Statistics and Its Interface Volume 9 (2016) 281–293

Nonlinear censored regression models with heavy-tailed distributions

Aldo M. Garay*, Victor H. Lachos, and Tsung-I Lin

In the framework of censored nonlinear regression models, the random errors are routinely assumed to have a normal distribution, mainly for mathematical convenience. However, this method has been criticized in the literature due to its sensitivity to deviations from the normality assumption. In practice, data such as income or viral load in AIDS studies, often violate this assumption because of heavy tails. In this paper, we establish a link between the censored nonlinear regression model and a recently studied class of symmetric distributions, which extends the normal one by the inclusion of kurtosis, called scale mixtures of normal (SMN) distributions. The Student-t, Pearson type VII, slash and contaminated normal, among others distributions, are contained in this class. Choosing a member of this class can be a good alternative to model this kind of data, because they have been shown its flexibility in several applications. We develop an analytically simple and efficient EM-type algorithm for iteratively computing maximum likelihood estimates of model parameters together with standard errors as a by-product. The algorithm uses nice expressions at the E-step, relying on formulae for the mean and variance of truncated SMN distributions. The usefulness of the proposed methodology is illustrated through applications to simulated and real data.

Keywords and phrases: Censored nonlinear regression model, EM-type algorithms, Scale mixtures of normal distributions, Outliers.

1. INTRODUCTION

Nonlinear (NL) regression models with censored dependent variable (hereafter NLCR models) are applied in many fields, like econometric analysis, clinical essays, medical surveys, engineering studies, among others. For example, in medical surveys, the relationship between the survival time and the age of a patient who has received a given treatment is often nonlinear, and the survival time is subject to right censoring because the patient may decide to leave the study. As such, he/she may die due to another cause than the disease from which he/she suffers, or the study itself can be stopped. In AIDS research, the viral load measures may be subjected to some upper and lower detection limits, namely below or above which they are not quantifiable. As a result, the viral load responses are either left or right censored depending on the diagnostic assays used (Vaida and Liu, 2009).

In general, for mathematical tractability reasons, it is assumed that the random errors in nonlinear models have a normal distribution, see Wei and Tanner (1990) and Heuchenne and Van Keilegom (2007). However, it is well-known that several phenomena are not always in agreement with this assumption, yielding data with a distribution with heavier tails. The problem of longer-than-normal tails (or outliers) can be circumvented by data transformations (namely, Box–Cox, etc.), which can render approximate normality with reasonable empirical results. However, some possible drawbacks of these methods are: (i) transformations provide reduced information on the underlying data generation scheme; (ii) component wise transformations may not guarantee joint normality; (iii) parameters may lose interpretability on a transformed scale and (iv) transformations may not be universal and usually vary with the data set. Hence, from a practical perspective, there is a necessity to seek an appropriate theoretical model that avoids data transformations, yet presenting a robustified "Gaussian" framework.

To deal with the problem of atypical observations in NL regression models, proposals have been made in the literature to replace normality with more flexible classes of distributions. For instance, Cysneiros and Vanegas (2008) studied the symmetrical NL regression model and performed an analytical and empirical study to describe the behavior of the standardized residuals. Vanegas and Cysneiros (2010) proposed diagnostic procedures based on case-deletion model for symmetrical NL regression models. Cancho et al. (2010) introduced the skew-normal (Azzalini, 1985) NL regression model and presented a complete likelihood based analysis, including an efficient EM algorithm for maximum likelihood (ML) estimation. A common feature of these classes of nonlinear models is that the Gaussian nonlinear model is a particular member of the class. Although there are some proposals that overcome the problem of atypical observations in NL regression models, there are no studies taking into account, at the same time, censored responses and observational errors modeled by a distribution in the scale mixture of normal (SMN) class, which is, maybe,

*Corresponding author.

the most important family of symmetric distributions. The SMN distributions are extensions of the normal one, while incorporating kurtosis. The Student-t (T), Pearson type VII (PVII), slash (SL), power exponential (PE), contaminated normal (CN) and, obviously, the normal (N) distributions are included in this class. Comprehensive surveys are available in Fang and Zhang (1990), Liu (1996), Meza et al. (2012), among others.

In this paper, we extend the NLCR model, called the SMN-NLCR model, by assuming a SMN distribution for the errors. A computationally feasible EM-type algorithm is carried out for ML estimation. We show that the E-step reduces to computing the first two moments of certain truncated SMN distributions. The general formulas for these moments were derived in closed form by Genç (2013). The likelihood function and the asymptotic standard errors are easily computed as a by-product of the E-step and are used for monitoring convergence and for model selection using the Akaike information criterion (AIC) or the Bayesian information criterion (BIC). The theoretical justification of the proposal rests on the facts that the SMN class stochastically attributes varying weights to each subject, i.e., lower weight for outliers and thus controls the influence of atypical observations on the overall inference. Moreover, every member of the SMN class tends to the normal case, for example, as the Student-t degrees of freedom tends to the infinity, it approaches normality.

The rest of the paper is organized as follows. Section 2 briefly outlines some preliminary properties of the SMN and truncated SMN distributions. Section 3 discusses the specification of the SMN-NLCR model and presents the Expectation Conditional Maximization Either (ECME) algorithm (Liu and Rubin, 1994) for obtaining the ML estimates of parameters. In Section 4, we derive approximate standard errors for the regression parameters of the SMN-NLCR model. Section 5, presents a simulation study to compare the performance of our methods with other normality-based methods. In Section 6, advantages of the proposed methodology is illustrated through the analysis of a real data set, previously analyzed under normal errors. Section 7 concludes with a short discussion on the issues raised by our study and some possible directions for future research.

2. PRELIMINARIES

Throughout this paper, $X \sim \mathrm{N}(\mu, \sigma^2)$ denotes that a random variable X follows the normal distribution with mean μ and variance σ^2; $\phi(\cdot|\mu, \sigma^2)$ stands for its probability density function (pdf); $\phi(\cdot)$ and $\Phi(\cdot)$ represent the pdf and the cumulative distribution function (cdf) of the standard normal distribution, respectively. Moreover, we follow the traditional convention to indicate a random variable (or a random vector) by an upper case letter and its realization by the corresponding lower case. Random vectors and matrices are denoted by boldface letters. The symbol $\mathbf{X}^\top$ means the transpose of $\mathbf{X}$ and '$X \perp Y$' indicates that the random variables X and Y are independent. We start by defining the SMN distributions and its hierarchical formulation and then introduce some further properties. A random variable X is said to follow a SMN distribution with location parameter μ and scale parameter $\sigma^2 > 0$ if it has the following stochastic representation:

$$X = \mu + U^{-1/2} Z, \quad Z \perp U, \tag{1}$$

where $Z \sim N(0, \sigma^2)$ and U is a positive random variable with cdf $H(\cdot|\boldsymbol{\nu})$. We use the notation $X \sim \mathrm{SMN}(\mu, \sigma^2, \boldsymbol{\nu})$. When $\mu = 0$ and $\sigma^2 = 1$, we have the so-called standard SMN distribution. Note from (1) that $X|U = u \sim N(\mu, u^{-1}\sigma^2)$.

$$\begin{aligned} &f_{SMN}\left(x|\mu, \sigma^2, \boldsymbol{\nu}\right) \\ &= (2\pi\sigma^2)^{-1/2} \int_0^\infty u^{1/2} \exp\left\{-(u/2\sigma^2)(x-\mu)^2\right\} dH\left(u|\boldsymbol{\nu}\right), \end{aligned} \tag{2}$$

where $-\infty < x < \infty$. The form of the SMN distribution is determined by the distribution of U. Herein, U is called the scale factor and $H(\cdot|\boldsymbol{\nu})$ is called the mixture distribution.

It is important to notice that there exists a relationship between the SMN and the elliptical distributions. We say that the random variable X has a univariate elliptical distribution with location parameter μ and scale parameter σ^2, when its density is given by

$$f(x) = \sigma^{-1} g\left(z\right), \tag{3}$$

where $z = (x-\mu)^2/\sigma^2$ and $g : \mathbb{R} \to [0, \infty)$ satisfies $\int_0^\infty z^{-1/2} g(z) dz < \infty$. It is easy to see that (2) has the form of (3). To obtain the standard errors of the regression parameters, the relation between SMN and elliptical distributions will be used in Section 4. Let $X \sim \mathrm{SMN}(\mu, \sigma^2, \boldsymbol{\nu})$ and $a < b$ such that $P(a < X < b) > 0$. A random variable Y has a truncated SMN distribution within the interval (a, b). In this case, we write $Y \sim \mathrm{TSMN}_{(a,b)}(\mu, \sigma^2, \boldsymbol{\nu})$ if it has the same distribution as $X|X \in (a, b)$. Thus, the density of Y is

$$\begin{aligned} &f_{\mathrm{TSMN}}(y|\mu, \sigma^2, \boldsymbol{\nu}; (a, b)) \\ &= f_{SMN}(y|\mu, \sigma^2, \boldsymbol{\nu}) \\ &\quad \times \left[F_{SMN}\left(\frac{b-\mu}{\sigma}\right) - F_{SMN}\left(\frac{a-\mu}{\sigma}\right)\right]^{-1}, \end{aligned} \tag{4}$$

where $a < y < b$ and $f_{\mathrm{TSMN}}(y|\mu, \sigma^2, \boldsymbol{\nu}; (a, b)) = 0$ otherwise, and $F_{SMN}(\cdot)$ denotes the cdf of the standard SMN distribution. Next, we establish the following proposition, which is crucial for the development of our proposed theory. This proposition is a natural extension of Theorem 1 (and Corollary 1) of Genç (2013). The proof is given in Appendix A.

Proposition 1. *Let* $X \sim \mathrm{SMN}(0,1)$ *with scale factor* U *and mixture distribution* $H(\cdot|\boldsymbol{\nu})$. *Then, for* $a<b$,

$$\begin{aligned}
&\mathrm{E}\left[U^r|X \in (a,b)\right] = \frac{1}{F_{SMN}(b) - F_{SMN}(a)} \\
&\times \left[\mathrm{E}_\Phi(r,b) - \mathrm{E}_\Phi(r,a)\right]; \\
&\mathrm{E}\left[U^r X|X \in (a,b)\right] = \frac{1}{F_{SMN}(b) - F_{SMN}(a)} \\
&\times \left[\mathrm{E}_\phi\left(r-\frac{1}{2},a\right) - \mathrm{E}_\phi\left(r-\frac{1}{2},b\right)\right]; \\
&\mathrm{E}\left[U^r X^2|X \in (a,b)\right] = \frac{1}{F_{SMN}(b) - F_{SMN}(a)} \\
&\times \Bigg[\mathrm{E}_\Phi(r-1,b) - \mathrm{E}_\Phi(r-1,a) + a\mathrm{E}_\phi\left(r-\frac{1}{2},a\right) \\
&- b\mathrm{E}_\phi\left(r-\frac{1}{2},b\right)\Bigg],
\end{aligned}$$

where

$$\mathrm{E}_\phi(r,h) = \mathrm{E}\left[U^r \phi\left(h\,U^{1/2}\right)\right], \tag{5}$$

and

$$\mathrm{E}_\Phi(r,h) = \mathrm{E}\left[U^r \Phi\left(h\,U^{1/2}\right)\right]. \tag{6}$$

When the distribution of U is available, this proposition gives closed form expressions for the expected values $\mathrm{E}[U^r X^s|X \in (a,b)]$, where $s=0,1,2$ and $r \geq 1$. Next, we compute the quantities $E_\phi(r,h)$ and $E_\Phi(r,h)$ for some elements of the SMN family. They are useful in the implementation of the ECME algorithm.

