	-0.0002	-0.0002	-0.0010	0.0091	-0.0007	-0.0002	-0.0002	0.8356	-0.0001	-0.0001
		SD	2.1911	1.0093	1.0093	1.0045	3.1032	1.0044	1.0022	1.0022	3.1333	1.0021	1.0021
		CS	-0.0054	-0.9419	-0.0057	-0.0023	-0.0018	-0.0028	-0.0035	-0.0035	-0.1484	-0.0033	-0.0033
		CK	-0.2145	-0.1194	-0.2044	-0.1149	-1.4337	-0.1129	-0.0611	-0.0611	-1.4539	-0.0606	-0.0606
	0.5	MN	-0.0006	-2.1173	-0.0001	-0.0004	2.2921	-0.0001	-0.0002	-0.0002	0.2326	-0.0002	-0.0002
		MD	-0.0012	-2.2758	-0.0008	-0.0010	3.4442	-0.0007	-0.0002	-0.0002	0.2018	-0.0001	-0.0001
		SD	1.0092	2.6909	1.0092	1.0045	2.8829	1.0044	1.0022	1.0022	2.9603	1.0021	1.0021
		CS	-0.0027	0.3635	-0.0030	-0.0023	-0.7669	-0.0028	-0.0035	-0.0035	-0.0082	-0.0033	-0.0033
		CK	-0.2150	-1.0596	-0.2074	-0.1149	-0.7860	-0.1129	-0.0611	-0.0611	-1.3435	-0.0606	-0.0606
0.5	0.75	MN	-0.0006	-0.2725	-0.0001	-0.0004	-1.1287	-0.0001	-0.0002	-0.0002	-4.5000	-0.0002	-0.0002
		MD	-0.0012	-0.4018	-0.0008	-0.0010	-1.7211	-0.0007	-0.0002	-0.0002	-4.9832	-0.0001	-0.0001
		SD	1.0092	2.8834	1.0092	1.0045	3.1633	1.0044	1.0022	1.0022	1.5547	1.0021	1.0021
		CS	-0.0027	0.0591	-0.0030	-0.0023	0.3594	-0.0029	-0.0035	-0.0035	1.2891	-0.0033	-0.0033
		CK	-0.2149	-1.3973	-0.2074	-0.1149	-1.2754	-0.1129	-0.0611	-0.0611	1.2414	-0.0606	-0.0606
	0.25	MN	-0.0006	-2.3923	-0.0001	-0.0004	-1.9300	-0.0001	-0.0001	-0.0001	0.7844	-0.0001	-0.0001
		MD	-0.0012	-2.5996	-0.0007	-0.0010	-2.0355	-0.0006	-0.0001	-0.0001	0.7682	0.0000	0.0000
		SD	1.0088	1.8164	1.0088	1.0042	1.8304	1.0042	1.0020	1.0020	1.8929	1.0020	1.0020
		CS	-0.0021	0.3022	-0.0023	-0.0023	0.2409	-0.0027	-0.0035	-0.0035	-0.0576	-0.0033	-0.0033
		CK	-0.2258	-1.0039	-0.2194	-0.1212	-0.8613	-0.1196	-0.0649	-0.0649	-0.7816	-0.0645	-0.0645
0.5	0.5	MN	-0.0006	-0.2474	-0.0001	-0.0004	-1.2189	-0.0001	-0.0001	-0.0001	-1.1277	-0.0000	-0.0000
		MD	-0.0012	-0.3091	-0.0007	-0.0010	-1.3267	-0.0006	-0.0001	-0.0001	-1.1230	0.0000	0.0000
		SD	1.0088	1.9014	1.0088	1.0042	1.8326	1.0042	1.0020	1.0020	1.7712	1.0020	1.0020
		CS	-0.0021	0.1187	-0.0023	-0.0023	0.2699	-0.0027	-0.0035	-0.0035	0.0708	-0.0033	-0.0033
		CK	-0.2258	-0.7672	-0.2194	-0.1212	-0.6692	-0.1196	-0.0649	-0.0649	-0.6857	-0.0645	-0.0645

Table 7 continued

α	q	$n = 50$			$n = 100$			$n = 200$		
		CN-BS	SL-BS	t_v -BS	CN-BS	SL-BS	t_v -BS	CN-BS	SL-BS	t_v -BS
1	0.75	MN	0.3435	-0.0001	-0.0004	0.9719	-0.0001	-0.0001	-0.4675	-0.0000
		MD	0.3910	-0.0007	-0.0010	0.9259	-0.0006	-0.0001	-0.3988	0.0000
		SD	1.0088	1.0088	1.0042	1.7567	1.0042	1.0020	1.9726	1.0020
		CS	-0.0021	-0.0023	-0.0023	-0.0706	-0.0027	-0.0035	-0.0207	-0.0033
		CK	-0.2258	-0.2194	-0.1212	-0.6657	-0.1196	-0.0649	-0.8315	-0.0645
1	0.25	MN	-0.0007	-0.0002	-0.0004	0.5441	-0.0000	0.0000	-1.0096	0.0001
		MD	-0.0012	-0.0008	-0.0009	0.4899	-0.0007	0.0001	-0.9013	0.0002
		SD	1.0071	1.0074	1.0034	1.3864	1.0035	1.0016	1.4481	1.0016
		CS	-0.0004	-0.0004	-0.0020	0.1054	-0.0022	-0.0034	-0.2055	-0.0032
		CK	-0.2586	-0.2562	-0.1399	-0.0067	-0.1402	-0.0758	-0.2410	-0.0763
0.5	0.5	MN	-0.0007	-0.0002	-0.0004	-0.1923	-0.0000	0.0000	0.8391	0.0001
		MD	-0.0012	-0.0008	-0.0009	-0.1726	-0.0007	0.0001	0.7316	0.0002
		SD	1.0071	1.0074	1.0034	1.4142	1.0035	1.0016	1.4253	1.0016
		CS	-0.0004	-0.0004	-0.0020	-0.0342	-0.0022	-0.0034	0.2282	-0.0032
		CK	-0.2587	-0.2562	-0.1399	-0.0014	-0.1402	-0.0758	-0.1461	-0.0763
0.75	0.75	MN	-0.0007	-0.0002	-0.0004	-1.4258	-0.0000	0.0000	0.3369	0.0001
		MD	-0.0012	-0.0008	-0.0009	-1.3223	-0.0007	0.0001	0.2991	0.0002
		SD	1.0071	1.0074	1.0034	1.4880	1.0035	1.0016	1.4284	1.0016
		CS	-0.0005	-0.0004	-0.0020	-0.1390	-0.0022	-0.0034	0.0640	-0.0032
		CK	-0.2586	-0.2562	-0.1399	-0.5011	-0.1402	-0.0758	0.0339	-0.0763

$H_0 : \beta_{3-\kappa+1} = \dots = \beta_3 = 0$, with $\kappa = 1, 3$ against $H_0 : \beta_{3-\kappa+1} = \dots = \beta_3 \neq 0$, with $\kappa = 1, 3$. The tests' nominal levels used are $\alpha = 0.01, 0.05, 0.1$.

Tables 8, 9 and 10 present the simulation results of null rejection rates for $H_0 : \beta_3 = 0$ and in the CN-BS, SL-BS and t_v -BS quantile regression models, respectively. From these tables, we observe the following results:

- For nominal levels of 1%, 5%, and 10%, the statistics S_T and S_W consistently show values closer to the expected rates.
- S_{LR} and S_R , while performing reasonably well, occasionally show slight rates above nominal values, especially for smaller sample sizes ($n = 50$) or higher quantiles ($q = 0.75$).
- As the sample size increases ($n = 50, 100, 200$), the alignment of all null rejection rates with nominal levels improves. Particularly, the statistic S_T shows the most consistent improvement.
- For lower quantiles ($q = 0.25$), all test statistics perform closer to nominal levels compared to higher quantiles ($q = 0.75$).
- Higher quantiles tend to show inflated rejection rates for S_R .
- S_W and S_T are the most reliable across all conditions, as their rejection rates remain close to nominal values even under varying quantiles and sample sizes.
- S_{LR} and S_R are more sensitive to changes in distributional properties, with slightly higher rejection rates at extreme quantiles.
- The tests' null rejection rates do not seem to be affected by the shape parameter α .

In general, Tables 8, 9 and 10 suggest that S_T and S_W are the preferred test statistics for hypothesis testing in this context. While S_{LR} and S_R are useful, they may require caution in small samples or at higher quantiles.

Tables 15, 16 and 17 (Appendix A) present the simulation results of null rejection rates for $H_0 : \beta_1 = \beta_2 = \beta_3 = 0$ in the CN-BS, SL-BS and t_v -BS quantile regression models, respectively. Analogously to the case where $\kappa = 1$, the tests' null rejection rates does not change according to the value of α . Moreover, the results confirm that S_T and S_W are the most reliable test statistics across different models and conditions.

Figures 2 and 3 display the test power curves for the CN-BS, SL-BS and t_v -BS models considering the hypotheses $H_0 : \beta_{3-\kappa+1} = \dots = \beta_3 = 0$ against $H_1 : \beta_{3-\kappa+1} = \dots = \beta_3 = \delta$ with $|\delta| = 0, 1, 2, 3, 4$ and $\kappa = 1, 3$, where $\pi(\delta)$ denotes the power function. The significance level considered was 1% with $n = 100$ and $q = 0.5$ (other settings present similar results). From this figure, we observe that there are no changes in the power performances of the tests, as well as for any of the CN-BS, SL-BS and t_v -BS models, so that we have $\pi(\delta) = 1$ for $|\delta| \geq 1$.

