Statistical Papers (2025) 66:50
https://doi.org/10.1007/s00362-024-01634-1

REGULAR ARTICLE

Modified maximum likelihood estimator for censored linear regression model with two-piece generalized *t* distribution

Chengdi Lian[1] · Camila Borelli Zeller[2] · Ke Yang[1] · Weihu Cheng[1]

Received: 22 August 2023 / Revised: 22 April 2024
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2025

Abstract

In many fields, limited or censored data are often collected due to limitations of measurement equipment or experimental design. Commonly used censored linear regression models rely on the assumption of normality for the error terms. However, this approach has faced criticism in literature due to its sensitivity to deviations from the normality assumption. In this paper, we propose an extension of the CR model under the two-piece generalized *t* (TPGT)-error distribution, called TPGT-CR model. The TPGT-CR model offers greater flexibility in modeling data by accommodating skewness and heavy tails. We developed a modified maximum likelihood (MML) estimator for the proposed model and introduced the modified deviance residual to detect outliers. The developed MML estimator under the TPGT assumption possesses several appealing merits, including robustness against outliers, asymptotic equivalence to the maximum likelihood estimator, and explicit functions of sample observations. Simulation studies are conducted to examine the finite sample performance, robustness, and effectiveness of both the classical and proposed estimators. The results from both the simulated and real data illustrate the usefulness of the proposed method.

Keywords Censored regression model · Two-piece generalized *t* distribution · Modified maximum likelihood · Residual analysis · Tobit model

Chengdi Lian, Camila Borelli Zeller, Ke Yang have contributed equally to this work.

✉ Weihu Cheng
chengweihu@bjut.edu.cn

Chengdi Lian
lianchengdi@emails.bjut.edu.cn

Camila Borelli Zeller
camila.zeller@ufjf.edu.br

Ke Yang
yangke@bjut.edu.cn

[1] School of Mathematics, Statistics and Mechanics, Beijing University of Technology, Pingleyuan NO.100, Beijing 100124, China

[2] Rua José Lourenço Kelmer, s/n - São Pedro - Juiz de Fora, MG 36036-900, Brazil

Published online: 10 February 2025

1 Introduction

In many practical situations, some subjects may be recorded only if the values fall within an interval range, so the responses are often subject to censoring. For example, in astronomical data, censoring can occur due to nondetections. Censored regression model has been widely used for analyzing censored data in the theoretical and applied econometric field [see Khan and Tamer (2009), Chen and Khan (2000), Khan and Powell (2001)].

The classical parametric estimators of this model rely on the assumption that the error terms follow a known parametric distribution. It is well-known that the Tobit estimator based on normally distributed errors is inefficient for non-normal error terms. In addition, the Tobit estimates has certain limitations when there are a few extreme observations in the data and look way off in some cases. To address these limitations and allow for valid analysis of censored data, various alternative estimation procedures have been developed for the CR model, which can be categorized into two types.

The first type is distribution-free method, which does not assume any specific error distribution. Examples of this type include censored least absolute deviations (CLAD) estimator proposed by Powell (1984), symmetrically censored least squares (SCLS) proposed by Powell (1986), and two-step estimators proposed by Khan and Powell (2001). The second type aims to identify a specific distribution for the censored data and then estimate the model parameters. In an earlier study, Amemiya (1985) introduced a bivariate sample selection model, extending the Tobit model with a censoring latent variable distinct from the outcome-generating latent variable (see also chapter 16 in Cameron and Trivedi (2005)). Arellano-Valle et al. (2012) introduced the T-CR model, where the error terms are independently distributed with a t distribution. Lewis and McDonald (2014) proposed partially adaptive estimators (PAE) for the CR model by considering two families of distributions, i.e., the exponential generalized beta of the second kind and the generalized t (GT) distributions. Garay et al. (2017) proposed a robust CR model where the errors are assumed to follow a scale mixture of normal (SMN) distribution. In the context of SMN censored regression model, Garay et al. (2015) introduced another robust parametric approach from a Bayesian perspective. More recently, Zeller et al. (2019) developed an efficient EM-type algorithm for the extension of the finite mixture Tobit model by assuming that the error terms follow a distribution from the SMN class independently. Massuia et al. (2017) provided a Bayesian framework for CR models by considering the scale mixtures of skew-normal (SMSN) class of distributions.