- *Pearson type VII distribution:* Consider $U \sim Gamma(\nu/2, \delta/2)$, with $\nu>0$ and $\delta>0$, where $Gamma(a,b)$ denotes the Gamma distribution with mean a/b. From (1), the density of X takes the form of

$$f_{PVII}(x|\nu,\delta) = \frac{1}{B(\nu/2,1/2)\sqrt{\delta}}\left(1+\frac{x^2}{\delta}\right)^{-\frac{\nu+1}{2}}, \tag{7}$$

where $x \in \mathbb{R}$, and $\delta>0$ and $\nu>0$ are shape parameters and $B(a,b)$ represents the beta function. We use the notation $X \sim PVII(0,1;\nu,\delta)$ to denote X has density (7). Moreover, we obtain

$$\mathrm{E}_\Phi(r,h) = \frac{\Gamma\left(\frac{\nu+2r}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)}\left(\frac{\delta}{2}\right)^{-r} F_{PVII}(h|\nu+2r,\delta) \quad \text{and}$$

$$\mathrm{E}_\phi(r,h) = \frac{\Gamma\left(\frac{\nu+2r}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)\sqrt{2\pi}}\left(\frac{\delta}{2}\right)^{\nu/2}\left(\frac{h^2+\delta}{2}\right)^{-\frac{(\nu+2r)}{2}}$$

where $\Gamma(a)$ is the Gamma function and $F_{PVII}(\cdot)$ is the cdf of the Pearson type VII distribution. Note that the Pearson type VII distribution reduces to the Student-t distribution with ν degrees of freedom when $\delta=\nu$ and the Cauchy distribution when $\delta=\nu=1$.

- *Slash distribution:* Consider $U \sim Beta(\nu,1)$ with positive shape parameter ν. It follows from (1) that the density of X is given by

$$f_{sl}(x|\nu) = \nu\int_0^1 u^{\nu-1}\phi(x\sqrt{u})du,\ x \in \mathbb{R}. \tag{8}$$

The notation $X \sim SL(0,1;\nu)$ indicates X has density (8). Moreover, we have

$$\mathrm{E}_\Phi(r,h) = \left(\frac{\nu}{\nu+r}\right) F_{SL}(h|\nu+r) \quad \text{and}$$

$$\mathrm{E}_\phi(r,h) = \frac{\nu}{\sqrt{2\pi}}\left(\frac{h^2}{2}\right)^{-(\nu+r)}\Gamma\left(\nu+r,\frac{h^2}{2}\right),$$

where $\Gamma(a,b)=\int_0^b e^{-t}t^{a-1}dt$ is the incomplete gamma function, see Lemma 6 in Genç (2013), and $F_{SL}(\cdot)$ is the cdf of the slash distribution.

- *Contaminated normal distribution:* Let U be a discrete random variable taking one of two states 1 or γ. In this case, the probability function of U is

$$h(u|\nu,\gamma) = \nu\mathbb{I}_\gamma(u) + (1-\nu)\mathbb{I}_1(u),\ \ \nu,\gamma \in (0,1),$$

where $\mathbb{I}_B(\cdot)$ is the indicator function of the set B. It follows immediately that the density of X is

$$f_{CN}(x|\nu,\gamma) = \nu\phi(x|0,\gamma^{-\frac{1}{2}}) + (1-\nu)\phi(x).$$

It follows that

$$\begin{aligned}
&\mathrm{E}_\Phi(r,h) = \gamma^r F_{CN}(h|\nu,\gamma) + (1-\gamma^r)(1-\nu)\Phi(h) \quad \text{and} \\
&\mathrm{E}_\phi(r,h) = \nu\gamma^r\phi(h\sqrt{\gamma}) + (1-\nu)\phi(h),
\end{aligned}$$

where $F_{CN}(\cdot)$ is the cdf of the contaminated normal distribution.

As a direct consequence of Proposition 1, we present an important corollary in Appendix. It is useful for parameter estimation in SMN-NLCR models via the ECME algorithm.

3. THE SMN CENSORED NONLINEAR REGRESSION MODEL

3.1 The model

Consider a nonlinear regression model where the responses are observed with errors which are independent and identically distributed according to the SMN distribution. The model considered is written as

$$Y_i = \eta(\mathbf{x}_i,\boldsymbol{\beta}) + \varepsilon_i,\quad \varepsilon_i \overset{\text{iid}}{\sim} \mathrm{SMN}(0,\sigma^2,\boldsymbol{\nu}),\quad i=1,\ldots,n, \tag{9}$$

where the Y_i is the response, $\eta(\mathbf{x}_i,\boldsymbol{\beta})$ is an injective and twice differentiable function with respect to the vector of regression parameters $\boldsymbol{\beta}=(\beta_1,\ldots,\beta_p)^\top$, the derivative matrix $\mathbf{D}_{i\beta}=\partial\eta_i(\boldsymbol{\beta})/\partial\boldsymbol{\beta}$ has rank p $(p<n)$ and $\mathbf{x}_i$ is a vector of explanatory variable for subject i. It follows from (1)

that $Y_i \stackrel{\text{ind}}{\sim} \text{SMN}(\eta(\mathbf{x}_i,\boldsymbol{\beta}),\sigma^2,\boldsymbol{\nu})$, for $i=1,\ldots,n$. We call (9) the SMN nonlinear regression (SMN-NLR) model. Following Vaida and Liu (2009), we consider the case in which the response Y_i is not fully observed for all i. Thus, let the observed data for the i-th subject be $(\mathbf{V}_i, C_i)$, where $\mathbf{V}_i$ represents the uncensored reading ($\mathbf{V}_i = V_{0i}$) or censoring interval ($\mathbf{V}_i = (V_{1i}, V_{2i})$), and C_i the censoring indicators:

$$V_{1i} \le Y_i \le V_{2i} \text{ if } C_i = 1, \text{ and } Y_i = V_{0i} \text{ if } C_i = 0, \tag{10}$$

so that the SMN-NLCR model is defined. Let $\boldsymbol{\theta} = (\beta^\top, \sigma^2, \boldsymbol{\nu})^\top$ be the vector containing all unknown parameters in the SMN-CR model; $\mathbf{V} = (\mathbf{V}_1^\top, \ldots, \mathbf{V}_n^\top)^\top$ and $\mathbf{C} = (C_1,\ldots,C_n)^\top$. Supposing that there are (possibly) m censored values of the characteristic of interest, we can partition the observed sample $\mathbf{y}_{\text{obs}}$ in two subsamples of m censored and $n-m$ uncensored values, such that $\mathbf{y}_{\text{obs}} = (\mathbf{C}^\top, \mathbf{V}^\top) = \{\mathbf{V}_1,\ldots,\mathbf{V}_m, y_{m+1},\ldots,y_n\}$. Then, the log-likelihood function is given by

$$\begin{aligned} \ell(\boldsymbol{\theta}|\mathbf{y}_{\text{obs}}) = & \sum_{i=1}^m \log\left[F_{SMN}\left(\frac{V_{2i}-\eta(\mathbf{x}_i,\boldsymbol{\beta})}{\sigma}\right)\right. \\ & - \left. F_{SMN}\left(\frac{V_{1i}-\eta(\mathbf{x}_i,\boldsymbol{\beta})}{\sigma}\right)\right] \\ & + \sum_{i=m+1}^n \log\left[f_{SMN}(y_i|\eta(\mathbf{x}_i,\boldsymbol{\beta}),\sigma^2,\boldsymbol{\nu})\right]. \end{aligned} \tag{11}$$

To estimate the parameters of the SMN-NLCR model, maximizing this log-likelihood function directly prevents the possibility of analytical solutions. One alternative is to maximize the complete likelihood using the EM algorithm (Dempster et al., 1977) or some other extensions like the ECM (Meng and Rubin, 1993) or the ECME algorithm (Liu and Rubin, 1994). We exploit the ECME algorithm for conducting ML estimation of the SMN-NLCR model because it is computationally more efficient than both EM and ECM algorithms. The Newton-Raphson procedure is not recommended for the implementation of the SMN-NLCR model because it is difficult to converge if the initial values are not very close to the ML estimates.

3.2 The ECME algorithm for the SMN-NLCR model

We develop an efficient ECME algorithm for maximum likelihood estimation of the parameters in the SMN-NLCR model. In so doing, we need a representation of the model in terms of missing data. In light of (1), the SMN-NLCR model has the following hierarchical representation:

$$Y_i|U_i = u_i \sim \text{N}\left(\eta(\mathbf{x}_i,\boldsymbol{\beta}), u_i^{-1}\sigma^2\right); \quad U_i \sim H(\cdot|\boldsymbol{\nu}). \tag{12}$$

If the observation i is censored, y_i can be viewed as a realization of the latent unobservable variable $Y_i \sim \text{SMN}(\eta(\mathbf{x}_i,\boldsymbol{\beta}),\sigma^2,\boldsymbol{\nu})$, $i=1,\ldots,m$. The key to the development of ECME algorithm is to consider the complete-data $\mathbf{z} = \{\mathbf{y}_{\text{obs}}, y_1,\ldots,y_m,u_1,\ldots,u_n\} = \{y_1,\ldots,y_n,u_1,\ldots,u_n\}$. From (12), the log-likelihood based on complete-data $\mathbf{z}$ is

$$\begin{aligned} \ell_c(\boldsymbol{\theta}|\mathbf{z}) \approx & -\frac{n}{2}\log\left(\sigma^2\right) \\ & + \frac{1}{2}\sum_{i=1}^n \log(u_i) - \frac{1}{2\sigma^2}\sum_{i=1}^n u_i(y_i-\eta(\mathbf{x}_i,\boldsymbol{\beta}))^2 \\ & + \sum_{i=1}^n \log(h(u_i|\boldsymbol{\nu})). \end{aligned} \tag{13}$$

In what follows the superscript (k) indicates the estimate of the related parameter at iteration k. In the E-step of the algorithm, we compute the Q-function, defined as

$$Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(k)}) = \text{E}_{\boldsymbol{\theta}^{(k)}}\left[\ell_c(\boldsymbol{\theta}|\mathbf{Z})|\mathbf{y}_{\text{obs}}\right],$$

where $\text{E}_{\boldsymbol{\theta}^{(k)}}[Y|X]$ denotes the conditional expectation of Y given X evaluated at $\boldsymbol{\theta} = \boldsymbol{\theta}^{(k)}$. It follows from (13) that the E-step involves the calculation of the following conditional expectations:

$$\mathcal{E}_{si}(\boldsymbol{\theta}^{(k)}) = \text{E}_{\boldsymbol{\theta}^{(k)}}[U_i Y_i^s|y_{\text{obs}_i}], \quad \text{E}_{\boldsymbol{\theta}^{(k)}}[\log(U_i)|y_{\text{obs}_i}] \quad \text{and}$$

$\text{E}_{\boldsymbol{\theta}^{(k)}}[\log(h(U_i|\boldsymbol{\nu}))|y_{\text{obs}_i}]$ for $s=0,1$ and 2.