5 Application to personal accident insurance data

QSBS quantile regression models are now used to analyze a data set related to the personal injury insurance claims that correspond to amounts of paid money by an insurance policy in Australian dollars (response variable, amount). The data is collected from 767 individuals and the claims occurred between January 1998 and January 1999; see De Jong and Heller (2008). The covariates considered in the study are: `optime`, denoting the operating time in percentage; `legrep`, with (1) or without (0) legal representation; and `month`, denoting the month of accident occurrence. Due to lack of correlation between the covariate `month` and the response variable `amount`, we removed this covariate from our analysis.

Table 8 Null rejection rates for $H_0 : \beta_3 = 0$ in the CN-BS quantile regression model ($\nu = 0.1, \delta = 0.3$)

α	q	$n = 50$			$n = 100$			$n = 200$		
		1%	5%	10%	1%	5%	10%	1%	5%	10%
0.2	S_W	0.0126	0.0516	0.1026	0.0150	0.0566	0.1024	0.0078	0.0488	0.1040
	S_{LR}	0.0214	0.0768	0.1324	0.0196	0.0728	0.1260	0.0120	0.0668	0.1262
	S_R	0.0386	0.1080	0.1764	0.0204	0.0782	0.1356	0.0142	0.0630	0.1232
	S_T	0.0120	0.0618	0.1156	0.0114	0.0502	0.1038	0.0100	0.0534	0.1026
0.5	S_W	0.0126	0.0514	0.1026	0.0150	0.0566	0.1022	0.0078	0.0488	0.1038
	S_{LR}	0.0186	0.0786	0.1386	0.0196	0.0730	0.1260	0.0120	0.0668	0.1262
	S_R	0.0258	0.0894	0.1486	0.0198	0.0740	0.1354	0.0116	0.0600	0.1102
	S_T	0.0086	0.0486	0.1012	0.0082	0.0504	0.1020	0.0126	0.0544	0.1040
0.75	S_W	0.0126	0.0514	0.1024	0.0148	0.0566	0.1020	0.0078	0.0490	0.1042
	S_{LR}	0.0186	0.0786	0.1386	0.0196	0.0728	0.1260	0.0120	0.0668	0.1262
	S_R	0.0380	0.1068	0.1824	0.0198	0.0758	0.1370	0.0164	0.0736	0.1320
	S_T	0.0124	0.0544	0.1112	0.0088	0.0508	0.1012	0.0108	0.0506	0.1038
0.5	S_W	0.0130	0.0482	0.0964	0.0142	0.0554	0.1024	0.0074	0.0488	0.1014
	S_{LR}	0.0178	0.0780	0.1374	0.0200	0.0716	0.1252	0.0114	0.0660	0.1236
	S_R	0.0378	0.1104	0.1834	0.0236	0.0760	0.1420	0.0120	0.0602	0.1112
	S_T	0.0128	0.0592	0.1110	0.0108	0.0480	0.1018	0.0086	0.0526	0.0940
0.5	S_W	0.0130	0.0482	0.0964	0.0142	0.0554	0.1024	0.0074	0.0488	0.1012
	S_{LR}	0.0178	0.0780	0.1374	0.0200	0.0716	0.1252	0.0114	0.0660	0.1236
	S_R	0.0216	0.0834	0.1486	0.0202	0.0704	0.1236	0.0122	0.0560	0.1056
	S_T	0.0062	0.0468	0.0982	0.0122	0.0518	0.1076	0.0092	0.0524	0.1044

Table 8 continued

α	q	$n = 50$				$n = 100$				$n = 200$			
		1%	5%	10%		1%	5%	10%		1%	5%	10%	
1	0.75	S_W	0.0130	0.0482	0.0964	0.0142	0.0556	0.1024	0.0074	0.0488	0.1012		
		S_{LR}	0.0178	0.0780	0.1374	0.0200	0.0716	0.1252	0.0114	0.0660	0.1236		
		S_R	0.0442	0.1186	0.1932	0.0200	0.0716	0.1342	0.0150	0.0668	0.1180		
		S_T	0.0144	0.0604	0.1122	0.0108	0.0556	0.1102	0.0134	0.0588	0.1158		
	0.25	S_W	0.0108	0.0418	0.0830	0.0122	0.0500	0.0932	0.0062	0.0460	0.0980		
		S_{LR}	0.0180	0.0728	0.1322	0.0176	0.0632	0.1160	0.0110	0.0634	0.1168		
		S_R	0.0424	0.1252	0.1890	0.0256	0.0834	0.1422	0.0158	0.0624	0.1162		
		S_T	0.0092	0.0544	0.1114	0.0084	0.0556	0.1104	0.0108	0.0494	0.1066		
	0.5	S_W	0.0108	0.0418	0.0830	0.0122	0.0500	0.0932	0.0062	0.0460	0.0980		
		S_{LR}	0.0180	0.0728	0.1322	0.0178	0.0632	0.1160	0.0110	0.0634	0.1168		
		S_R	0.0238	0.0898	0.1594	0.0182	0.0724	0.1286	0.0136	0.0582	0.1096		
		S_T	0.0086	0.0484	0.1018	0.0070	0.0444	0.0972	0.0096	0.0458	0.0928		
0.75	0.75	S_W	0.0108	0.0418	0.0830	0.0122	0.0500	0.0932	0.0062	0.0460	0.0980		
		S_{LR}	0.0180	0.0728	0.1322	0.0176	0.0632	0.1160	0.0110	0.0634	0.1168		
		S_R	0.0384	0.1166	0.1834	0.0210	0.0774	0.1342	0.0120	0.0600	0.1154		
		S_T	0.0130	0.0538	0.1070	0.0112	0.0516	0.1052	0.0106	0.0520	0.1066		

Table 9 Null rejection rates for $H_0 : \beta_3 = 0$ in the SL-BS quantile regression model ($\nu = 4$)

α	q	$n = 50$				$n = 100$				$n = 200$			
		1%		5%	10%	1%		5%	10%	1%		5%	10%
		S_W	S_{LR}	S_R	S_T	S_W	S_{LR}	S_R	S_T	S_W	S_{LR}	S_R	S_T
0.2	0.25					0.0116	0.0170	0.0284	0.0088	0.0106	0.0122	0.0554	0.1060
						0.0146	0.0284	0.0088	0.0106	0.0112	0.0122	0.0554	0.1060
						0.0366	0.0078	0.0522	0.1034	0.0112	0.0122	0.0554	0.1060
						0.0078	0.0106	0.0490	0.0908	0.0112	0.0122	0.0554	0.1060
	0.5					0.0116	0.0170	0.0284	0.0088	0.0106	0.0122	0.0554	0.1060
						0.0146	0.0284	0.0088	0.0106	0.0112	0.0122	0.0554	0.1060
						0.0366	0.0078	0.0522	0.1034	0.0112	0.0122	0.0554	0.1060
						0.0078	0.0106	0.0490	0.0908	0.0112	0.0122	0.0554	0.1060
	0.75					0.0116	0.0170	0.0284	0.0088	0.0106	0.0122	0.0554	0.1060
						0.0146	0.0284	0.0088	0.0106	0.0112	0.0122	0.0554	0.1060
						0.0366	0.0078	0.0522	0.1034	0.0112	0.0122	0.0554	0.1060
						0.0078	0.0106	0.0490	0.0908	0.0112	0.0122	0.0554	0.1060
0.5	0.25					0.0116	0.0170	0.0284	0.0088	0.0106	0.0122	0.0554	0.1060
						0.0146	0.0284	0.0088	0.0106	0.0112	0.0122	0.0554	0.1060
						0.0366	0.0078	0.0522	0.1034	0.0112	0.0122	0.0554	0.1060
						0.0078	0.0106	0.0490	0.0908	0.0112	0.0122	0.0554	0.1060
	0.5					0.0116	0.0170	0.0284	0.0088	0.0106	0.0122	0.0554	0.1060
						0.0146	0.0284	0.0088	0.0106	0.0112	0.0122	0.0554	0.1060
						0.0366	0.0078	0.0522	0.1034	0.0112	0.0122	0.0554	0.1060
						0.0078	0.0106	0.0490	0.0908	0.0112	0.0122	0.0554	0.1060
	0.75					0.0116	0.0170	0.0284	0.0088	0.0106	0.0122	0.0554	0.1060
						0.0146	0.0284	0.0088	0.0106	0.0112	0.0122	0.0554	0.1060
						0.0366	0.0078	0.0522	0.1034	0.0112	0.0122	0.0554	0.1060
						0.0078	0.0106	0.0490	0.0908	0.0112	0.0122	0.0554	0.1060

Table 9 continued

α	q	$n = 50$			$n = 100$			$n = 200$		
		1%	5%	10%	1%	5%	10%	1%	5%	10%
1	0.75	S_W	0.0100	0.0452	0.0852	0.0106	0.0486	0.0980	0.0446	0.0926
		S_{LR}	0.0142	0.0590	0.1112	0.0126	0.0644	0.0108	0.0562	0.1096
		S_R	0.0348	0.1076	0.1820	0.0248	0.0836	0.0162	0.0606	0.1132
		S_T	0.0086	0.0522	0.1080	0.0104	0.0492	0.0102	0.0450	0.0938
	0.25	S_W	0.0072	0.0310	0.0696	0.0088	0.0422	0.0076	0.0434	0.0926
		S_{LR}	0.0140	0.0624	0.1206	0.0108	0.0546	0.0088	0.0486	0.0950
		S_R	0.0402	0.1184	0.1794	0.0196	0.0730	0.0178	0.0616	0.1146
		S_T	0.0084	0.0510	0.1066	0.0124	0.0606	0.0094	0.0528	0.1042
	0.5	S_W	0.0072	0.0310	0.0696	0.0088	0.0420	0.0076	0.0434	0.0926
		S_{LR}	0.0108	0.0530	0.1098	0.0144	0.0574	0.0110	0.0442	0.0944
		S_R	0.0246	0.0874	0.1500	0.0186	0.0684	0.0130	0.0642	0.1162
		S_T	0.0080	0.0474	0.0986	0.0076	0.0476	0.0080	0.0470	0.0994
0.75	0.75	S_W	0.0072	0.0310	0.0696	0.0088	0.0422	0.0076	0.0434	0.0926
		S_{LR}	0.0124	0.0568	0.1102	0.0130	0.0550	0.0100	0.0454	0.0894
		S_R	0.0344	0.1048	0.1744	0.0230	0.0816	0.0144	0.0596	0.1218
		S_T	0.0104	0.0638	0.1130	0.0110	0.0524	0.0130	0.0520	0.1066