The vast majority of the mentioned works either requires computationally intensive numerical techniques, or is too complex for practitioners to implement. Additionally, some proposals are not suitable for modeling highly skewed data with censoring and heavy tails, which is common among actuarial, financial, social, epidemiological, and medical studies (Carrasco et al. 2008; Wang et al. 2016; Guzmán et al. 2021). For such complex censored data structure, traditional Tobit model may lead to biased and inconsistent estimation results. Thus, there is a need to seek an appropriate theoretical model and develop effective inference techniques that can provide valid analysis. Lian et al. (2024) proposed the TPGT distribution, which is inspired by the two-piece distribution family. Two-piece distributions are a two-component mixtures of truncated

densities, allowing for flexible fitting to various types of data. As Arellano-Valle et al. (2005), Mudholkar and Hutson (2000), Arellano-Valle et al. (2020), all of these studies are based on this distribution family.

In this context, we propose a CR model for censored data structure where the observational errors have a two-piece generalized t distribution, called TPGT-CR model. First, our TPGT-CR model assumes a data distribution called TPGT, while the censoring thresholds and mechanisms of the model are determined by the data. This distributional assumption enhances flexibility in modeling data by accommodating skewness and heavy tails. The TPGT-CR model is capable of capturing more features of the data, offering higher interpretability, and modeling capabilities for left/right-censored data, thus making it applicable to a wider range of domains. Based on the proposed TPGT-CR model, we use the modified maximum likelihood method to estimate the parameters [see Balci et al. (2013), Yalçınkaya et al. (2018)]. The MML estimators have several appealing merits: (i) They have explicit forms to facilitate calculation; (ii) The MML estimators are asymptotic equivalence to ML estimators [see also Bhattacharyya (1985), Vaughan and Tiku (2000), Tiku and Suresh (1992) and Sect. 4 of Yalçınkaya et al. (2018)]. This ensures that both MML estimators and ML estimators have the same asymptotic properties necessary for hypothesis testing, such as the consistency and asymptotic normality. Besides, simulation studies indicate that the developed MML estimators under the TPGT assumption are robust to the presence of outliers. We demonstrate that they are more efficient (unbiased and smaller variance) than some existing estimators for small samples, especially for non-normal error distribution. Additionally, we consider residual analysis for the TPGT-CR model based on the modified deviance residual to assess model fit and/or identify outliers in the data.

The structure of the paper is as follows. Section 2 briefly outlines some basic features of the TPGT distribution. In Sect. 3, we propose a linear regression model based on the two-piece generalized t distribution for censored data and provide a modified maximum likelihood method to estimate the TPGT-CR model. The fisher information matrix are derived to obtain the standard errors of the MML estimates. In Sect. 4, we introduce two types of residuals to check the model assumptions and the presence of outliers. In Sect. 5, we provide various simulation studies to examine the performance of our proposed MML method. In Sect. 6, a real data set is analyzed to illustrate the proposed methodology. Finally, some concluding remarks are given in the last section.

2 Preliminaries

In this section, some basic definitions, theorems and properties of the TPGT distribution are outlined. The TPGT distribution introduced by Lian et al. (2024) has the following probability density function (pdf)

$$f_{TPGT}(x;\mu,\sigma,r,a,b) = \frac{b}{2\sigma(2a)^{1/b}B(a,1/b)}\left\{1+\frac{|x-\mu|^b}{2a\sigma^b[1+r\,\mathrm{sign}(x-\mu)]^b}\right\}^{-(a+1/b)}, \quad (1)$$

where $x \in \mathbb{R}$, $B(\cdot,\cdot)$ denotes the beta function, $\mu \in \mathbb{R}$, $\sigma > 0$ and $|r| < 1$ determine the location, scale and the skewness, respectively, and $a,\ b > 0$ control the shape of the density function. A random variable X with pdf as in formula (1) will be denoted by $X \sim TPGT(\mu, \sigma, r, a, b)$.