Thus, dropping terms which are unrelated parameters, the Q-function takes the form of

$$\begin{aligned} Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(k)}) = & -\frac{n}{2}\log\left(\sigma^2\right) - \frac{1}{2\sigma^2}\sum_{i=1}^n\Big[\mathcal{E}_{2i}(\boldsymbol{\theta}^{(k)}) \\ & - 2\mathcal{E}_{1i}(\boldsymbol{\theta}^{(k)})\eta(\mathbf{x}_i,\boldsymbol{\beta}) + \mathcal{E}_{0i}(\boldsymbol{\theta}^{(k)})\eta(\mathbf{x}_i,\boldsymbol{\beta})^2\Big] \\ & + \frac{1}{2}\sum_{i=1}^n \text{E}_{\boldsymbol{\theta}^{(k)}}[\log(U_i)|\mathbf{V}_i, C_i] \\ & + \sum_{i=1}^n \text{E}_{\boldsymbol{\theta}^{(k)}}[\log(h(U_i|\boldsymbol{\nu}))|\mathbf{V}_i, C_i]. \end{aligned} \tag{14}$$

At each step, the conditional expectations $\mathcal{E}_{si}(\boldsymbol{\theta}^{(k)})$ can be easily derived from the results given in Proposition 1. Thus,

- *for an uncensored observation* i, we have

  $$\mathcal{E}_{si}(\boldsymbol{\theta}^{(k)}) = y_i^s \text{E}_{\boldsymbol{\theta}^{(k)}}[U_i|y_i], \tag{15}$$

  where $\text{E}_{\boldsymbol{\theta}^{(k)}}[U_i|y_i]$ can be obtained using results in Osorio et al. (2007). Thus, for example,

  - If $Y_i \sim PVII(\eta(\mathbf{x}_i,\boldsymbol{\beta}),\sigma^2,\nu,\delta)$, we have $\text{E}_{\boldsymbol{\theta}^{(k)}}[U_i|y_i] = (\nu+1)/(\delta + d(\boldsymbol{\theta}^{(k)}, y_i))$,

– If $Y_i \sim SL(\eta(\mathbf{x}_i, \boldsymbol{\beta}), \sigma^2, \nu)$, we have

$$\mathrm{E}_{\boldsymbol{\theta}^{(k)}}[U_i|y_i] = \frac{\Gamma\left(\nu + 1.5, d(\boldsymbol{\theta}^{(k)}, y_i)/2\right)}{\Gamma\left(\nu + 0.5, d(\boldsymbol{\theta}^{(k)}, y_i)/2\right)}$$

– If $Y_i \sim CN(\eta(\mathbf{x}_i, \boldsymbol{\beta}), \sigma^2, \nu, \gamma)$, we have

$$\mathrm{E}_{\boldsymbol{\theta}^{(k)}}[U_i|y_i] = \frac{\left(1 - \nu + \nu\gamma^{1.5} e^{0.5(1-\gamma)d(\boldsymbol{\theta}^{(k)}, y_i)}\right)}{\left(1 - \nu + \nu\gamma^{0.5} e^{0.5(1-\gamma)d(\boldsymbol{\theta}^{(k)}, y_i)}\right)},$$

where $d(\boldsymbol{\theta}^{(k)}, y_i) = (\frac{y_i - \eta(\mathbf{x}_i, \boldsymbol{\beta}^{(k)})}{\sigma^{(k)}})^2$,

- *for an interval censored observation* i, we have $Y_i \leq \kappa_i$, so that

$$(16) \qquad \mathcal{E}_{si}(\boldsymbol{\theta}^{(k)}) = \mathrm{E}_{\boldsymbol{\theta}^{(k)}}[U_i Y_i^s | V_{i1} \leq Y_i \leq V_{i2}],$$

  which can be obtained by using Proposition 1 along with the results of (5) and (6) with $r = 1$.

When the M-step turns out to be analytically intractable, it can be replaced with a sequence of conditional maximization (CM) steps. The resulting procedure in known as the *ECM algorithm* (Meng and Rubin, 1993). The *ECME algorithm* (Liu and Rubin, 1994) is a faster extension of EM and ECM algorithm by maximizing the constrained Q-function with some CM-steps that maximize the corresponding constrained actual marginal likelihood function, called the *CML-steps*. In summary, the ECME algorithm for estimating the parameters of the SMN-NLCR model can be proceeded in the following way:

E-step: Given $\boldsymbol{\theta} = \boldsymbol{\theta}^{(k)}$, for $i = 1, \ldots, n$;

– If observation i is uncensored, then for $s = 0, 1, 2$, compute $\mathcal{E}_{si}(\boldsymbol{\theta}^{(k)})$ given in (15);
– If observation i is censored, then for $s = 0, 1, 2$, compute $\mathcal{E}_{si}(\boldsymbol{\theta}^{(k)})$ given in (16).

CM-step: Update $\boldsymbol{\theta}^{(k)}$ by maximizing $Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(k)})$ over $\boldsymbol{\theta}$, which leads to the following expressions:

$$(17) \quad \widehat{\boldsymbol{\beta}}^{(k+1)} = \mathrm{argmin}_{\boldsymbol{\beta}}(\boldsymbol{\tau}^{(k)} - \boldsymbol{\eta}(\boldsymbol{\beta}, \mathbf{x}))^\top \widehat{\mathbf{U}}^{(k)} (\boldsymbol{\tau}^{(k)} - \boldsymbol{\eta}(\boldsymbol{\beta}, \mathbf{x})),$$

$$(18) \quad \widehat{\sigma^2}^{(k+1)} = \frac{1}{n}\sum_{i=1}^{n}\Big[\mathcal{E}_{2i}(\boldsymbol{\theta}^{(k)}) - 2\mathcal{E}_{1i}(\boldsymbol{\theta}^{(k)})\eta(\mathbf{x}_i, \boldsymbol{\beta}^{(k+1)}) + \mathcal{E}_{0i}(\boldsymbol{\theta}^{(k)})(\eta(\mathbf{x}_i, \boldsymbol{\beta}^{(k+1)}))^2\Big].$$

CML-step: Update $\nu^{(k)}$ by maximizing the actual marginal log-likelihood function, obtaining

$$(19) \quad \boldsymbol{\nu}^{(k+1)} = \mathrm{argmax}_{\nu}\Bigg\{\sum_{i=1}^{m}\log\Big[F_{SMN}\Big(\frac{V_{2i} - \eta(\mathbf{x}_i, \boldsymbol{\beta}^{(k+1)})}{\sigma^{(k+1)}}\Big) - F_{SMN}\Big(\frac{V_{1i} - \eta(\mathbf{x}_i, \boldsymbol{\beta}^{(k+1)})}{\sigma^{(k+1)}}\Big)\Big] + \sum_{i=m+1}^{n}\log[f_{SMN}(y_i|\eta(\mathbf{x}_i, \boldsymbol{\beta}^{(k+1)}), \sigma^{2(k+1)}, \boldsymbol{\nu})]\Bigg\},$$

where $\boldsymbol{\eta}(\boldsymbol{\beta}, \mathbf{x}) = (\eta(\boldsymbol{\beta}, \mathbf{x}_1), \ldots, \eta(\boldsymbol{\beta}, \mathbf{x}_n))^\top$, $\widehat{\mathbf{U}}^{(k)} = \mathrm{Diag}(\mathcal{E}_{01}(\boldsymbol{\theta}^{(k)}), \ldots, \mathcal{E}_{0n}(\boldsymbol{\theta}^{(k)}))$ and $\widehat{\boldsymbol{\tau}}^{(k)} = (\widehat{\tau}_1^{(k)}, \ldots, \widehat{\tau}_n^{(k)})^\top$ is the corrected observed response with $\widehat{\tau}_i^{(k)} = \mathcal{E}_{1n}(\boldsymbol{\theta}^{(k)})/\mathcal{E}_{0n}(\boldsymbol{\theta}^{(k)})$. Given a set of suitable initial values $\widehat{\boldsymbol{\theta}}^{(0)}$ described in the next subsection, the ECME procedure is performed iteratively until some distance involving two successive evaluations of the actual log-likelihood $\ell(\boldsymbol{\theta}|\mathbf{y}_{obs})$, like $||\ell(\boldsymbol{\theta}^{(k+1)}|\mathbf{y}_{obs}) - \ell(\boldsymbol{\theta}^{(k)}|\mathbf{y}_{obs})||$ or $||\ell(\boldsymbol{\theta}^{(k+1)}|\mathbf{y}_{obs})/\ell(\boldsymbol{\theta}^{(k)}|\mathbf{y}_{obs}) - 1||$, is small enough. We have adopted this strategy to update the estimate of $\boldsymbol{\nu}$ by directly maximizing the marginal log-likelihood. In this way, we circumvent the cumbersome problem of computing $\mathrm{E}_{\boldsymbol{\theta}^{(k)}}[\log(U_i)|y_{\mathrm{obs}_i}]$ and $\mathrm{E}_{\boldsymbol{\theta}^{(k)}}[\log(h(U_i|\boldsymbol{\nu}))|y_{\mathrm{obs}_i}]$. Upon convergence, the ML estimates of $\boldsymbol{\theta}$ is denoted by $\widehat{\boldsymbol{\theta}} = (\widehat{\boldsymbol{\beta}}, \widehat{\sigma}^2, \widehat{\boldsymbol{\nu}})$.

3.3 Notes on implementation

It is well known that ML estimation in nonlinear models may face some computational hurdles in the sense that the method may not give maximum global solutions when the starting values are far from the true parameter values. Thus, the choice of starting values for the EM algorithm in the non-linear context plays an important role in parameter estimation. In our example we consider the following procedure for the SMN-NLCR.