Table 10 Null rejection rates for $H_0 : \beta_3 = 0$ in the t_{ν} -BS quantile regression model ($\nu = 11$)

α	q	$n = 50$				$n = 100$				$n = 200$			
		1%		5%		1%		5%		1%		5%	
			10%		10%		10%		10%		10%		10%
0.2	0.25	S_W	0.0126	0.0534	0.1036	0.0156	0.0546	0.1062	0.0070	0.0502	0.1052	0.0070	0.0502
		S_{LR}	0.0162	0.0696	0.1296	0.0182	0.0664	0.1230	0.0092	0.0598	0.1164	0.0092	0.0598
		S_R	0.0332	0.0966	0.1634	0.0204	0.0786	0.1428	0.0140	0.0666	0.1222	0.0140	0.0666
		S_T	0.0096	0.0534	0.1100	0.0086	0.0544	0.1032	0.0104	0.0490	0.1054	0.0104	0.0490
	0.5	S_W	0.0126	0.0536	0.1038	0.0156	0.0546	0.1062	0.0070	0.0502	0.1052	0.0070	0.0502
		S_{LR}	0.0162	0.0696	0.1296	0.0182	0.0662	0.1230	0.0092	0.0598	0.1162	0.0092	0.0598
		S_R	0.0210	0.0812	0.1446	0.0150	0.0646	0.1208	0.0126	0.0568	0.1070	0.0126	0.0568
		S_T	0.0086	0.0480	0.1010	0.0116	0.0484	0.1036	0.0122	0.0494	0.1054	0.0122	0.0494
	0.75	S_W	0.0126	0.0536	0.1038	0.0156	0.0546	0.1064	0.0070	0.0502	0.1052	0.0070	0.0502
		S_{LR}	0.0162	0.0694	0.1296	0.0184	0.0662	0.1232	0.0092	0.0598	0.1160	0.0092	0.0598
		S_R	0.0424	0.1118	0.1776	0.0202	0.0742	0.1394	0.0184	0.0676	0.1274	0.0184	0.0676
		S_T	0.0106	0.0548	0.1082	0.0102	0.0536	0.1028	0.0098	0.0520	0.1052	0.0098	0.0520
0.5	0.25	S_W	0.0124	0.0514	0.0980	0.0138	0.0560	0.1024	0.0066	0.0482	0.1046	0.0066	0.0482
		S_{LR}	0.0160	0.0666	0.1274	0.0178	0.063	0.1216	0.0088	0.0576	0.1126	0.0088	0.0576
		S_R	0.0320	0.1006	0.1736	0.0212	0.0748	0.1400	0.0146	0.0578	0.1162	0.0146	0.0578
		S_T	0.012	0.0608	0.1174	0.0112	0.0562	0.1036	0.0104	0.0528	0.0974	0.0104	0.0528
	0.5	S_W	0.0124	0.0514	0.0980	0.0138	0.0560	0.1024	0.0066	0.0482	0.1046	0.0066	0.0482
		S_{LR}	0.0160	0.0666	0.1274	0.0178	0.0630	0.1216	0.0088	0.0576	0.1126	0.0088	0.0576
		S_R	0.0218	0.0840	0.1466	0.0144	0.0630	0.1178	0.0100	0.0576	0.1072	0.0100	0.0576
		S_T	0.0092	0.0530	0.1026	0.0114	0.0504	0.1034	0.0098	0.0510	0.0980	0.0098	0.0510

Table 10 continued

α	q	$n = 50$			$n = 100$			$n = 200$			
		1%	5%	10%	1%	5%	10%	1%	5%	10%	
1	0.75	S_W	0.0124	0.0514	0.0980	0.0138	0.0560	0.1024	0.0066	0.0482	0.1046
		S_{LR}	0.0160	0.0666	0.1274	0.0178	0.0630	0.1216	0.0088	0.0576	0.1126
		S_R	0.0326	0.1030	0.1688	0.0180	0.0690	0.1326	0.0126	0.0648	0.1234
		S_T	0.0106	0.0576	0.1136	0.0146	0.0586	0.1050	0.0104	0.0528	0.1044
	0.25	S_W	0.0104	0.0426	0.0846	0.0120	0.0510	0.0952	0.0064	0.0438	0.0992
		S_{LR}	0.0136	0.0630	0.1230	0.0154	0.0584	0.1144	0.0078	0.0508	0.1072
		S_R	0.0338	0.1046	0.1722	0.0196	0.0744	0.1364	0.0124	0.0592	0.1208
		S_T	0.0112	0.0586	0.1092	0.0120	0.0526	0.1014	0.0112	0.0548	0.1054
	0.5	S_W	0.0104	0.0426	0.0846	0.0120	0.0510	0.0952	0.0064	0.0438	0.0992
		S_{LR}	0.0136	0.0630	0.1230	0.0154	0.0584	0.1144	0.0078	0.0508	0.1072
		S_R	0.0246	0.0888	0.1424	0.0188	0.0654	0.1188	0.0122	0.0566	0.1104
		S_T	0.0070	0.0498	0.1016	0.0094	0.0492	0.1016	0.0108	0.0502	0.1028
0.75	S_W	0.0104	0.0426	0.0846	0.0120	0.0510	0.0952	0.0064	0.0438	0.0992	
	S_{LR}	0.0136	0.0630	0.1230	0.0154	0.0584	0.1144	0.0078	0.0508	0.1072	
	S_R	0.0360	0.1148	0.1822	0.0250	0.0818	0.1356	0.0140	0.0632	0.1212	
	S_T	0.0124	0.0582	0.1082	0.0120	0.0568	0.1076	0.0096	0.0470	0.1002	

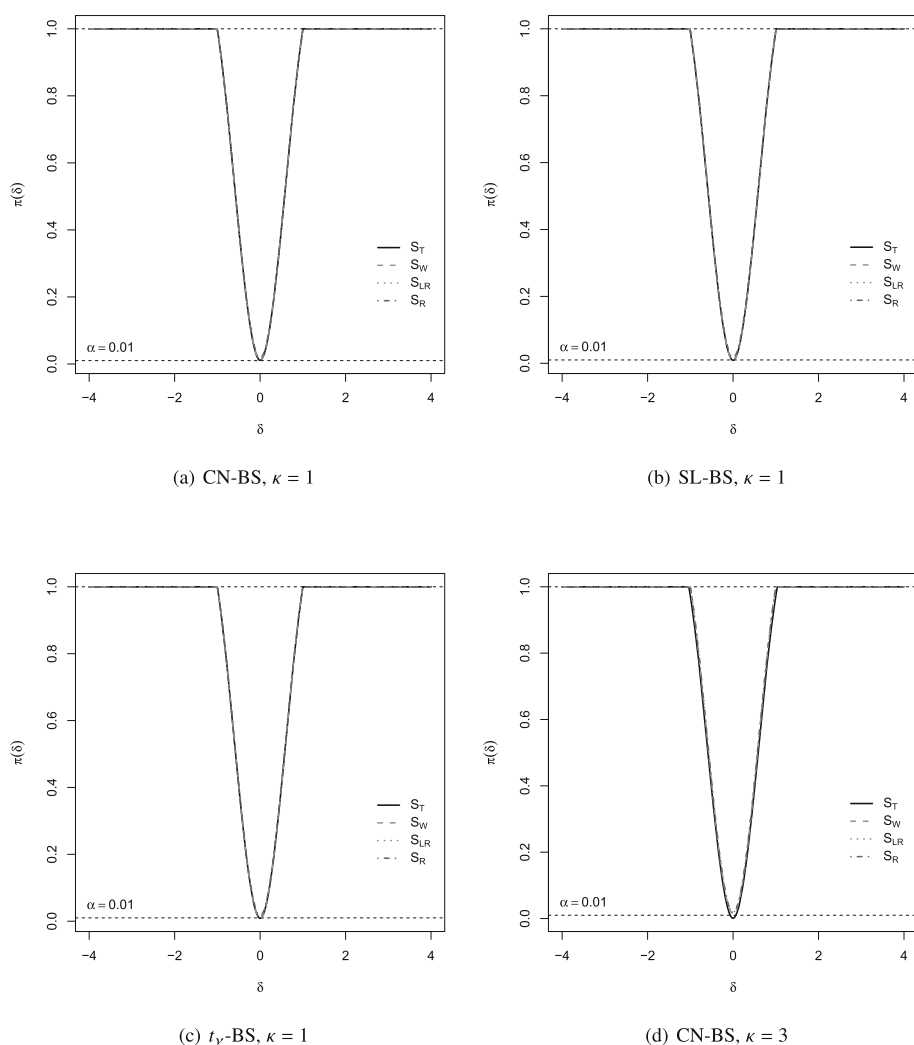


Fig. 2 Power curves of the S_W , S_{LR} , S_R and S_T tests for the CN-BS, SL-BS and t_V -BS quantile regression models (nominal level = 1%, $n = 100$, $q = 0.5$)

Table 11 reports the descriptive statistics of the observed amounts of paid money that includes MN, MD, SD, coefficient of variation (CV), CS, CK (excess) and range. From this table, we observe skewed and high kurtosis features in the data. Figure 4(a) confirms the skewness observed in Table 11. Figure 4(b) displays the usual and adjusted boxplots (Saulo et al. (2019)), where the latter is useful when the data is distributed as skewed. Note that the adjusted boxplot indicates that some potential outliers identified by the usual boxplot are not outliers.