The TPGT density (1) is continuous and unimodal with a mode at μ. It is positively (negatively) skewed for $r > 0$ (< 0). Following the definition, the TPGT distribution nests many important distributions as special cases. As $a \to +\infty$, the TPGT distribution reduces to the two-piece generalized normal distribution. As $a \to +\infty, r = 0$ and $b = 1$, it gives $Laplace(\mu, 2\sigma)$ distribution. As $\mu = 0, \sigma = 1, r = 0, a = n/2$ and $b = 2$, it gives a Student's t distribution with n degrees of freedom. As $r \to 1$ and $b = 1$, it gives $Pareto(II)(\mu, 2\sigma a, a)$ distribution. As $r = 0$, it gives the symmetric generalized t distribution.

By allowing a skewing factor to describe the skewness and introducing two shape parameters to control kurtosis and tail heaviness, we can potentially capture more subtle features of the distribution compared to the one proposed by Azzalini and Capitanio (2003). This enhances our ability to describe tail phenomena and improves predictions of quantities like Expected Shortfall which rely on the shape of the tail. The TPGT distribution provides a powerful tool to model data having both heavy tails and high kurtosis due to its flexibility.

The stochastic representation of $X \sim TPGT(\mu, \sigma, r, a, b)$, as introduced in Lian et al. (2024), is given by

$$X \overset{d}{=} \mu + \sigma 2^{1/b} W Y^{1/b} Z^{1/b}, \tag{2}$$

where $W,\ Y$ and Z are independent random variables. W is a discrete variable assuming two states with the following probability function $f_W(w|r) = \frac{(r+1)}{2}\ \mathbb{I}\{w = r+1\} + \frac{(1-r)}{2}\ \mathbb{I}\{w = r-1\}$, $Y \sim Ga(1/b, 1)$ and $Z \sim IG(a, a)$. Here, $\mathbb{I}\{A\}$ denotes the indicator function of A, $Ga(\alpha, \lambda)$ represents the gamma distribution with mean value α/λ and variance α/λ^2, and $IG(\cdot,\cdot)$ denotes the inverse gamma distribution. It should be noted that this stochastic representation is essential for generating random numbers and calculating the kth moments of the TPGT distribution. For further details, please refer to Lian et al. (2024).

Let $X_0 = (X - \mu)/\sigma$, then the pdf and cumulative distribution function (cdf) of the standard TPGT distribution are respectively given by

$$f_0(x; r, a, b) = \frac{b}{2(2a)^{1/b} B(a, 1/b)} \left\{1 + \frac{|x|^b}{2a[1 + r\,\mathrm{sign}(x)]^b}\right\}^{-(a+1/b)}, \tag{3}$$

$$F_0(x; r, a, b) = \begin{cases} \left(\frac{1-r}{2}\right) I\left(\left(1 + \frac{|x|^b}{2a[1+r\cdot\mathrm{sign}(x)]^b}\right)^{-1}; a, 1/b\right), & x \le 0, \\ 1 - \left(\frac{1+r}{2}\right) I\left(\left(1 + \frac{|x|^b}{2a[1+r\cdot\mathrm{sign}(x)]^b}\right)^{-1}; a, 1/b\right), & x > 0, \end{cases} \tag{4}$$

where $I(y; a, b) = \frac{1}{B(a,b)} \int_0^y t^{a-1}(1-t)^{b-1} \mathrm{d}t$ denotes the incomplete beta function.