- Compute the estimates $\widehat{\boldsymbol{\beta}}^{(0)}$ and $(\widehat{\sigma}^2)^{(0)}$ using the nonlinear least squares (NLLS) method, which can be computed through the R function `nls()`. These values should be computed by considering the complete data, that is, censoring is not present in the data.
- We use the NLLS estimates of the regression parameter and scale parameters as initial values for the corresponding parameter under the N-NLCR, T-NLCR, SL-NLCR and the CN-NLCR models.
- In order to estimate the mixture parameter ν, we assuming 3, 3, and $(0.1, 0.1)$ as initial values for the T-NLCR, SL-NLCR and CN-NLCR respectively.

Notice that all the computational procedures were coded and implemented using the statistical software package R (R Core Team, 2013). The computer programs are available from the first author upon request.

4. STANDARD ERRORS ESTIMATES

Standard errors of the ML estimates can be approximated by the inverse of the observed information matrix, but there is generally no closed form, see Meilijson (1989) and Lin

(2010). Writing $\boldsymbol{\theta} = (\boldsymbol{\beta}, \sigma^2, \boldsymbol{\nu})$, the empirical information matrix is defined as

$$\mathbf{I}_e(\boldsymbol{\theta}|\mathbf{y}_{obs}) = \sum_{i=1}^{n} \mathbf{w}(\mathbf{y}_{obs_i}|\boldsymbol{\theta})\,\mathbf{w}^\top(\mathbf{y}_{obs_i}|\boldsymbol{\theta}) - \frac{1}{n}\mathbf{W}(\mathbf{y}_{obs}|\boldsymbol{\theta})\,\mathbf{W}^\top(\mathbf{y}_{obs}|\boldsymbol{\theta}),$$

where $\mathbf{W}^\top(\mathbf{y}_{obs}|\boldsymbol{\theta}) = \sum_{i=1}^{n} \mathbf{w}(\mathbf{y}_{obs_i}|\boldsymbol{\theta})$. It is noted from the result of Louis (1982) the individual score can be determined as

$$\mathbf{w}(\mathbf{y}_{obs_i}|\boldsymbol{\theta}) = \frac{\partial \ell(\boldsymbol{\theta}|\mathbf{y}_{\mathrm{obs}_i})}{\partial \boldsymbol{\theta}} = \mathrm{E}\left[\frac{\partial \ell_c(\boldsymbol{\theta}|\mathbf{z}_i)}{\partial \boldsymbol{\theta}}|\mathbf{y}_{obs_i}, \boldsymbol{\theta}\right]. \tag{20}$$

Substituting the ML estimates $\boldsymbol{\theta}$ in (20), $\mathbf{I}_e(\boldsymbol{\theta}|\mathbf{y}_{obs})$ is reduced to

$$\mathbf{I}_e\left(\widehat{\boldsymbol{\theta}}|\mathbf{y}_{obs}\right) = \sum_{i=1}^{n} \widehat{\mathbf{w}}_i \widehat{\mathbf{w}}_i^\top, \tag{21}$$

where $\widehat{\mathbf{w}}_i = (\widehat{w}_{\boldsymbol{\beta}i}, \widehat{w}_{\sigma^2 i}, \widehat{w}_{\boldsymbol{\nu}i})$ is an individual score vector and

$$\begin{aligned}
\widehat{w}_{\boldsymbol{\beta}i} &= \mathrm{E}\left[\frac{\partial \ell_c(\boldsymbol{\theta}|\mathbf{z}_i)}{\partial \boldsymbol{\beta}}|\mathbf{y}_{obs_i}, \widehat{\boldsymbol{\theta}}\right] \\
&= \frac{1}{\sigma^2}\mathbf{D}_{i\beta}\left(\mathcal{E}_{1i}(\widehat{\boldsymbol{\theta}}) - \mathcal{E}_{0i}(\widehat{\boldsymbol{\theta}})\eta(\mathbf{x}_i, \widehat{\boldsymbol{\beta}})\right), \\
\widehat{w}_{\sigma^2 i} &= \mathrm{E}\left[\frac{\partial \ell_c(\boldsymbol{\theta}|\mathbf{z}_i)}{\partial \sigma^2}|\mathbf{y}_{obs_i}, \widehat{\boldsymbol{\theta}}\right] \\
&= -\frac{1}{2\widehat{\sigma}^2} + \frac{1}{2\widehat{\sigma}^4}\Big(\mathcal{E}_{2i}(\widehat{\boldsymbol{\theta}}) - 2\mathcal{E}_{1i}(\widehat{\boldsymbol{\theta}})\eta(\mathbf{x}_i, \widehat{\boldsymbol{\beta}}) \\
&\quad + \mathcal{E}_{0i}(\widehat{\boldsymbol{\theta}})\eta(\mathbf{x}_i, \widehat{\boldsymbol{\beta}})^2\Big),
\end{aligned}$$

and

$$\begin{aligned}
\widehat{w}_{\boldsymbol{\nu}i} &= \mathrm{E}\left[\frac{\partial \ell_c(\boldsymbol{\theta}|\mathbf{z}_i)}{\partial \boldsymbol{\nu}}|\mathbf{y}_{obs_i}, \widehat{\boldsymbol{\theta}}\right] \\
&= \mathrm{E}\left[\frac{\partial \log\left(h(U_i|\boldsymbol{\nu})\right)}{\partial \boldsymbol{\nu}}|\mathbf{y}_{obs_i}, \widehat{\boldsymbol{\theta}}\right]
\end{aligned} \tag{22}$$

where $\ell_c(\boldsymbol{\theta}|\mathbf{z}_i)$ be the log-likelihood formed from the single complete observation $\mathbf{z}_i = (y_{\mathrm{obs}_i}, y_i, u_i)^\top$, $\mathbf{D}_{i\widehat{\beta}} = \partial\eta_i(\widehat{\boldsymbol{\beta}})/\partial\boldsymbol{\beta}$ and $\mathcal{E}_{si}(\boldsymbol{\theta}^{(k)}) = \mathrm{E}_{\boldsymbol{\theta}^{(k)}}[U_i Y_i^s | y_{\mathrm{obs}_i}]$. It is important to notice that the values of Equation (22) depend of the distribution of U. Thus for example:

- *For the Student-t distribution:* We consider $U \sim Gamma(\nu/2, \delta/2)$, with $\nu > 0$, then

$$\begin{aligned}
\widehat{w}_{\boldsymbol{\nu}i} &= -\psi\left(\frac{\widehat{\nu}}{2}\right) + \frac{1}{2}\left(\log\left(\frac{\widehat{\nu}}{2}\right) + 1\right) \\
&\quad + \frac{1}{2}\left(\mathrm{E}\left[\log\left(U_i\right)|\mathbf{y}_{obs_i}, \widehat{\boldsymbol{\theta}}\right] - \mathcal{E}_{0i}(\widehat{\boldsymbol{\theta}})\right)
\end{aligned}$$

  where $\psi(x)$ represents the digamma function of x.

- *For the Slash distribution:* We consider $U \sim Beta(\nu, 1)$ with positive shape parameter ν, then

$$\widehat{w}_{\boldsymbol{\nu}i} = \frac{1}{\widehat{\nu}} + \mathrm{E}\left[\log\left(U_i\right)|\mathbf{y}_{obs_i}, \widehat{\boldsymbol{\theta}}\right].$$

It is important to stress that the standard error of ν depends heavily on the calculation of $\mathrm{E}[\log(U_i)|\mathbf{y}_{obs_i}, \widehat{\boldsymbol{\theta}}]$, which relies on computationally intensive Monte Carlo integrations. In our analysis, we focus solely on comparing the standard errors of β and σ^2.

5. SIMULATION STUDIES

In order to study the performance of our proposed model and algorithm, we present three simulation studies. The first part of this simulation study shows that the parameter estimates based on the ECME algorithm from the SMN-NLCR models provides good asymptotic properties. The goal of the second part is to show the consistency of the standard errors for the fixed effects. The performance of the parameter estimates in the presence of outliers on the response variable is presented in the third simulation study. The computational procedures were implemented using the R software (R Core Team, 2013).

5.1 Asymptotic properties

The goal of this simulation study is to evaluate the finite-sample performance of the parameter estimates using the ECME algorithm developed in the Subsection 3.2. We performed a Monte Carlo simulation study with the nonlinear growth-curve model defined by,

$$Y_i = \frac{\beta_1}{1 + \exp(\beta_2 + \beta_3 x_i)} + \varepsilon_i, \quad i = 1, \ldots, n, \tag{23}$$

where $\varepsilon_i \sim SMN(0, \sigma^2, \nu)$. Following Labra et al. (2012), the variable x_i is a sequence of equally spaced values ranging from 0.1 to 20. These values were used throughout the simulations. The true values of the regression parameters were taken as $\beta_1 = 330$, $\beta_2 = 6.5$, $\beta_3 = -0.7$ and $\sigma^2 = 3$.

We generated 500 artificial samples from the SMN-NLCR model with censoring level $p = 10\%$ (i.e., 10% of the observations in each data set were censored). The sample sizes are $n = 30, 50, 100, 150, 200, 300, 400, 500, 700$ and 800. The main purpose here is the evaluation of bias (Bias) and mean square error (MSE). For β_i $(i = 1, 2, 3)$, these quantities are defined, respectively, by

$$\mathrm{Bias}\left(\beta_i\right) = \frac{1}{500}\sum_{j=1}^{500}\left(\widehat{\beta}_i^{(j)} - \beta_i\right) \quad \text{and}$$

$$\mathrm{MSE}\left(\beta_i\right) = \frac{1}{500}\sum_{j=1}^{500}\left(\widehat{\beta}_i^{(j)} - \beta_i\right)^2,$$

where $\widehat{\beta}_i^{(j)}$ is the estimate of β_i for the j-th sample. Observing Figure 1, we found that the bias and MSE tend to

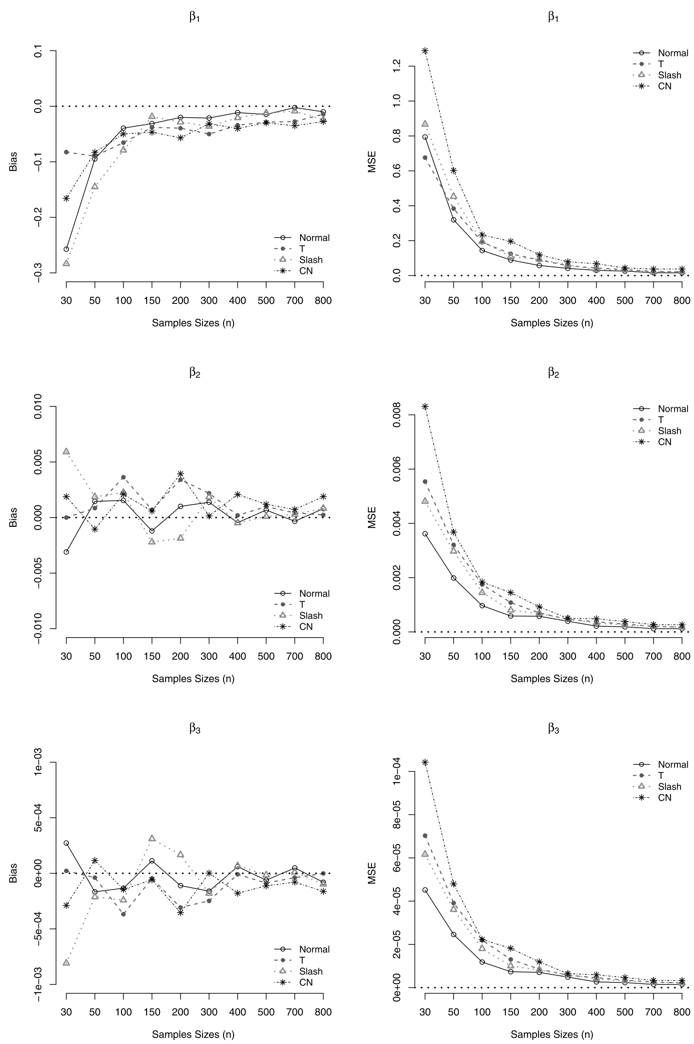

Figure 1. Simulation data. Bias and MSE from the parameter estimates considering $p = 10\%$.

approach to zero when the sample size n increases. As a general rule, the results indicate that the ML estimates of the model do provide good asymptotic properties. Simulations are also conducted under two higher censoring rates ($p = 30\%$ and 45%) and the patterns of convergence still behave well. See Figures 4 and 5 in Appendix B.