We then analyze the personal accident insurance data using the QSBS quantile regression model, expressed as

$$\log(Q_i) = \beta_0 + \beta_1 \text{optime}_i + \beta_2 \text{legrep}_i, \quad i = 1, 2, \dots, 767. \quad (16)$$

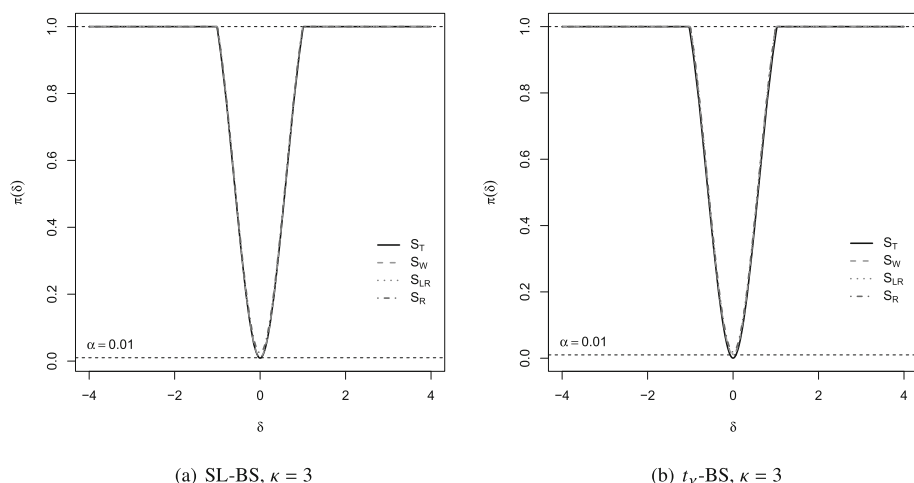


Fig. 3 Power curves of the S_W , S_{LR} , S_R and S_T tests for the SL-BS and t_v -BS quantile regression models (nominal level = 1%, $n = 100$, $q = 0.5$)

Table 11 Summary statistics for the personal accident insurance data

MN	MD	SD	CV	CS	CK	Range	Min	Max	n
7820.59	6000	8339.26	106.63%	5.09	47.77	116556.7	30	116586.7	767

Table 12 reports the the averages of the AIC, BIC, AICc and HQIC values based on $q \in \{0.01, 0.02, \dots, 0.99\}$ for the CN-BS, SL-BS, t_v -BS, Weibull (Sánchez et al. (2021)) and semiparametric (Koenker and Bassett (1978)) quantile regression models. The results indicate that the lowest values of AIC, BIC, AICc and HQIC are those based on the CN-BS model. Figure 5 displays the estimates of the CN-BS model parameters across $q \in \{0.01, 0.02, \dots, 0.99\}$. From this figure, we observe that the estimate of β_0 tends to increase with an increase in q . Moreover, the estimates of β_1 , β_2 , α and ν present a cyclic behavior. These results show the relevance of considering a quantile approach rather than traditional mean/median approaches.

Table 13 presents the ML estimates computed by the EM algorithm and SEs of the CN-BS quantile regression model parameters considering the quantiles $q = 0.025, 0.25, 0.5, 0.75, 0.975$. For the quantiles analyzed, we can observe that the impact of operational time and legal representation is greater for the quantile $q = 0.75$, the same is true with the estimates of α and β_0 .

The regression coefficients in the proposed QSBS quantile regression model (16) describe the effect of covariates on the dependent variable variable T_i . The conditional quantile of T_i is given by:

$$Q(T_i | \text{optime}_i, \text{legrep}_i) = \exp(\beta_0 + \beta_1 \text{optime}_i + \beta_2 \text{legrep}_i).$$

If optime_i increases by one unit while legrep_i remains fixed, the new conditional quantile is:

$$Q(T_i | \text{optime}_i + 1, \text{legrep}_i) = \exp(\beta_1) Q(T_i | \text{optime}_i, \text{legrep}_i).$$

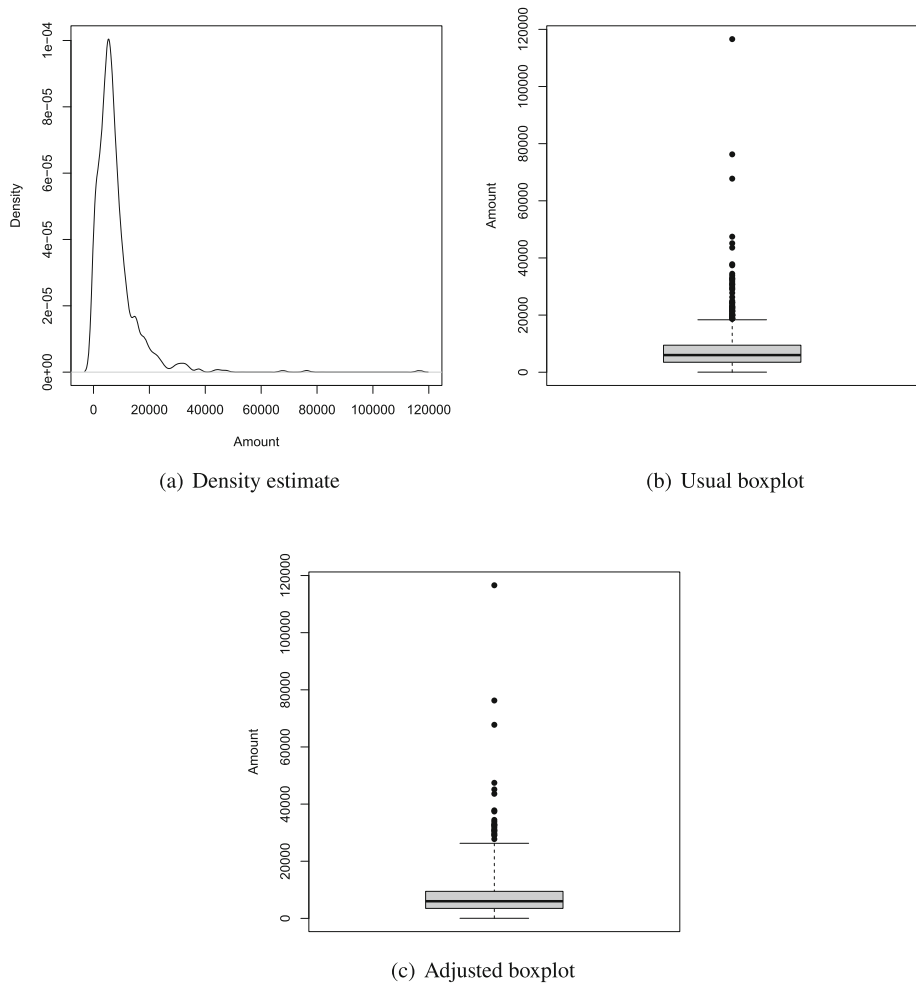


Fig. 4 Density estimate and boxplots for the personal accident insurance data

Table 12 Averages of the AIC, BIC, AICc and HQIC values based on $q \in \{0.01, 0.02, \dots, 0.99\}$ for different models for the personal accident insurance data

Model	AIC	BIC	AICc	HQIC
CN-BS	8204.63	8223.20	8204.69	8211.78
SL-BS	8220.06	8238.63	8220.11	8227.21
t_V -BS	8213.11	8231.68	8213.16	8220.26
Weibull	16,350.38	16,364.30	16,350.41	16,355.74
QR	15,648.84	15,662.77	15,648.87	15,654.20