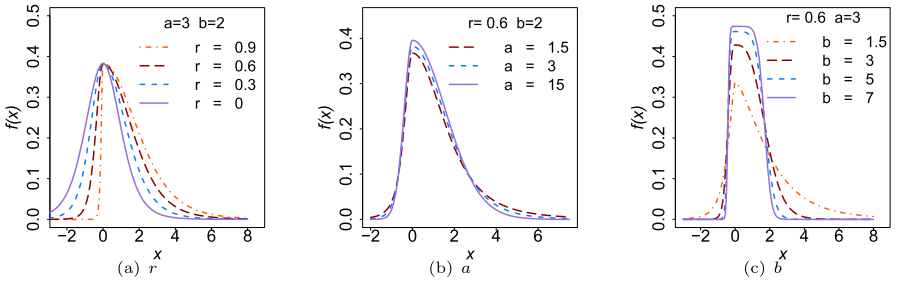


Fig. 1 The probability density functions of the standard TPGT distribution with different values of r, a and b

In addition for the purpose of visualization, we have also plotted the pdfs of the standard TPGT distribution with different values of r, a and b in Fig. 1. Figure 1a illustrates that the density curve is skewed to the right for $r > 0$ and to the left for $r < 0$. As r approaches $+1$ (-1), the density function becomes a right (or left) half-density. Figure 1b shows the pdfs of the standard TPGT distribution with $r = 0.6$, $b = 2$ and varying a. As a decreases, the tails of the pdfs become heavier. Figure 1c shows the pdfs of the standard TPGT distribution with $r = 0.6$, $a = 3$ and varying b. As b increases, the tails of the pdfs become shorter and lighter, the pdfs becomes flatter around the center and the peaks turn to be higher. Figure 1b and c show that two shape parameters (a and b) directly regulate the tail behavior and the peakedness of the density.

3 The TPGT-CR model

3.1 The model

In the CR model, we assume that the error terms follow a TPGT distribution rather than the normal distribution. Consider first the uncensored scenario and the multiple linear regression model

$$Y_i = \beta_0 + \boldsymbol{x}_i^\top \boldsymbol{\beta} + \epsilon_i, \quad \epsilon_i \overset{iid}{\sim} \text{TPGT}\,(0, \sigma, r, a, b)\,, \quad i = 1, \cdots, n, \tag{5}$$

where Y_i is a continuous response variables for the ith subject, $\boldsymbol{x}_i = (x_{i1}, \ldots, x_{ip})^\top$ is a known covariate vector, $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_p)^\top$ denotes an unknown parameter vector and $\{\epsilon_i\}_{i=1}^n$ is a sequence of independent random errors. Let $\mu_i = \beta_0 + \boldsymbol{x}_i^\top \boldsymbol{\beta}$, then $Y_i \sim TPGT(\mu_i, \sigma, r, a, b)$.

We assume that the response variable Y_i is not fully observed for all subjects. Specifically, Y_i is a latent variable and the observed data (V_i, ρ_i) take the form

$$V_i = \max\{Y_i, c_i\} = \begin{cases} c_i, & \text{if } \rho_i = 1 \text{ (i.e. } Y_i \le c_i) \\ Y_i, & \text{if } \rho_i = 0 \text{ (i.e. } Y_i > c_i) \end{cases}, \; i = 1, \ldots, n, \tag{6}$$

where c_i is the known left censoring point. The censoring indicator $\rho_i = 1$ (or $\rho_i = 0$) means that the ith observation is censored (or not censored). Since the response Y_i is defined within the real numbers, extending to right-censored data is straightforward. It's noteworthy that if $Y_i \sim TPGT(\mu, \sigma, r, a, b)$, then $-Y_i \sim TPGT(-\mu, \sigma, -r, a, b)$. This allows us to treat right-censoring as left-censoring by transforming Y_i and c_i to $-Y_i$ and $-c_i$ respectively. We call the structure defined by (5) and (6) as the TPGT-CR model.