5.2 Consistency of the estimates of the standard errors for the fixed effects

The design considered in this simulation study is the same that we used in Subsection 5.1. In the study, we examine the consistency of the approximation method suggested in Section 4 for the standard errors (SE) of the MLE of the regression parameters $\boldsymbol{\theta}^* = (\boldsymbol{\beta}, \sigma^2)$. We generated 1,000 samples of size $n = 150$ from four different SMN-NLCR models, including N-NLCR, T-NLCR with $\nu = 4$, SL-NLCR with $\nu = 3$ and CN-NLCR with $\boldsymbol{\nu} = (0.1, 0.1)$. For each sample, eight different censoring levels (0%, 5%, 10%, 15%, 20%, 30%, 45% or 55%) were considered.

For each setup, we compute the MLE's of $\boldsymbol{\theta}^*$ along with the associated SE estimates and the 95% normal-approximation confidence intervals. Table 1 presents the sample standard errors of $\widehat{\theta^*}_i$

$$\text{MC SE} = \frac{1}{999}\sum_{j=1}^{1000}\left(\widehat{\theta^*}_i^{(j)} - \overline{\widehat{\theta^*}_i}\right)^2, \quad \text{where}$$

$$\overline{\widehat{\theta^*}_i} = \frac{1}{1000}\sum_{j=1}^{1000}\widehat{\theta^*}_i^{(j)},$$

which are calculated using the average values (across 1000 samples) of the standard errors computed using the information method (IM MC SE) and the percentage coverage of the resulting 95% confidence intervals (COV MC). It can be observed from Table 1 that the COV MC for $\boldsymbol{\beta}$ is quite stable, but the COV MC of σ^2 tend to be lower than the nominal level (95%).

5.3 Performance evaluation

The goal of this simulation study is to compare the performance of the parameter estimates for the SMN-NLCR regression models in the presence of outliers on the response variable. We considered a two-parameter Michaelis-Menten model, presented by Vanegas et al. (2012)

$$Y_i = \frac{\beta_1 x_i}{\beta_2 + x_i} + \varepsilon_i, \quad i = 1, \ldots, n, \tag{24}$$

where $\varepsilon_i \sim SMN(0, \sigma^2, \nu)$. True regression parameter values were taken as $\beta_1 = 3$, $\beta_2 = 0.5$ and $\sigma^2 = 1$. Sample size was $n = 300$ and the censoring level was fixed at $p = 8\%$. Covariates were generated from uniform distribution in (0,1) and those values were kept constant through the experiment. The number of Monte Carlo replications was 500.

To assess how much the EM estimates are influenced by the presence of outliers, we adopted six different levels ($\vartheta =$ 1%, 2%, 3%, 4%, 5% and 10%) of outliers present on the data sets. The outliers are created replacing the uncensored observations y_i chosen randomly by $y_i + 2Sd(y)$, where Sd represents the standard deviation of the data set.

Following Fagundes et al. (2013), the performance assessment of the parameter estimates is based on the mean magnitude of relative error (MMRE), defined as

$$\text{MMRE} = \frac{1}{3}\left\{\left|\frac{\widehat{\beta}_1(\vartheta) - \widehat{\beta}_1}{\widehat{\beta}_1}\right| + \left|\frac{\widehat{\beta}_2(\vartheta) - \widehat{\beta}_2}{\widehat{\beta}_2}\right| + \left|\frac{\widehat{\sigma}_2^2(\vartheta) - \widehat{\sigma}_2^2}{\widehat{\sigma}_2^2}\right|\right\},$$

where $\widehat{\theta}_i(\vartheta)$ is the MLE of θ_i after the contamination ϑ, with $\theta = \beta_1, \beta_2, \sigma^2$.

Table 2 shows the values (across 500 samples) of average and standard deviation of the MMRE obtained for the

Table 1. Simulation data. MC SE, IM MC SE and COV MC of $\widehat{\boldsymbol{\theta}^*}_i$

Cens level	Measure	N-NLCR				T-NLCR			
		$\widehat{\beta}_1$	$\widehat{\beta}_2$	$\widehat{\beta}_3$	$\widehat{\sigma}^2$	$\widehat{\beta}_1$	$\widehat{\beta}_2$	$\widehat{\beta}_3$	$\widehat{\sigma}^2$
0%	MC SE	0.260	0.024	0.002	0.349	0.305	0.028	0.003	0.430
	IM MC SE	0.265	0.024	0.002	0.357	0.315	0.029	0.003	0.453
	COV MC	96.0%	96.0%	96.0%	93.2%	94.6%	95.4%	95.8%	93.6%
5%	MC SE	0.280	0.025	0.002	0.369	0.336	0.030	0.003	0.477
	IM MC SE	0.276	0.025	0.002	0.363	0.321	0.030	0.003	0.470
	COV MC	94.2%	95.4%	94.8%	92.8%	93.2%	95.4%	96.0%	93.0%
10%	MC SE	0.308	0.026	0.002	0.471	0.347	0.031	0.003	0.508
	IM MC SE	0.285	0.026	0.002	0.381	0.332	0.030	0.003	0.481
	COV MC	93.4%	94.0%	94.0%	93.2%	93.0%	92.8%	94.6%	92.0%
15%	MC SE	0.299	0.027	0.003	0.392	0.368	0.032	0.003	0.507
	IM MC SE	0.293	0.027	0.003	0.388	0.341	0.031	0.003	0.496
	COV MC	95.0%	94.0%	94.8%	93.0%	93.2%	94.6%	94.6%	94.0%
20%	MC SE	0.300	0.027	0.003	0.391	0.391	0.033	0.003	0.573
	IM MC SE	0.300	0.028	0.003	0.396	0.349	0.032	0.003	0.506
	COV MC	94.2%	95.6%	96.0%	92.8%	92.2%	93.8%	94.2%	91.6%
30%	MC SE	0.328	0.029	0.003	0.405	0.379	0.035	0.003	0.536
	IM MC SE	0.322	0.030	0.003	0.429	0.373	0.035	0.003	0.537
	COV MC	94.8%	95.0%	94.4%	92.4%	94.6%	94.0%	95.4%	91.0%
45%	MC SE	0.435	0.034	0.003	0.590	0.512	0.042	0.004	0.621
	IM MC SE	0.370	0.034	0.003	0.496	0.422	0.039	0.004	0.600
	COV MC	94.2%	94.2%	94.0%	90.0%	90.8%	92.4%	92.6%	91.6%
55%	MC SE	0.415	0.036	0.004	0.562	0.478	0.043	0.004	0.685
	IM MC SE	0.409	0.039	0.004	0.550	0.473	0.044	0.004	0.681
	COV MC	95.4%	95.8%	96.6%	91.0%	94.0%	95.0%	95.6%	90.6%
Cens level	Measure	SL-NLCR				CN-NLCR			
		$\widehat{\beta}_1$	$\widehat{\beta}_2$	$\widehat{\beta}_3$	$\widehat{\sigma}^2$	$\widehat{\beta}_1$	$\widehat{\beta}_2$	$\widehat{\beta}_3$	$\widehat{\sigma}^2$
0%	MC SE	0.303	0.027	0.003	0.368	0.445	0.043	0.004	0.538
	IM MC SE	0.310	0.028	0.003	0.370	0.315	0.029	0.003	0.509
	COV MC	94.2%	96.6%	96.2%	92.0%	96.6%	94.4%	94.8%	92.8%
5%	MC SE	0.329	0.030	0.003	0.381	0.518	0.049	0.003	0.427
	IM MC SE	0.316	0.029	0.003	0.382	0.420	0.039	0.003	0.417
	COV MC	93.8%	94.2%	94.4%	92.6%	93.4%	93.2%	93.2%	91.4%
10%	MC SE	0.360	0.029	0.003	0.436	0.594	0.043	0.004	0.537
	IM MC SE	0.324	0.030	0.003	0.398	0.434	0.040	0.004	0.522
	COV MC	92.4%	95.4%	95.8%	92.4%	91.6%	94.2%	94.6%	90.8%
15%	MC SE	0.348	0.029	0.003	0.394	0.533	0.053	0.005	0.489
	IM MC SE	0.333	0.031	0.003	0.402	0.437	0.041	0.004	0.419
	COV MC	94.2%	95.8%	95.4%	93.8%	92.2%	90.4%	91.0%	91.8%
20%	MC SE	0.358	0.031	0.003	0.383	0.578	0.054	0.005	0.472
	IM MC SE	0.343	0.032	0.003	0.414	0.469	0.043	0.004	0.434
	COV MC	94.0%	95.6%	95.8%	93.2%	92.0%	91.6%	91.4%	91.0%
30%	MC SE	0.386	0.034	0.003	0.440	0.579	0.049	0.005	0.543
	IM MC SE	0.373	0.034	0.003	0.445	0.466	0.043	0.004	0.517
	COV MC	95.2%	94.6%	95.4%	91.6%	92.6%	92.8%	92.4%	90.8%
45%	MC SE	0.493	0.038	0.004	0.642	0.747	0.060	0.006	0.615
	IM MC SE	0.422	0.040	0.004	0.512	0.625	0.049	0.005	0.632
	COV MC	93.6%	96.2%	96.2%	91.6%	90.8%	92.4%	92.8%	90.0%
55%	MC SE	0.529	0.042	0.004	0.639	0.919	0.070	0.007	0.692
	IM MC SE	0.470	0.044	0.004	0.577	0.597	0.057	0.006	0.658
	COV MC	94.4%	95.0%	94.8%	93.6%	93.0%	92.8%	94.0%	90.2%

Table 2. Simulation data. Average and standard deviation (in parenthesis) of the MMRE

Outlier quantity (%)	Models			
	N-NLCR		T-NLCR	
1	0.06052	(0.0296)	0.04578	(0.0314)
2	0.09498	(0.0406)	0.06933	(0.0413)
3	0.13238	(0.0495)	0.09428	(0.0464)
4	0.16531	(0.0584)	0.11085	(0.0580)
5	0.19635	(0.0718)	0.13419	(0.0690)
10	0.32163	(0.0763)	0.26046	(0.0965)
(%)	SL-NLCR		CN-NLCR	
1	0.02517	(0.0336)	0.04797	(0.0481)
2	0.06085	(0.0497)	0.07034	(0.0471)
3	0.09012	(0.0521)	0.09613	(0.0648)
4	0.12058	(0.0716)	0.11725	(0.0723)
5	0.14844	(0.0724)	0.13914	(0.0814)
10	0.27139	(0.0764)	0.25357	(0.0814)

different SMN-NLCR models. In the the N-NLCR case, we observe that influence increases when the different quantities of outliers increases. In contrast, for the SMN-NLCR models with heavy tails, namely the T-NLCR, SL-NLCR and CN-NLCR, the measures vary little, indicating that they are more robust than the N-NLCR model in the ability to accommodate discrepant observations.