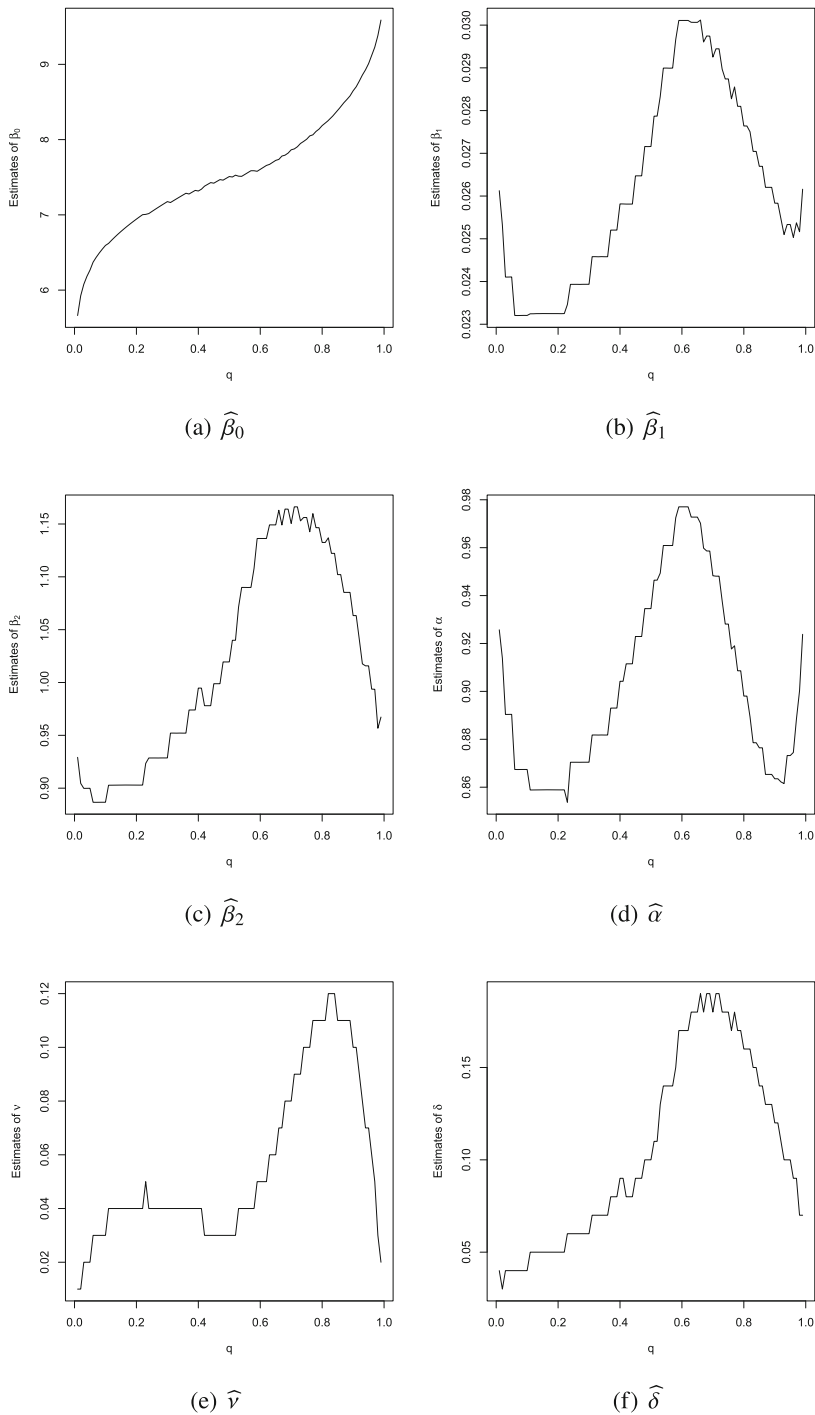


Fig. 5 Estimated parameters in the CN-BS quantile regression models across q for the personal accident insurance data

Table 13 ML estimates (with SE in parentheses) for the CN-BS quantile regression model for the personal accident insurance data

q	Estimate (SE)					
	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\alpha}$	$\hat{\nu}$	$\hat{\delta}$
0.025	5.9965(0.0541)	0.0253(0.0033)	0.9045(0.0630)	0.9138(0.0205)	0.01	0.03
0.25	7.0447(0.0543)	0.0239(0.0033)	0.9286(0.0638)	0.8704(0.0217)	0.04	0.06
0.5	7.5092(0.0543)	0.0272(0.0033)	1.0195(0.0671)	0.9356(0.0220)	0.03	0.10
0.75	8.0053(0.0614)	0.0287(0.0030)	1.1561(0.0694)	0.9281(0.0217)	0.10	0.18
0.975	9.3056(0.0698)	0.0254(0.0033)	0.9936(0.0658)	0.8885(0.0220)	0.05	0.09

Table 14 Observed values of the indicated test statistics and their p -values (in parentheses) for the CN-BS quantile regression model for the personal accident insurance data

q	Hypothesis	Test statistics			
		S_W	S_{LR}	S_R	S_T
0.25	$H_0 : \beta_1 = 0$	53.47(< 0.0001)	61.97(< 0.0001)	149.80(< 0.0001)	144.23(< 0.0001)
	$H_0 : \beta_2 = 0$	212.12(< 0.0001)	13.44(0.0002)	36.39(< 0.0001)	40.17(< 0.0001)
0.5	$H_0 : \beta_1 = 0$	68.51(< 0.0001)	159.35(< 0.0001)	202.04(< 0.0001)	165.38(< 0.0001)
	$H_0 : \beta_2 = 0$	230.97(< 0.0001)	17.20(< 0.0001)	42.47(< 0.0001)	47.92(< 0.0001)
0.75	$H_0 : \beta_1 = 0$	90.68(< 0.0001)	231.28(< 0.0001)	262.37(< 0.0001)	190.90(< 0.0001)
	$H_0 : \beta_2 = 0$	275.56(< 0.0001)	36.01(< 0.0001)	40.56(< 0.0001)	51.65(< 0.0001)

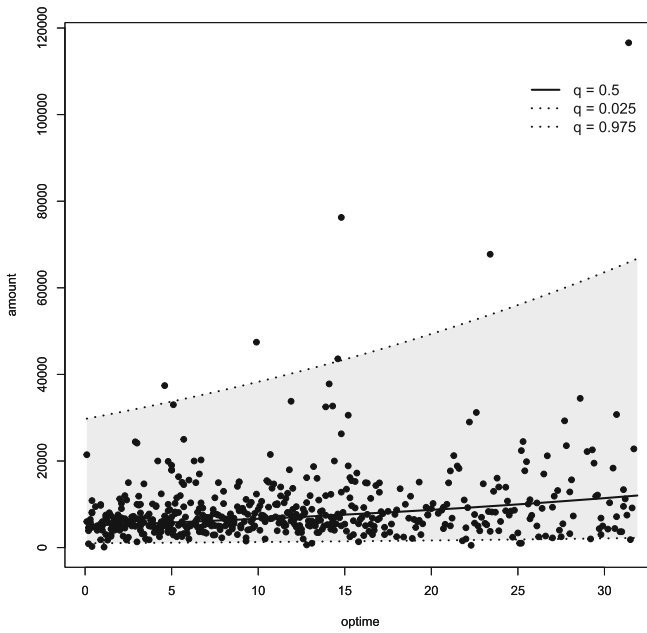
This implies that increasing optime_i by one unit multiplies the conditional quantile of T_i by $\exp(\beta_1)$. This change can also be interpreted as a percentage difference:

$$\frac{Q(T_i | \text{optime}_i + 1, \text{legrep}_i) - Q(T_i | \text{optime}_i, \text{legrep}_i)}{Q(T_i | \text{optime}_i, \text{legrep}_i)} \times 100\% \\ = (\exp(\beta_1) - 1) \times 100\%.$$

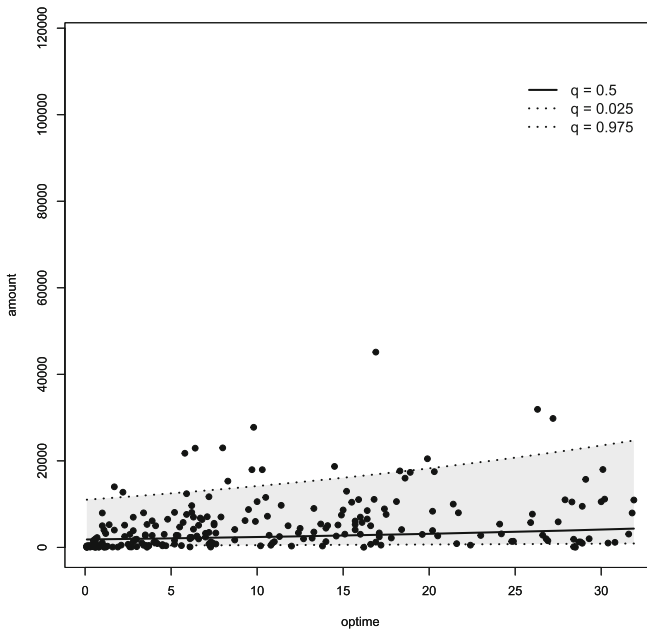
When β_1 is small (approximately $-0.4 \leq \beta_1 \leq 0.4$), the approximation $\exp(\beta_1) - 1 \approx \beta_1$ can be applied. For a binary covariate, this expression represents the percentage change in the quantile of T_i when the covariate value changes from 0 to 1.

From Table 13, we observe, for instance, that one additional percentage point of operating time (optime), increases in $(\exp(0.0239) - 1) \times 100\% = 2.42\%$ the 25° percentile ($q = 0.25$) of the amount of paid money by the insurance, while it increases by $(\exp(0.0287) - 1) \times 100\% = 2.91\%$ the 75° percentile ($q = 0.75$). For individuals with legal representation (legrep), there is an increase in the 25° percentile ($q = 0.25$) of the amount of paid money by the insurance of $(\exp(0.9286) - 1) \times 100\% = 153.10\%$ when compared to individuals without legal representation. The increase goes to $(\exp(1.1561) - 1) \times 100\% = 217.75\%$ in the 75° percentile ($q = 0.75$). In other words, the effect of legal representation on the amount of paid money by the insurance is greater for individuals with larger amounts to receive (higher quantiles).

Table 14 presents the results for the CN-BS quantile regression model to test the null hypotheses $H_0 : \beta_1 = 0$ and $H_0 : \beta_2 = 0$ using the S_W , S_{LR} , S_R and S_T test statistics. At a 5% significance level, we reject the null hypothesis in all cases tested in Table 14, indicating that coefficients associated with the operating time and legal representation are significant. In addition, from Table 14, we observe that the values of S_W , S_{LR} , S_R and S_T statistics increase as q increases.