It is important to note that if ϵ_i are independent identically distributed TPGT with $\mu = 0,\ r = 0,\ a = +\infty,\ b = 2$, (i.e., $\epsilon_i \overset{iid}{\sim} N(0, \sigma^2)$), and all censoring point $c_i = 0$, TPGT-CR model corresponds to the well-known ″Tobit model″ studied by Tobin (1958). The resulting maximum likelihood (ML) estimator is commonly referred to as the Tobit estimator.

3.2 Maximum likelihood estimation

With the observed data $(v_i, \rho_i, \boldsymbol{x}_i)$, the likelihood function for the parameters $\beta_0, \boldsymbol{\beta}$ and σ can be expressed as follows

$$\ln L = \sum_{y_i \le c_i} \ln F_0\left(\frac{c_i - \beta_0 - \boldsymbol{x_i}^\top \boldsymbol{\beta}}{\sigma}\right) + \sum_{y_i > c_i} \left[\ln f_0\left(\frac{y_i - \beta_0 - \boldsymbol{x_i}^\top \boldsymbol{\beta}}{\sigma}\right) - \ln(\sigma)\right], \tag{7}$$

where $f_0(\cdot)$ and $F_0(\cdot)$ denote the pdf and cdf of $TPGT(0, 1, r, a, b)$, respectively. Then, the ML estimators of the parameters of the TPGT-CR model can be obtained by maximizing Eq. (7). However, due to the complexity of the log-likelihood function, it is difficult to solve the explicit solution of the ML estimator and make some analytical studies, especially with small sample sizes. In the next subsection, we will use the modified maximum likelihood method to estimate the parameters of the TPGT-CR model.

Note that the skewing factor r, the tail behavior parameter a and the shape parameter b are first taken to be fixed, then the MML methodology works; see for example Arslan and Genc (2009) and Acitas et al. (2021). The details of finding plausible values for r, a and b are reported in Sect. 3.4. After taking the partial derivatives of the $\ln L$ function in Eq. (7) with respect to $\beta_0, \boldsymbol{\beta}$ and σ, we obtain the following likelihood equation:

$$\begin{cases} \dfrac{\partial \ln L}{\partial \beta_0} = -\dfrac{1}{\sigma}\displaystyle\sum_{y_i \le c_i} h_0(z_i) + \dfrac{1}{\sigma}(a + \dfrac{1}{b})\sum_{y_i > c_i} h_1(z_i) = 0, \\ \dfrac{\partial \ln L}{\partial \boldsymbol{\beta}} = -\dfrac{1}{\sigma}\displaystyle\sum_{y_i \le c_i} h_0(z_i)\boldsymbol{x}_i + \dfrac{1}{\sigma}(a + \dfrac{1}{b})\sum_{y_i > c_i} h_1(z_i)\boldsymbol{x}_i = 0, \\ \dfrac{\partial \ln L}{\partial \sigma} = -\dfrac{1}{\sigma}\displaystyle\sum_{y_i \le c_i} h_0(z_i)z_i + \dfrac{1}{\sigma}(a + \dfrac{1}{b})\sum_{y_i > c_i} h_1(z_i)z_i - \dfrac{n - n_c}{\sigma} = 0, \end{cases} \tag{8}$$

where $z_i = \frac{v_i - \beta_0 - \boldsymbol{x}_i^\top \boldsymbol{\beta}}{\sigma}$, n_c is the number of censored subjects and

$$h_0(z_i) = \frac{f_0(z_i)}{F_0(z_i)}, \quad h_1(z_i) = \frac{b \cdot \text{sign}(z_i)|z_i|^{b-1}}{|z_i|^b + 2a[1 + r\text{sign}(z_i)]^b}. \tag{9}$$

The likelihood Eq. (8) involve nonlinear functions, $h_0(z_i)$ and $h_1(z_i)$, and do not have explicit solutions. Solving them through iteration presents challenges (Chen and Oliver 2012; Lee and Zhu 2002). The modified maximum likelihood estimation method addresses these problems.