6. APPLICATION

In this section, we apply the proposed techniques to ultrasonic calibration data previously analyzed by Lin et al. (2009). These data are the result of the NIST study related to ultrasonic calibration on 214 samples. The response variable is the ultrasonic response (Y) and the predictor variable is metal distance (X). Following Lin et al. (2009), we consider the following non-linear model:

$$Y_i = \frac{\exp(-\beta_1 x_i)}{\beta_2 + \beta_3 x_i} + \varepsilon_i, \qquad i = 1, 2, \ldots, 214.$$

We utilize the same non-linear function to evaluate the performance of SMN-NLCR models. To conduct experimental studies, we choose randomly $p = 8\%$ (18 observations) as censoring interval level and replaced each observation chosen randomly, Y_j, by the interval (Y_{L_j}, Y_{U_j}) where $Y_{L_j} = max(0, Y_j - \frac{1}{4}Sd(Y))$, $Y_{U_j} = Y_j + \frac{1}{4}Sd(Y)$ and Sd is the Standard deviation of Y. Thus, the cases ♯13, ♯30, ♯49, ♯50, ♯61, ♯74, ♯79, ♯106, ♯110, ♯118, ♯130, ♯137, ♯147, ♯166, ♯185, ♯195, ♯204 and ♯213 were selected as censored, see Figure 2 (panel a).

To identify atypical observations and/or model misspecification, we follow Barros et al. (2010) approach to examine the transformed martingale residual, r_{MT_i}, defined as

$$r_{MT_i} = \text{sign}(r_{M_i})\sqrt{-2\left[r_{M_i} + \delta_i \log\left(\delta_i - r_{M_i}\right)\right]}, \quad i = 1, \ldots, n$$

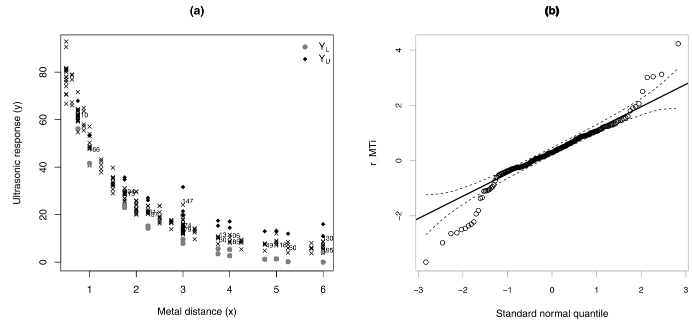

Figure 2. (a) Scatter-plot of the ultrasonic calibration data with censoring. (b) Envelope of the martingale-type residuals, r_{MT_i}, for the N-NLCR model.

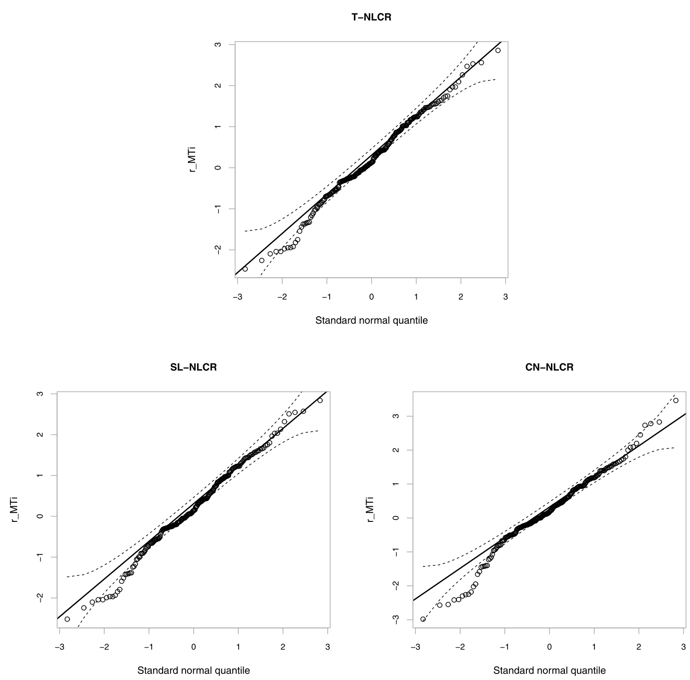

Figure 3. Envelopes of the martingale-type residuals for the SMN-NLCR models.

where $r_{M_i} = \delta_i + \log(S(y_i, \widehat{\boldsymbol{\theta}}))$ is the martingale residual (Ortega et al., 2003) with $\delta_i = 0, 1$ indicating whether the i-th observation is censored or not, respectively, sign(r_{M_i}) denoting the sign of r_{M_i} and $S(y_i, \widehat{\boldsymbol{\theta}}) = P_{\widehat{\boldsymbol{\theta}}}(Y_i > y_i)$ representing the survival function evaluated at y_i. A more detailed account of the martingale residual can be found in Therneau et al. (1990). The plot of r_{MT_i} for the N-NLCR model together with the 95% confidence envelopes are presented in Figure 3 (panel b). This Figure exhibits a heavy-tailed behavior, suggesting that the normality assumption might be inappropriate. In order to study the validity of the hypothesized N-NLCR model, we perform the Shapiro-Wilk normality test and obtained a *p-value* ≈ 0, which rejects the hypothesis of normality. The non-normality of the

*Table 3. Real data. EM estimates, estimated standard errors (SE) and a confidence interval asymptotic (IC) for the SMN-NLCR models. * indicates parameter significance*

	Estimate	SE	IC (95%)	
	N-NLCR			
	Estimate	SE	IC (95%)	
β_1	0.1953	0.0274	(0.1415; 0.2492)	*
β_2	0.0061	0.0003	(0.0055; 0.0067)	*
β_3	0.0103	0.0008	(0.0087; 0.0120)	*
σ^2	11.1801	0.7151	(9.7785;12.5817)	*
	T-NLCR			
	Estimate	SE	IC (95%)	
β_1	0.1803	0.0165	(0.1478; 0.2127)	*
β_2	0.0059	0.0002	(0.0054; 0.0064)	*
β_3	0.0111	0.0006	(0.0099; 0.0123)	*
σ^2	3.6470	0.5448	(2.5792; 4.7149)	*
ν	2.4562	——	——	
	SL-NLCR			
	Estimate	SE	IC (95%)	
β_1	0.1846	0.0173	(0.1507; 0.2186)	*
β_2	0.0060	0.0002	(0.0054; 0.0065)	*
β_3	0.0109	0.0006	(0.0097; 0.0121)	*
σ^2	2.1936	0.3080	(1.5897; 2.7974)	
ν	1.0100	——	——	
	CN-NLCR			
	Estimate	SE	IC (95%)	
β_1	0.1868	0.0199	(0.1477; 0.2258)	*
β_2	0.0060	0.0002	(0.0054; 0.0066)	*
β_3	0.0108	0.0006	(0.0095; 0.0122)	*
σ^2	4.7709	0.52462	(3.7426; 5.7991)	*
ν	0.2	——	——	
γ	0.2	——	——	

distribution gives an indication that some atypical observations or outliers might exist in the data. Consequently, we revisited the censored ultrasonic calibration dataset with the implementation of T-NLCR, SL-NLCR and CN-NLCR models using the ECME algorithm described in Section 3.2. Table 3 shows the parameter estimates together with the corresponding SE and the 95% normal-approximation confidence intervals. The SE obtained by T-NLCR, SL-NLCR and CN-NLCR models are smaller than that of the N-NLCR model. Note that the estimates of all the coefficients $\boldsymbol{\beta}$ are significant for all the SMN-NLCR models since all 95% confidence intervals of $\boldsymbol{\beta}$ do not include zero. Table 4 presents some model selection criteria, together with the values of the log-likelihood. The AIC (Akaike, 1974), BIC (Schwarz, 1978), EDC (Bai et al., 1989) and AIC_{SUR} (Liang and Zou, 2008) values indicate that the three models with longer than normal tails do likely produce more accurate estimates. Figure 3 show the plots of r_{MT_i} for the T-NLCR, SL-NLCR and CN-NLCR models along with their 95% confidence envelopes. Clearly, the SMN-NLCR models with heavy tails contain fewer observations outside the envelopes, signifying a better fit than the N-NLCR model. Moreover, the model

Table 4. Real data. Comparison between the SMN-NLCR models

	Models			
Criteria	N-NLCR	**T-NLCR**	SL-NLCR	CN-NLCR
log-likelihood	-520.783	**-497.106**	-497.683	-498.743
AIC	1049.566	**1004.210**	1005.367	1009.488
BIC	1063.030	**1021.042**	1022.197	1029.684
EDC	1053.269	**1008.841**	1009.996	1005.911
AIC_{SUR}	1049.972	**1004.756**	1005.911	1010.190

selection criteria shown in Table 4 indicate that the T-NLCR presents the best fit, followed closely by SL-NLCR and CN-NLCR models. The fit of N-NLCR is the worst, indicating a lack of adequacy of normality assumptions for this dataset. As suggested by a referee, the comparison process is conducted for the ultrasonic calibration data without censored observations. The T-NLCR still presents a better overall fit than the other three models (see Table 5 in Appendix C).