(a) legrep = Yes



(b) legrep = No

Fig. 6 95% confidence bands in the CN-BS quantile regression model for the personal accident insurance data

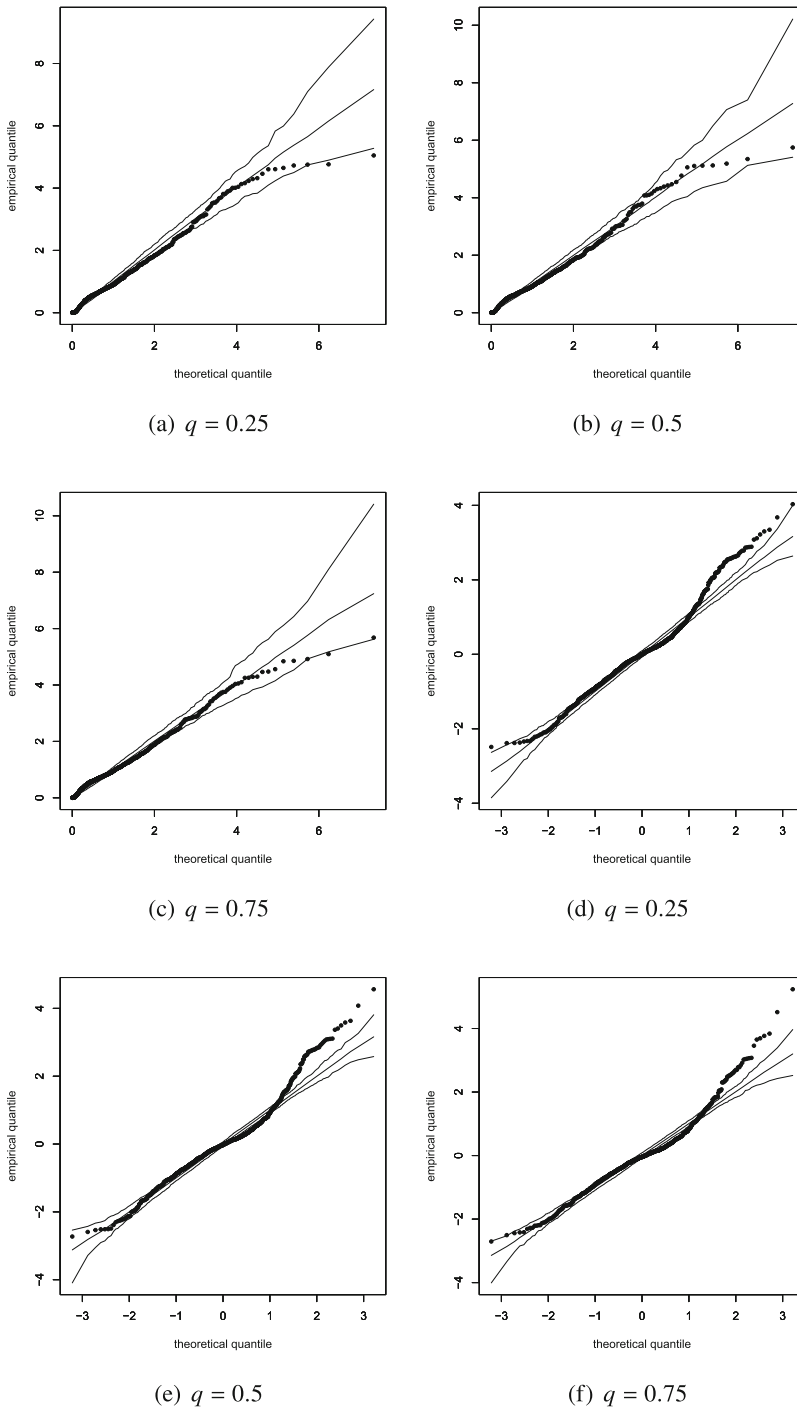


Fig. 7 QQ plots and their envelopes for the GCS (a, b and c) and RQ (d, e and f) residuals in the CN-BS quantile regression model for the personal accident insurance data

Figure 6 presents the scatterplot of the amount against the optime according to the group with legal representation and without legal representation with 95% confidence bands in the CN-BS quantile regression model for the personal accident insurance data. We use ML estimates based on the quantiles 2.5% and 97.5% to build these confidence bands. We can see that the upper bound of the confidence bands increases exponentially.

Figure 7 shows the QQ plots with simulated envelope of the GCS and RQ residuals for the CN-BS quantile regression model considered in Table 13. We can see clearly that the CN-BS model provides a good fit based on the GCS residual. On the other hand, the plots for the RQ residuals present some more extreme points outside the region delimited by the confidence bands.

6 Concluding remarks

In this paper, we have proposed a class of quantile regression models based on the reparameterized scale-mixture Birnbaum-Saunders distributions, which have the distribution quantile as one of their parameters. The relevance of this work lies in the proposal of a new quantile regression model that presents a new alternative capable of providing a very flexible fit for strictly positive and asymmetric data. The results obtained in the Monte Carlo simulation studies showed that (a) the maximum likelihood estimates obtained by the EM algorithm, in terms of bias, MSE and CP, perform well; (b) the GCS and RQ residuals conform well with their respective reference distributions with the exception of the SL-BS model; and (c) the Wald, likelihood ratio, score and gradient tests have similar performances. However, in terms of power, the Wald and gradient tests show null rejection rates that are closer to nominal levels. Thus, the Wald and gradient tests are preferable to the other tests analyzed. We have applied the proposed models to a real data set related to personal accident insurance. While the application of the proposed class of quantile regression models to personal accident insurance data demonstrates its usefulness to heavy-tailed and skewed data, some limitations persist. Notably, the model's assumptions, such as the choice of link functions or distribution families, can influence results and may not perfectly capture all data nuances. In addition, the computational demands of the EM algorithm, particularly with large datasets or multiple covariates, can be challenging. Future studies could extend the model to incorporate the presence of nonlinearity. Therefore, tackling non-linear frameworks that integrate both parametric and nonparametric elements continues to be an unresolved issue. Moreover, exploring alternative optimization techniques or Bayesian frameworks may enhance computational efficiency. Additionally, it will be of interest to study influence diagnostic tools as well as multivariate versions. Furthermore, Bartlett and Bartlett-type corrections can be used to attenuate the size distortion of hypothesis tests. Finally, expanding applications to other domains, such as health or environmental risk data, could further validate and refine the model's versatility.

Appendix A Simulation results: hypothesis tests

See Tables 15, 16 and 17.

Table 15 Null rejection rates for $H_0 : \beta_1 = \beta_2 = \beta_3 = 0$ in the CN-BS quantile regression model ($\nu = 0.1, \delta = 0.3$)

α	q	$n = 50$				$n = 100$				$n = 200$			
		1%		5%		10%		5%		10%		5%	
0.2	0.25	S_W	0.0156	0.0570	0.1082	0.0124	0.0540	0.1046	0.0108	0.0512	0.1028	0.0108	0.0512
		S_{LR}	0.0210	0.0792	0.1402	0.0202	0.0778	0.1344	0.0188	0.0742	0.1362	0.0188	0.0742
		S_R	0.0388	0.1204	0.1972	0.0232	0.0836	0.1488	0.0150	0.0662	0.1214	0.0150	0.0662
		S_T	0.0066	0.0476	0.1024	0.0102	0.0472	0.0978	0.0090	0.0468	0.1012	0.0090	0.0468
	0.5	S_W	0.0190	0.0608	0.1078	0.0158	0.0558	0.1034	0.0110	0.0478	0.0956	0.0110	0.0478
		S_{LR}	0.0200	0.0850	0.1402	0.0146	0.0726	0.1290	0.0142	0.0666	0.1220	0.0142	0.0666
		S_R	0.0294	0.1032	0.1704	0.0188	0.0706	0.1266	0.0140	0.0646	0.1284	0.0140	0.0646
		S_T	0.0066	0.0446	0.0944	0.0094	0.0458	0.0958	0.0084	0.0460	0.100	0.0084	0.0460
	0.75	S_W	0.0184	0.0614	0.1066	0.0128	0.0536	0.1046	0.0142	0.0504	0.0998	0.0142	0.0504
		S_{LR}	0.0212	0.0764	0.1332	0.0166	0.0700	0.1320	0.0172	0.0728	0.1390	0.0172	0.0728
		S_R	0.0496	0.1384	0.2166	0.0284	0.1004	0.1636	0.0190	0.0610	0.1178	0.0190	0.0610
		S_T	0.0074	0.0436	0.1002	0.0106	0.0450	0.1010	0.0092	0.0470	0.1008	0.0092	0.0470
0.5	0.25	S_W	0.0170	0.0544	0.0990	0.0124	0.0510	0.0994	0.0094	0.0492	0.1022	0.0094	0.0492
		S_{LR}	0.0192	0.0798	0.1410	0.0166	0.0662	0.1226	0.0142	0.0710	0.1230	0.0142	0.0710
		S_R	0.0664	0.1646	0.2450	0.0228	0.0814	0.1430	0.0134	0.0596	0.1160	0.0134	0.0596
		S_T	0.0074	0.0464	0.0980	0.0102	0.0474	0.0978	0.0086	0.0470	0.1002	0.0086	0.0470
	0.5	S_W	0.0128	0.0568	0.1044	0.0114	0.0536	0.1014	0.0104	0.0514	0.1002	0.0104	0.0514
		S_{LR}	0.0250	0.0832	0.1364	0.0166	0.0656	0.1252	0.0144	0.0638	0.1182	0.0144	0.0638
		S_R	0.0266	0.0994	0.1678	0.0176	0.0668	0.1266	0.0138	0.0638	0.1202	0.0138	0.0638
		S_T	0.0068	0.0446	0.0944	0.0094	0.0458	0.0956	0.0084	0.0458	0.0996	0.0084	0.0458