3.3 Modified maximum likelihood estimation

We use MML methodology (Tiku 1967; Tiku and Suresh 1992) which alleviates computational difficulties and allows to obtain the explicit forms of the ML estimators. The steps of the MML methodology are explained as follows:

Step1. Take the rank of z_i and denote it as r_i for $i = 1, 2, \ldots, n$. Use $rank(\cdot)$ function in $\mathcal{R}$ for this purpose. Let $\{z_{(i)}\}_{i=1}^n$ denote the order statistics obtained by ordering $\{z_i\}_{i=1}^n$, then $z_i = z_{(r_i)}$.

Step2. Let $f_0^{'}(t_i) = -b(ab+1)\text{sign}(t_i)|t_i|^{b-1}[u(t_i)]^{a+\frac{1}{b}+1}/\{2(2a)^{1/b+1}[1 + r \cdot \text{sign}(t_i)]^b B(a, \frac{1}{b})\}$ and $u(t_i) = \left(1 + \frac{|t_i|^b}{2a[1+r\cdot\text{sign}(t_i)]^b}\right)^{-1}$. Linearize the nonlinear functions $h_0(\cdot)$ and $h_1(\cdot)$ around $t_i = E(z_{(r_i)})$ using the first two terms of Taylor series expansion:

$$h_0(z_i) \approx a_{0i} - b_{0i} z_i, \quad i = 1, \ldots, n_c, \quad h_1(z_i) \approx a_{1i} - b_{1i} z_i, \quad i = n_c + 1, \ldots, n, \tag{10}$$

where

$$\begin{aligned} a_{0i} &= h_0(t_i) + b_{0i} t_i, \quad b_{0i} = -\frac{f_0^{'}(t_i)}{F_0(t_i)} + \left[\frac{f_0(t_i)}{F_0(t_i)}\right]^2, \\ a_{1i} &= h_1(t_i) + b_{1i} t_i, \quad b_{1i} = -\frac{b(b \cdot u(t_i) - 1)|t_i|^{b-2}}{2a[1 + r \cdot \text{sign}(t_i)]^b + |t_i|^b}. \end{aligned} \tag{11}$$

Step 3. By incorporating the linearized versions of $h_0(\cdot)$ and $h_1(\cdot)$ functions into the likelihood equations, the following modified likelihood equation can be obtained

$$\begin{cases} \dfrac{\partial \ln L^*}{\partial \beta_0} \propto \displaystyle\sum_{y_i \le c_i} (a_{0i} - b_{0i} z_i) - (a + \frac{1}{b}) \sum_{y_i > c_i} (a_{1i} - b_{1i} z_i) = 0, \\ \dfrac{\partial \ln L^*}{\partial \boldsymbol{\beta}} \propto \displaystyle\sum_{y_i \le c_i} (a_{0i} - b_{0i} z_i)\boldsymbol{x}_i - (a + \frac{1}{b}) \sum_{y_i > c_i} (a_{1i} - b_{1i} z_i)\boldsymbol{x}_i = 0, \\ \dfrac{\partial \ln L^*}{\partial \sigma} \propto \displaystyle\sum_{y_i \le c_i} (a_{0i} - b_{0i} z_i) z_i - (a + \frac{1}{b}) \sum_{y_i > c_i} (a_{1i} - b_{1i} z_i) z_i + (n - n_c) = 0. \end{cases} \tag{12}$$

Step 4. The solutions of Eq. (12) are the following MML estimators:

$$\begin{cases} \hat{\beta}_{0MML} = \bar{y} - \bar{\boldsymbol{x}}^\top \hat{\boldsymbol{\beta}}_{MML} - \dfrac{a^*}{b^*}\hat{\sigma}_{MML}, \\ \hat{\boldsymbol{\beta}}_{MML} = \mathbf{K} + \mathbf{L}\hat{\sigma}_{MML}, \\ \hat{\sigma}_{MML} = \dfrac{-J_1 + \sqrt{J_1^2 + 4J_0J_2}}{2\sqrt{J_0(J_0-(p+1))}}, \end{cases} \tag{13}$$