7. CONCLUSIONS

We studied the nonlinear regression models with censored responses based on scale mixtures of normal distributions, called SMN-NLCR models. This class of distributions offers a high degree of flexibility, allowing us to deal properly with censored data in the presence of outliers. For parameter estimation, a computational efficient ECME algorithm was developed using formulas for the moments of the truncated SMN distribution. Simulation studies revealed that our proposed method is quite robust against outlying and influential observations. The consistency of the EM-based estimator is also demonstrated. Experimental results show that the use of SMN-NLCR models with heavy tails offer a better fitting as well as a better protection against outliers than the N-NLCR model.

There are a number of possible extensions of the current work. It is of interest to generalize the SMN-NLCR model for a single response to the case of multiple responses and make allowance of missing data. The non-identifiability problem due to missing and censored data will be studied in depth in our future work. Due to the popularity of Markov chain Monte Carlo techniques, another potential work is to pursue a fully Bayesian treatment in this context for producing posterior inference. The methodology can also be extended to mixtures of nonlinear regressions with skewed and heavy-tailed censored responses based on recent approaches by Rossin et al. (2011), Da Silva Ferreira et al. (2011) and Liu and Lin (2014).

ACKNOWLEDGMENTS

We are grateful to the editor, the associate editor and three anonymous referees for providing insightful comments that led to an improved presentation of this paper. Aldo M. Garay and Víctor H. Lachos acknowledges the partial financial support from Fundação de Amparo à Pesquisa do

Estado de São Paulo (by grant 2013/21468-0 and 2014/02938-9) and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq-Brazil), and Tsung-I Lin acknowledges the partial financial support from the Ministry of Science and Technology of Taiwan (Grant no. MOST 103-2118-M-005-001-MY2).

APPENDIX A. PROOF OF PROPOSITION 1

In what follows E_X denotes an expectation relative to the distribution of X and, for the sake of notation simplicity, we denote all pdf's by f. Thus, for example, $f(u)$ denotes the pdf of U, $f(u,x)$ denotes the joint pdf of U and X and $f(u|X \in \mathcal{A})$ denotes the pdf of U given the event $\{X \in \mathcal{A}\}$. Let $\mathcal{A} = (a,b)$. From (1), we have

$$X|U = u \sim \mathrm{N}(0, u^{-1}), \quad X|X \in \mathcal{A} \sim \mathrm{TSMN}_{\mathcal{A}}(0,1) \quad \text{and}$$
$$X|U = u, X \in \mathcal{A} \sim \mathrm{TN}_{\mathcal{A}}(0, u^{-1}),$$

where $\mathrm{TN}_{\mathcal{A}}(0,u^{-1})$ denotes the truncated normal distribution in $\mathcal{A}$, being 0 and u^{-1} the mean and variance, respectively, before truncation. We have that

$$\begin{aligned} (25)\qquad & \mathrm{E}\left[U^r X^s | X \in \mathcal{A}\right] \\ &= \mathrm{E}_U\left[U^r \mathrm{E}_X\left[X^s | U = u, X \in \mathcal{A}\right] | X \in \mathcal{A}\right] \\ &= \int_0^\infty U^r \mathrm{E}_X\left[X^s | U = u, X \in \mathcal{A}\right] f(u|X \in \mathcal{A}) du. \end{aligned}$$

We have that

$$\begin{aligned} f(u|X \in \mathcal{A}) &= \int f(u,x|X \in \mathcal{A}) dx \\ &= \int f(u|X = x, X \in \mathcal{A}) f(x|X \in \mathcal{A}) \\ (26)\qquad &= \frac{1}{P(X \in \mathcal{A})} \int f(u|X = x, X \in \mathcal{A}) f(x) \mathbb{I}_{\mathcal{A}}(x) \\ (27)\qquad &= \frac{1}{F_{SMN}(b) - F_{SMN}(a)} \int f(u,x) \mathbb{I}_{\mathcal{A}}(x) dx \\ &= \frac{1}{F_{SMN}(b) - F_{SMN}(a)} \int_{\mathcal{A}} f(u) \phi\left(x|0, u^{-1}\right) dx \\ &= \frac{f(u)}{F_{SMN}(b) - F_{SMN}(a)} \int_{\mathcal{A}^*} \phi(z) dz \\ &= \frac{f(u)\left[\Phi(b\sqrt{u}) - \Phi(a\sqrt{u})\right]}{F_{SMN}(b) - F_{SMN}(a)}, \end{aligned}$$

where $\mathcal{A}^* = (a\sqrt{u}, b\sqrt{u})$. Equation (26) is obtained using the pdf's expression of $X|X \in \mathcal{A}$. Equation (27) is consequence of the fact that, if $x \in \mathcal{A}$, then $\{X \in \mathcal{A}, X = x\} = \{X = x\}$, implying that $f(u,x) = f(u|X = x) f(x) = f(u|X = x, X \in \mathcal{A}) f(x)$. If $x \notin \mathcal{A}$ then $\mathbb{I}_{\mathcal{A}}(x) = 0$ and the integrands in (26) and (27) are equal to zero. By (25) and Lemma 1 given in this Appendix, it follows that

- for $s = 0$,

$$\begin{aligned} \mathrm{E}\left[U^r | X \in \mathcal{A}\right] &= \int_0^\infty U^r f(u|X \in \mathcal{A}) du \\ &= \frac{\mathrm{E}_U\left[U^r \Phi\left(b\sqrt{U}\right) - U^r \Phi\left(a\sqrt{U}\right)\right]}{F_{SMN}(b) - F_{SMN}(a)}. \end{aligned}$$

- for $s = 1$,

$$\begin{aligned} & \mathrm{E}\left[U^r X | X \in \mathcal{A}\right] \\ &= \int_0^\infty U^r \frac{1}{U^{1/2}} \frac{\left(\phi\left(a\sqrt{U}\right) - \phi\left(b\sqrt{U}\right)\right)}{\Phi\left(b\sqrt{U}\right) - \Phi\left(a\sqrt{U}\right)} \\ &\quad \times f(u|X \in \mathcal{A})(u) du \\ &= \frac{\mathrm{E}_U\left[U^{r-1/2}\phi\left(a\sqrt{U}\right) - U^{r-1/2}\phi\left(b\sqrt{U}\right)\right]}{F_{SMN}(b) - F_{SMN}(a)}. \end{aligned}$$

- for $s = 2$,

$$\begin{aligned} & \mathrm{E}\left[U^r X^2 | X \in \mathcal{A}\right] \\ &= \int_0^\infty U^r f(u|X \in \mathcal{A}) \\ &\quad \times \left[\frac{1}{U} + \frac{aU^{-1/2}\phi\left(a\sqrt{U}\right) - bU^{-1/2}\phi\left(b\sqrt{U}\right)}{\Phi\left(b\sqrt{U}\right) - \Phi\left(a\sqrt{U}\right)}\right] du \\ &= \frac{1}{F_{SMN}(b) - F_{SMN}(a)} \mathrm{E}_U\left[U^{r-1}\Phi\left(b\sqrt{U}\right)\right. \\ &\quad - U^{r-1}\Phi\left(a\sqrt{U}\right) + aU^{r-1/2}\phi\left(a\sqrt{U}\right) \\ &\quad \left. - bU^{r-1/2}\phi\left(b\sqrt{U}\right)\right]. \end{aligned}$$

The following Lemmas, provided by Kim (2008) and Genç (2013), are useful for evaluating some integrals used in this paper as well as for the implementation of the proposed EM-type algorithm.

Lemma 1: If $Z \sim \mathrm{TN}_{(a,b)}(0,1)$, then

$$(k+1) E\left[Z^k\right] - E\left[Z^{k+2}\right] = \frac{b^{k+1}\phi(b) - a^{k+1}\phi(a)}{\Phi(b) - \Phi(a)},$$

for $k = -1, 0, 1, 2, \ldots$.

Proof. See Lemma 2.3 in Kim (2008).

Lemma 2: Let U be a positive random variable. Then $F_{SMN}(a) = E_U[\Phi(a\sqrt{U})]$, where $F_{SMN}(\cdot)$ denotes the cdf of a standard SMN random variable, that is, when $\mu = 0$ and $\sigma^2 = 1$.

Proof. See Lemma 3 in Genç (2013).

Lemma 3: For $\nu > 0$, $\int_0^u x^{v-1} e^{-\mu x} dx = \mu^{-v} \gamma(v, \mu u)$, where $\gamma(a,x) = \int_0^x e^{-t} t^{a-1} dt$ is the incomplete gamma function.

Proof. See Lemma 6 in Genç (2013).

The following Corollary is a direct consequence of Proposition 1 given in Section 2.

Corollary: Let $Y \sim SMN(\mu, \sigma^2, \boldsymbol{\nu})$ with scale factor U and $\mathcal{A} = (\mathrm{a}, \mathrm{b})$. Then, for $r \geq 1$,

$$\begin{aligned} E[U^r|Y \in \mathcal{A}] &= E[U^r|X \in \mathcal{A}^*] \\ E[U^rY|Y \in \mathcal{A}] &= \mu E[U^r|X \in \mathcal{A}^*] + \sigma E[U^rX|X \in \mathcal{A}^*] \\ E\left[U^rY^2|Y \in \mathcal{A}\right] &= \mu^2 E[U^r|X \in \mathcal{A}^*] \\ &\quad + 2\mu\sigma E[U^rX|X \in \mathcal{A}^*] \\ &\quad + \sigma^2 E\left[U^rX^2|X \in \mathcal{A}^*\right], \end{aligned}$$

where $X \sim SMN(0, 1, \boldsymbol{\nu})$ and $\mathcal{A}^* = (a^*, b^*)$, with $a^* = (a - \mu)/\sigma$ and $b^* = (b - \mu)/\sigma$.

APPENDIX B. SIMULATION 1

In this Appendix, we present the results of the simulation study 1 for the levels censoring: $p = 30\%$ and $p = 45\%$.