Table 15 continued

α	q	$n = 50$			$n = 100$			$n = 200$			
		1%	5%	10%	1%	5%	10%	1%	5%	10%	
1	0.75	S_W	0.0154	0.0636	0.1052	0.0136	0.0532	0.1004	0.0118	0.0530	0.0996
		S_{LR}	0.0166	0.0746	0.1338	0.0154	0.0782	0.1350	0.0166	0.0728	0.1300
		S_R	0.0406	0.1244	0.1922	0.0258	0.0834	0.1410	0.0130	0.0594	0.1104
		S_T	0.0076	0.0430	0.1008	0.0108	0.0444	0.0992	0.0088	0.0468	0.1012
	0.25	S_W	0.0128	0.0452	0.0866	0.0108	0.0460	0.0862	0.0084	0.0448	0.0954
		S_{LR}	0.0170	0.0710	0.1252	0.0150	0.0642	0.1160	0.0146	0.0632	0.1136
		S_R	0.0700	0.1586	0.2410	0.0224	0.0798	0.1482	0.0142	0.0668	0.1256
		S_T	0.0074	0.0460	0.0988	0.0104	0.0472	0.0980	0.0088	0.0462	0.1016
	0.5	S_W	0.0122	0.0478	0.0908	0.0130	0.0470	0.0938	0.0104	0.0488	0.0944
		S_{LR}	0.0206	0.0772	0.1318	0.0138	0.0622	0.1220	0.0144	0.0578	0.1072
		S_R	0.0292	0.1054	0.1740	0.0186	0.0718	0.1276	0.0140	0.0644	0.1208
		S_T	0.0068	0.0444	0.0948	0.0094	0.0462	0.0962	0.0084	0.0454	0.1004
0.75	S_W	0.0194	0.0534	0.0922	0.0106	0.0486	0.0888	0.0092	0.0514	0.1010	
	S_{LR}	0.0208	0.0726	0.1276	0.0138	0.0622	0.1150	0.0158	0.0642	0.1130	
	S_R	0.0454	0.1348	0.2070	0.0208	0.0866	0.1506	0.0126	0.0592	0.1128	
	S_T	0.0078	0.0430	0.1012	0.0110	0.0462	0.1022	0.0086	0.0464	0.1014	

Table 16 Null rejection rates for $H_0 : \beta_1 = \beta_2 = \beta_3 = 0$ in the SL-BS quantile regression model ($\nu = 4$)

α	q	$n = 50$			$n = 100$			$n = 200$		
		1%		5%	10%		5%	10%		5%
		1%	5%	10%	1%	5%	10%	1%	5%	10%
0.2	0.25	S_W	0.0146	0.0584	0.1000	0.0118	0.0506	0.0974	0.0096	0.0468
		S_{LR}	0.0160	0.0658	0.1250	0.0130	0.0616	0.1170	0.0126	0.0542
		S_R	0.0446	0.1258	0.2072	0.0222	0.0892	0.1520	0.0170	0.0702
		S_T	0.0072	0.0436	0.0916	0.0078	0.0436	0.0960	0.0090	0.0524
	0.5	S_W	0.0148	0.0496	0.0902	0.0118	0.0506	0.0974	0.0096	0.0468
		S_{LR}	0.0154	0.0666	0.1284	0.0164	0.0632	0.1190	0.0118	0.0548
		S_R	0.0664	0.1626	0.2434	0.0188	0.0736	0.1406	0.0148	0.0664
		S_T	0.0066	0.0426	0.0914	0.0082	0.0494	0.1004	0.0108	0.0450
	0.75	S_W	0.0150	0.0496	0.0902	0.0118	0.0508	0.0974	0.0096	0.0468
		S_{LR}	0.0150	0.0676	0.1236	0.0118	0.0574	0.1106	0.0104	0.0572
		S_R	0.0422	0.1240	0.1978	0.0260	0.0882	0.1554	0.0176	0.0640
		S_T	0.0070	0.0474	0.1028	0.0062	0.0458	0.0926	0.0086	0.0482
0.5	0.25	S_W	0.0136	0.0446	0.0826	0.0114	0.0470	0.0916	0.0096	0.0454
		S_{LR}	0.0160	0.0708	0.1290	0.0152	0.0544	0.1066	0.0102	0.0500
		S_R	0.0418	0.1262	0.2064	0.0270	0.0890	0.1538	0.0158	0.0684
		S_T	0.0058	0.0424	0.0950	0.0078	0.0488	0.0974	0.0094	0.0486
	0.5	S_W	0.0136	0.0446	0.0826	0.0114	0.0470	0.0916	0.0096	0.0454
		S_{LR}	0.0144	0.0666	0.1210	0.0134	0.0584	0.1090	0.0130	0.0566
		S_R	0.0436	0.1268	0.2104	0.0332	0.1036	0.1690	0.0130	0.0602
		S_T	0.0042	0.0438	0.0994	0.0104	0.0456	0.0928	0.0112	0.0476

Table 16 continued

α	q	$n = 50$			$n = 100$			$n = 200$			
		1%	5%	10%	1%	5%	10%	1%	5%	10%	
1	0.75	S_W	0.0136	0.0446	0.0824	0.0114	0.0470	0.0916	0.0096	0.0454	0.0902
		S_{LR}	0.0156	0.0608	0.1194	0.0124	0.0574	0.1072	0.0120	0.0566	0.1064
		S_R	0.0366	0.1098	0.1852	0.0260	0.0896	0.1560	0.0178	0.0694	0.1288
		S_T	0.0062	0.0460	0.0984	0.0070	0.0408	0.0944	0.0094	0.0562	0.1030
	0.25	S_W	0.0088	0.0310	0.0608	0.0102	0.0366	0.0738	0.0088	0.0394	0.0834
		S_{LR}	0.0130	0.0618	0.1130	0.0090	0.0504	0.0992	0.0100	0.0450	0.0946
		S_R	0.0450	0.1288	0.2100	0.0270	0.0942	0.1568	0.0122	0.0638	0.1268
		S_T	0.0080	0.0478	0.0928	0.0080	0.0508	0.1056	0.0112	0.0486	0.0966
	0.5	S_W	0.0088	0.0308	0.0608	0.0102	0.0366	0.0738	0.0088	0.0394	0.0834
		S_{LR}	0.0130	0.0578	0.1120	0.0114	0.0500	0.0992	0.0084	0.0428	0.0934
		S_R	0.0338	0.1080	0.1830	0.0206	0.0762	0.1396	0.0114	0.0602	0.1142
		S_T	0.0064	0.0460	0.0974	0.0076	0.0470	0.0974	0.0088	0.0502	0.1040
0.75	S_W	0.0088	0.0308	0.0608	0.0102	0.0366	0.0738	0.0088	0.0394	0.0834	
	S_{LR}	0.0144	0.0594	0.1148	0.0084	0.0468	0.0894	0.0094	0.0528	0.0978	
	S_R	0.0586	0.1494	0.2310	0.0216	0.0826	0.1496	0.0128	0.0644	0.1218	
	S_T	0.0054	0.0436	0.0916	0.0088	0.0468	0.0934	0.0086	0.0498	0.0994	

Table 17 Null rejection rates for $H_0 : \beta_1 = \beta_2 = \beta_3 = 0$ in the t_ν -BS quantile regression model ($\nu = 11$)

α	q	$n = 50$				$n = 100$				$n = 200$			
		1%		5%		1%		5%		1%		5%	
			10%				10%				10%		10%
0.2	0.25	S_W	0.0160	0.0630	0.1128	0.0112	0.0522	0.0982	0.0104	0.0494	0.0984	0.0494	0.0984
		S_{LR}	0.0178	0.0712	0.1304	0.0156	0.0668	0.1214	0.0138	0.0604	0.1164	0.0604	0.1164
		S_R	0.0452	0.1308	0.2092	0.0154	0.0704	0.1290	0.0150	0.0692	0.1254	0.0692	0.1254
		S_T	0.0086	0.0498	0.1016	0.0068	0.0430	0.0918	0.0082	0.0478	0.0962	0.0478	0.0962
	0.5	S_W	0.0156	0.0662	0.1232	0.0160	0.0566	0.1074	0.0120	0.0538	0.1060	0.0538	0.1060
		S_{LR}	0.0192	0.0790	0.1418	0.0146	0.0660	0.1240	0.0130	0.0580	0.1102	0.0580	0.1102
		S_R	0.0422	0.1188	0.1908	0.015	0.0662	0.1230	0.0122	0.0568	0.1126	0.0568	0.1126
		S_T	0.0058	0.0394	0.0886	0.0094	0.0454	0.0934	0.0096	0.0506	0.1020	0.0506	0.1020
0.5	0.75	S_W	0.0214	0.0674	0.1206	0.0138	0.0572	0.1084	0.0116	0.0580	0.1032	0.0580	0.1032
		S_{LR}	0.0180	0.0676	0.1364	0.0146	0.0594	0.1156	0.0138	0.0630	0.1176	0.0630	0.1176
		S_R	0.0604	0.1486	0.2198	0.0168	0.0668	0.1270	0.0156	0.0618	0.1204	0.0618	0.1204
		S_T	0.0092	0.0522	0.1064	0.0086	0.0486	0.0982	0.0092	0.0452	0.0950	0.0452	0.0950
	0.25	S_W	0.0144	0.0612	0.1110	0.0112	0.0492	0.0958	0.0110	0.0554	0.1026	0.0554	0.1026
		S_{LR}	0.0150	0.0678	0.1310	0.0164	0.0674	0.1220	0.0112	0.0576	0.1062	0.0576	0.1062
		S_R	0.0432	0.1368	0.2080	0.0194	0.0806	0.1418	0.0164	0.0668	0.1248	0.0668	0.1248
		S_T	0.0086	0.0502	0.1046	0.0068	0.0438	0.0916	0.0080	0.0454	0.0922	0.0454	0.0922
0.5	0.5	S_W	0.0188	0.0614	0.1116	0.0128	0.0544	0.0970	0.0106	0.0518	0.0996	0.0518	0.0996
		S_{LR}	0.0156	0.0694	0.1284	0.0140	0.0586	0.1118	0.0108	0.0626	0.1174	0.0626	0.1174
		S_R	0.0410	0.1124	0.1846	0.0210	0.0758	0.1430	0.0124	0.0608	0.1186	0.0608	0.1186
		S_T	0.0060	0.0456	0.0988	0.0094	0.0500	0.1024	0.0066	0.0452	0.0948	0.0452	0.0948