where $a^* = \sum_{y_i \le c_i} a_{0i} - (a+\frac{1}{b})\sum_{y_i > c_i} a_{1i}$, $b^* = \sum_{y_i \le c_i} b_{0i} - (a+\frac{1}{b})\sum_{y_i > c_i} b_{1i}$, $\bar{y} = \frac{\sum_{y_i \le c_i} b_{0i}c_i - (a+\frac{1}{b})\sum_{y_i > c_i} b_{1i}y_i}{b^*}$, $\bar{\boldsymbol{x}}^\top = (\bar{x}_{\cdot 1}, \ldots, \bar{x}_{\cdot p})$, $\bar{x}_{\cdot j} = \frac{\sum_{y_i \le c_i} b_{0i}x_{ij} - (a+\frac{1}{b})\sum_{y_i > c_i} b_{1i}x_{ij}}{b^*}$, $j = 1, \ldots, p$, and

$$\mathbf{K} = \mathbf{D}^{-1}\mathbf{B}_2, \quad \mathbf{L} = -\mathbf{D}^{-1}\mathbf{B}_1,$$

$$\mathbf{B}_1 = \mathbf{M}^\top(\boldsymbol{\rho} \circ \boldsymbol{a}_0) - \left(a + \frac{1}{b}\right)\mathbf{M}^\top((\mathbf{1}_n - \boldsymbol{\rho}) \circ \boldsymbol{a}_1),$$

$$\mathbf{B}_2 = \mathbf{X}^\top(\boldsymbol{\rho} \circ \boldsymbol{b}_0 \circ (\mathbf{v} - \mathbf{1}_n\bar{y})) - \left(a + \frac{1}{b}\right) \cdot \mathbf{X}^\top(\mathbf{I}_n - D(\boldsymbol{\rho}))(\boldsymbol{b}_1 \circ (\boldsymbol{y} - \mathbf{1}_n\bar{y})),$$

$$\mathbf{D} = \mathbf{M}^\top \cdot D(\boldsymbol{\rho} \circ \boldsymbol{b}_0) \cdot \mathbf{M} - \left(a + \frac{1}{b}\right) \cdot \mathbf{M}^\top \cdot D(\boldsymbol{b}_1 \circ (\mathbf{1}_n - \boldsymbol{\rho})) \cdot \mathbf{M},$$

$$\boldsymbol{\rho} = (\rho_1, \rho_2, \cdots, \rho_n)^\top, \quad \mathbf{X}_{n\times p} = (\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_n)^\top,$$

$$\boldsymbol{a}_k = (a_{k1}, a_{k2}, \cdots, a_{kn})^\top, \quad \boldsymbol{b}_k = (b_{k1}, b_{k2}, \cdots, b_{kn})^\top, k = 0, 1,$$

$$\mathbf{M} = \left(m_{ij}\right)_{n\times p}, \quad m_{ij} = x_{ij} - \bar{x}_{\cdot j},$$

$$\boldsymbol{y} = (y_1, y_2, \cdots, y_n)^\top, \quad J_0 = n - n_c,$$

$$J_1 = (\boldsymbol{\rho} \circ \boldsymbol{a}_0)^\top(\mathbf{v} - \mathbf{1}_n\bar{y} + \mathbf{MK}) - \left(a + \frac{1}{b}\right) \cdot ((\mathbf{1}_n - \boldsymbol{\rho}) \circ \boldsymbol{a}_1)^\top(\boldsymbol{y} - \mathbf{1}_n\bar{y} + \mathbf{MK}),$$

$$J_2 = (\boldsymbol{\rho} \circ \boldsymbol{b}_0)^\top[(\mathbf{v} - \mathbf{1}_n\bar{y} - \mathbf{MK}) \circ (\mathbf{v} - \mathbf{1}_n\bar{y} - \mathbf{MK})]$$
$$- \left(a + \frac{1}{b}\right) \cdot ((\mathbf{1}_n - \boldsymbol{\rho}) \circ \boldsymbol{b}_1)^\top[(\boldsymbol{y} - \mathbf{1}_n\bar{y} - \mathbf{MK}) \circ (\boldsymbol{y} - \mathbf{1}_n\bar{y} - \mathbf{MK})].$$