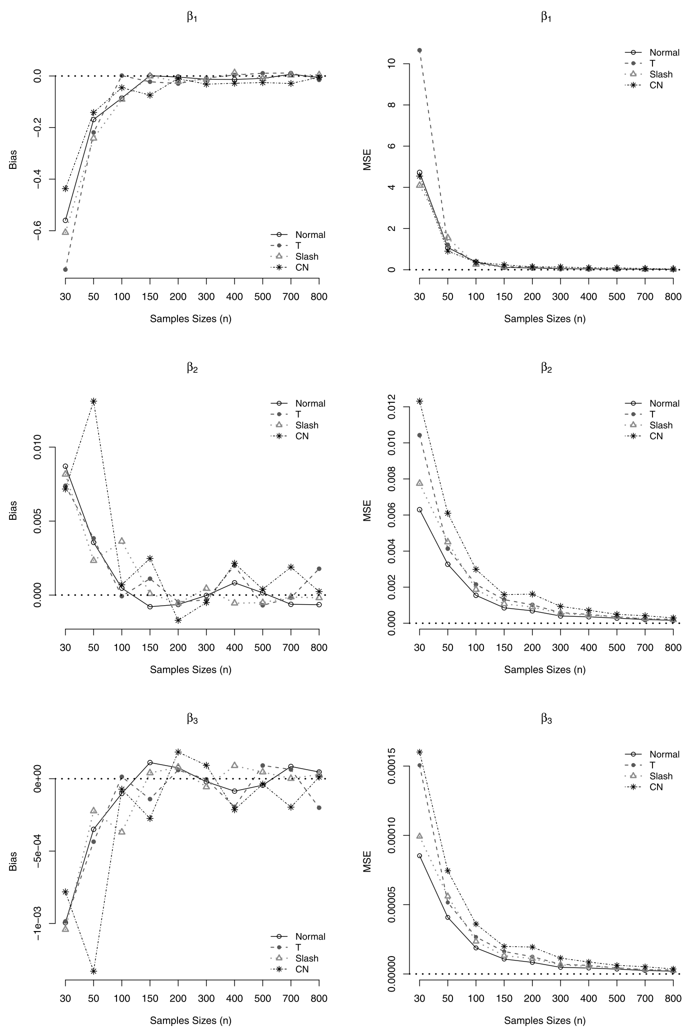

Figure 4. Simulation data. Bias and MSE from the parameter estimates considering $p = 30\%$.

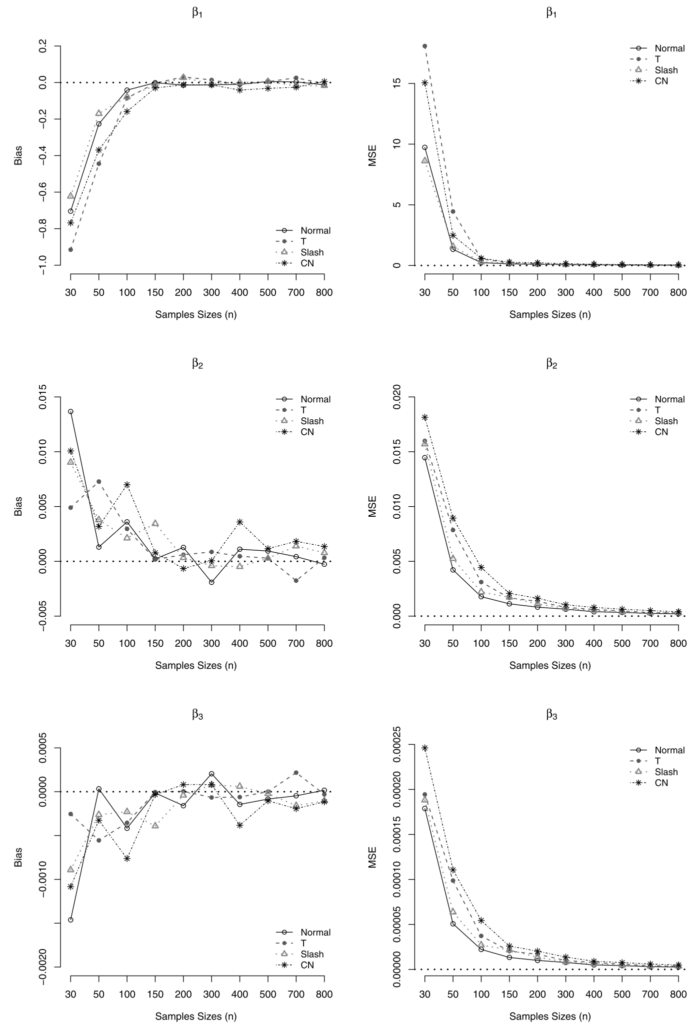

Figure 5. Simulation data. Bias and MSE from the parameter estimates considering $p = 45\%$.

APPENDIX C. ULTRASONIC CALIBRATION DATA WITHOUT CENSORED DATA

In this Appendix, we present the comparison between the SMN-NLCR models, considering the ultrasonic calibration data set, without censored data.

Table 5. Ultrasonic calibration data without censored data. Comparison between the SMN-NLCR models

	Models			
Criteria	N-NLCR	**T-NLCR**	SL-NLCR	CN-NLCR
log-likelihood	-561.604	**-531.526**	-532.679	-561.505
AIC	1131.208	**1073.053**	1075.359	1135.011
BIC	1144.672	**1089.883**	1092.189	1155.207
EDC	1134.911	**1077.682**	1079.987	1140.565
AIC_{SUR}	1131.614	**1073.597**	1075.902	1135.713

Received 2 February 2014

REFERENCES

Akaike, H., 1974. A new look at the statistical model identification. IEEE Transactions on Automatic Control 19, 716–723. MR0423716

Azzalini, A., 1985. A class of distributions which includes the normal ones. Scandinavian Journal of Statistics 12, 171–178. MR0808153

Bai, Z. D., Krishnaiah, P. R., Zhao, L. C., 1989. On rates of convergence of efficient detection criteria in signal processing with white noise. IEEE Transactions on Information Theory 35, 380–388. MR0999652

Barros, M., Galea, M., González, M., Leiva, V., 2010. Influence diagnostics in the Tobit censored response model. Statistical Methods & Applications 19, 716–723. MR2673350

Cancho, V. G., Lachos, V. H., Ortega, E. M., 2010. A nonlinear regression model with skew-normal errors. Statistical Papers 51 (3), 547–558. MR2679333

Cysneiros, F. J. A., Vanegas, L. H., 2008. Residuals and their statistical properties in symmetrical nonlinear models. Statistics & Probability Letters 78, 3269–3273. MR2479487

Da Silva Ferreira, C., Bolfarine, H., Lachos, V. H., 2011. Skew scale mixtures of normal distributions: properties and estimation. Statistical Methodology 8, 154–171. MR2769277

Dempster, A., Laird, N., Rubin, D., 1977. Maximum likelihood from incomplete data via the EM algorithm. Journal of the Royal Statistical Society, Series B 39, 1–38. MR0501537

Fagundes, R. A., de Souza, R. M., Cysneiros, F. J. A., 2013. Robust regression with application to symbolic interval data. Engineering Applications of Artificial Intelligence 26 (1), 564–573.

Fang, K. T., Zhang, Y. T., 1990. Generalized Multivariate Analysis. Springer. MR1079542

Genç, A. İ., 2013. Moments of truncated normal/independent distributions. Statistical Papers 54, 741–764. MR3072898

Heuchenne, C., Van Keilegom, I., 2007. Nonlinear regression with censored data. Technometrics 49, 34–44. MR2345450

Kim, H. J., 2008. Moments of truncated Student-t distribution. Journal of the Korean Statistical Society 37, 81–87. MR2409373

Labra, F. V., Garay, A. M., Lachos, V. H., Ortega, E. M. M., 2012. Estimation and diagnostics for heteroscedastic nonlinear regression models based on scale mixtures of skew-normal distributions. Journal of Statistical Planning and Inference 142, 2149–2165. MR2903419

Liang, H., Zou, G., 2008. Improved AIC selection strategy for survival analysis. Computational Statistics & Data Analysis 52 (5), 2538–2548. MR2411957

Lin, J., Xie, F., Wei, B., 2009. Statistical diagnostics for skew-t-normal nonlinear models. Communications in Statistics-Simulation and Computation 38 (10), 2096–2110. MR2751190

Lin, T. I., 2010. Robust mixture modeling using multivariate skew t distributions. Statistics and Computing 20 (3), 343–356. MR2725392

Liu, C., 1996. Bayesian robust multivariate linear regression with incomplete data. Journal of the American Statistical Association 91, 1219–1227. MR1424619

Liu, C., Rubin, D. B., 1994. The ECME algorithm: a simple extension of EM and ECM with faster monotone convergence. Biometrika 80, 267–278. MR1326414

Liu, M., Lin, T. I., 2014. A skew-normal mixture regression model. Educational and Psychological Measurement 74 (1), 139–162.

Louis, T., 1982. Finding the observed information matrix when using the em algorithm. Journal of the Royal Statistical Society, Series B, 226–233. MR0676213

Meilijson, I., 1989. A fast improvement to the em algorithm to its own terms. J. R. Stat. Soc. Ser. B 51, 127–138. MR0984999

Meng, X. L., Rubin, B. D., 1993. Maximum likelihood estimation via the ECM algorithm: a general framework. Biometrika 80, 267–278. MR1243503

Meza, C., Osorio, F., la Cruz, R. D., 2012. Estimation in nonlinear mixed-effects models using heavy-tailed distributions. Statistics and Computing 22, 121–139. MR2865060

Ortega, E. M. M., Bolfarine, H., Paula, G. A., 2003. Influence diagnostics in generalized log-gamma regression models. Computational Statistics & Data Analysis 42, 165–186. MR1963013

Osorio, F., Paula, G. A., Galea, M., 2007. Assessment of local influence in elliptical linear models with longitudinal structure. Computational Statistics & Data Analysis 51, 4354–4368. MR2364450

R Core Team, 2013. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org.

Rossin, E., Lin, T. I., Ho, H. J., Mentzer, S. J., Pyne, S., 2011. A framework for analytical characterization of monoclonal antibodies based on reactivity profiles in different tissues. Bioinformatics 27, 2746–2753.

Schwarz, G., 1978. Estimating the dimension of a model. Annals of Statistics 6, 461–464. MR0468014

Therneau, T. M., Grambsch, P. M., Fleming, R. T., 1990. Martingale-based residuals for survival models. Biometrika, 147–160. MR1049416

Vaida, F., Liu, L., 2009. Fast implementation for normal mixed effects models with censored response. Journal of Computational and Graphical Statistics 18, 797–817. MR2750442

Vanegas, L. H., Cysneiros, F. J. A., 2010. Assessment of diagnostic procedures in symmetrical nonlinear regression models. Computational Statistics & Data Analysis 54, 1002–1016. MR2580934

Vanegas, L. H., Rondón, L. M., Cysneiros, F. J. A., 2012. Diagnostic procedures in Birnbaum–Saunders nonlinear regression models. Computational Statistics & Data Analysis 56 (6), 1662–1680. MR2892367

Wei, C. G., Tanner, M. A., 1990. Posterior computations for censored regression data. Journal of the American Statistical Association 85, 829–839.

Aldo M. Garay
Department of Statistics
Campinas State University
Brazil

Rua Sérgio Buarque de Holanda, 651
Cidade Universitária "Zeferino Vaz"
Campinas, São Paulo
Brazil
E-mail address: medina_garay@yahoo.com

Victor H. Lachos
Department of Statistics
Campinas State University
Brazil
E-mail address: hlachos@gmail.com

Tsung-I Lin
National Chung Hsing University
Taichung
Taiwan

China Medical University
Taichung
Taiwan
E-mail address: tilin@nchu.edu.tw