Table 17 continued

α	q	$n = 50$			$n = 100$			$n = 200$			
		1%	5%	10%	1%	5%	10%	1%	5%	10%	
1	0.75	S_W	0.0172	0.0548	0.1032	0.0100	0.0566	0.1038	0.0108	0.0494	0.0972
		S_{LR}	0.0144	0.0720	0.1354	0.0142	0.0590	0.1112	0.0126	0.0624	0.1218
		S_R	0.0474	0.1302	0.2032	0.0180	0.0736	0.1328	0.0138	0.0624	0.1186
		S_T	0.0084	0.0496	0.1016	0.0086	0.0436	0.0912	0.0084	0.0490	0.0946
	0.25	S_W	0.0118	0.0478	0.0898	0.0118	0.0420	0.0864	0.0106	0.0446	0.0874
		S_{LR}	0.0138	0.0604	0.1128	0.0134	0.0580	0.1058	0.0138	0.0550	0.1016
		S_R	0.0544	0.1374	0.2124	0.0156	0.0708	0.1356	0.0144	0.0638	0.1252
		S_T	0.0086	0.0514	0.1050	0.0068	0.0440	0.0916	0.0098	0.0544	0.1016
0.5	S_W	0.0102	0.0430	0.0848	0.0102	0.0480	0.0940	0.0100	0.0452	0.0936	
	S_{LR}	0.0160	0.0664	0.1232	0.0136	0.0552	0.1074	0.0116	0.0498	0.0970	
	S_R	0.0196	0.0844	0.1572	0.0122	0.0698	0.1370	0.0134	0.0592	0.1218	
	S_T	0.0080	0.0452	0.0998	0.0086	0.0508	0.0974	0.0090	0.0472	0.0960	
0.75	S_W	0.0144	0.0484	0.0882	0.0094	0.0464	0.0940	0.0110	0.0482	0.0974	
	S_{LR}	0.0178	0.0696	0.1236	0.0096	0.0526	0.1000	0.0122	0.0528	0.1074	
	S_R	0.0416	0.1196	0.1964	0.0224	0.0820	0.1416	0.0114	0.0592	0.1132	
	S_T	0.0072	0.0490	0.0982	0.0082	0.0518	0.1028	0.0110	0.0510	0.0990	

Acknowledgements Helton Saulo and Roberto Vila gratefully acknowledge financial support from the University of Brasilia (UnB), CNPq, CAPES, and FAP-DF, Brazil.

Declarations

Conflict of interest There are no conflicts of interest to disclose.

References

- Andrews DF, Mallows CL (1974) Scale mixtures of normal distributions. *J R Stat Soc Ser B (Methodological)* 36(1):99–102
- Baker SG (1992) A simple method for computing the observed information matrix when using the em algorithm with categorical data. *J Comput Graph Stat* 1(1):63–76
- Balakrishnan N, Kundu D (2019) Birnbaum-saunders distribution: A review of models, analysis, and applications. *Appl Stochast Models Business Indust* 35:4–49
- Balakrishnan N, Leiva V, Sanhueza A, Vilca F, et al (2009) Estimation in the birnbaum-saunders distribution based on scale-mixture of normals and the em-algorithm. *SORT-Stat Oper Res Trans*
- Cox DR, Snell E (1968) A general definition of residuals. *J R Stat Soc B* 2:248–275
- Dasilva A, Dias R, Leiva V, Marchant C, Saulo H (2020). [Invited tutorial] Birnbaum-Saunders regression models: a comparative evaluation of three approaches. *J Stat Comput Simul.* pages 1–19
- De Jong P, Heller GZ, et al (2008) Generalized linear models for insurance data. Cambridge Books
- Dempster AP, Laird NM, Rubin DB (1977) Maximum likelihood from incomplete data via the em algorithm. *J R Stat Soc Ser B* 39(1):1–22
- Dunn P, Smyth G (1996) Randomized quantile residuals. *J Comput Graph Stat* 5:236–244
- Efron B, Olshen RA (1978) How broad is the class of normal scale mixtures? *Annal Stat.* pages 1159–1164
- Fernández C, Steel MF (1999) Multivariate student-t regression models: Pitfalls and inference. *Biometrika* 86(1):153–167
- Gneiting T (1997) Normal scale mixtures and dual probability densities. *J Stat Comput Simul* 59(4):375–384
- Huang S, Qu Y (2006) The loss in power when the test of differential expression is performed under a wrong scale. *J Comput Biol* 13:786–797
- Jamshidian M, Jennrich RI (1997) Acceleration of the em algorithm by using quasi-newton methods. *J R Stat Soc Ser B (Methodological)* 59(3):569–587
- Koenker R (2004) Quantile regression for longitudinal data. *J Multivar Anal* 91(1):74–89
- Koenker R, Bassett Jr G (1978) Regression quantiles. *Econ J Econ Soc* pages 33–50
- Koenker R, Hallock KF (2001) Quantile regression. *J Econ Perspect* 15(4):143–156
- Lachos VH, Dey DK, Cancho VG, Louzada F (2017) Scale mixtures log-birnbaum-saunders regression models with censored data: a bayesian approach. *J Stat Comput Simul* 87(10):2002–2022
- Lachos VH, Labra F (2007) Skew-normal/independent distributions, with applications. Universidade Estadual de Campinas, Instituto de Matemática, Estatística e
- Lange K, Sinsheimer JS (1993) Normal/independent distributions and their applications in robust regression. *J Comput Graph Stat* 2(2):175–198
- Lange KL, Little RJ, Taylor JM (1989) Robust statistical modeling using the t distribution. *J Am Stat Assoc* 84(408):881–896
- Lee ET, Wang J (2003) Statistical methods for survival data analysis, vol 476. John Wiley & Sons
- Leiva V (2016) The Birnbaum-Saunders Distribution. Academic Press, New York, US
- Louis TA (1982) Finding the observed information matrix when using the em algorithm. *J R Stat Soc Ser B (Methodological)* 44(2):226–233
- Lucas A (1997) Robustness of the student t based m-estimator. *Commun Stat Theo Methods* 26(5):1165–1182
- McLachlan GJ, Krishnan T (2007) The EM algorithm and extensions, vol 382. John Wiley & Sons
- Meilijson I (1989) A fast improvement to the em algorithm on its own terms. *J R Stat Soc Ser B (Methodological)* 51(1):127–138
- Naderi M, Hashemi F, Bekker A, Jamalizadeh A (2020) Modeling right-skewed financial data streams: A likelihood inference based on the generalized Birnbaum-Saunders mixture model. *Appl Math Comput* 376:125109
- Oakes D (1999) Direct calculation of the information matrix via the em. *J R Stat Soc Ser B* 61(2):479–482
- Paula GA, Leiva V, Barros M, Liu S (2012) Robust statistical modeling using the birnbaum-saunders-t distribution applied to insurance. *Appl Stochast Models Business Indust* 28(1):16–34

- R Core Team (2020) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria
- Rieck JR, Nedelman JR (1991) A log-linear model for the birnbaum-saunders distribution. *Technometrics* 33(1):51–60
- Rothenberg T (1971) Identification in parametric models. *Econometrica* 39:577–591
- Sánchez L, Leiva V, Galea M, Saulo H (2020) Birnbaum-Saunders quantile regression and its diagnostics with application to economic data. *Appl Stochast Models Business Indust* 37:53–73
- Saulo H, Dasilva A, Leiva V, Sánchez L, Fuente-Mella HL (2022) Log-symmetric quantile regression models. *Stat Neerl* 76:124–163
- Saulo H, Leão J, Leiva V, Aykroyd RG (2019) Birnbaum-Saunders autoregressive conditional duration models applied to high-frequency financial data. *Stat Pap* 60:1605–1629
- Saulo H, Leão J, Nobre L, Balakrishnan N (2021) A class of asymmetric regression models for left-censored data. *Braz J Probab Stat* 35:62–84
- Sánchez L, Leiva V, Saulo H, Marchant C, Sarabia JM (2021) A new quantile regression model and its diagnostic analytics for a weibull distributed response with applications. *Mathematics* 9:2768
- Taylor J, Verbyla A (2004) Joint modelling of location and scale parameters of the t distribution. *Stat Modell* 4(2):91–112
- Terrell GR (2002) The gradient statistic. *Comput Sci. Stat* 34(34):206–215
- Walker SG, Gutiérrez-Peña E (2007) Bayesian parametric inference in a nonparametric framework. *Test* 16(1):188–197
- West M (1987) On scale mixtures of normal distributions. *Biometrika* 74(3):646–648
- Yu K, Moyeed RA (2001) Bayesian quantile regression. *Statistics & Probab Lett* 54(4):437–447

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.