Here $\mathbf{v} = (v_1, \cdots, v_n)$ and $D(\mathbf{x})$ is the diagonal matrix of the vector $\mathbf{x}$, $\circ$ denotes the Hadamard product, $\mathbf{I}_n$ is a $n \times n$ identity matrix, and $\mathbf{1}_n$ denotes a $n \times 1$ vector with element 1. It is clear that the MML estimator can be regarded as a weighted sum of data, with weights given by b_{0i} and b_{1i}.

It is well-known that the MML estimators are asymptotic equivalence to ML estimators. Theorem 1 establishes the equivalence for the TPGT-CR model. Theorem 2 gives the asymptotic distribution of the MML estimators defined in Eq. (13).

Theorem 1 *The MML estimators $\hat{\beta}_{0MML}$, $\hat{\boldsymbol{\beta}}_{MML}$ and $\hat{\sigma}_{MML}$ are asymptotically equivalent to the ML estimators.*

Proof See Appendix A.1. □

Theorem 2 (Limiting distribution of the MML estimators) *The MML estimators asymptotically follow a normal distribution, with the mean parameter as* $\boldsymbol{\theta}$ *and the covariance matrix as* $I^{-1}(\boldsymbol{\theta})$*, where* $I(\boldsymbol{\theta})$ *is defined in Sect.* 3.5.

Proof See Appendix A.2. □

Note that the differences $h_k(z_{(i)}) - (a_{ki} - b_{ki} z_{(i)})$, $k = 0, 1$ tend to zero as $n \to \infty$. This confirms the conclusion that the MML estimators in formula (13) are asymptotically equivalent to the ML estimators.

The plots of $h_0(z_{(i)})$, $h_1(z_{(i)})$, and the approximations obtained from Eq. (10) are shown in Fig. 2 for the TPGT distribution with sample size $n = 100$ and 10% censoring level. We observe that in symmetric cases with $r = 0$, the approximations closely match the original values for $h_0(z_{(i)})$ and $h_1(z_{(i)})$. On the other hand, in skewed cases with $r = 0.5$, the results indicate that due to the decreased smoothness of h_0 when $z_{(i)}$ is small, the effectiveness of the Taylor approximation diminishes. Similarly, the non-differentiability of h_1 near 0 leads to a decrease in the effect of the Taylor approximation. However, overall, Eq. (10) demonstrates satisfactory approximations for $h_0(\cdot)$ and $h_1(\cdot)$.

In summary, the MML estimators have several appealing properties: (i) The MML method is computationally more efficient than the ML method because the MML estimators are explicitly formulated; (ii) The MML estimators are asymptotically unbiased, asymptotically equivalent to the ML estimators, and have minimum variance bound; see Tiku and Suresh (1992); (iii) Fig. 3 shows that the MML estimators are highly robust to outliers since the weights b_{0i} and b_{1i} are small for larger and smaller observations.

From the theoretical perspective, the MML estimators developed under the assumption of TPGT distribution provide robust estimators and are insensitive to the initial estimators. In practical terms, the estimators obtained in this paper are reliable and. The proposed TPGT-CR model can be used for fitting continuous data with high kurtosis and strong skewness. It should be noted that the MML estimator was first presented in the context of the multiple linear regression model with TPGT distribution, which is another contribution of the current paper.

For comparison, we also consider the Tobit and SCLS (Powell 1986) estimators. The SCLS relies on conditional symmetry and unimodality of the error distribution in the case of truncation in CR model. This assumption leads to the SCLS estimator minimizing

$$\frac{1}{n}\sum_{i=1}^{n}\left\{\left[V_i - \max\left\{\frac{V_i}{2}, \beta_0 + \boldsymbol{x}_i^\top \boldsymbol{\beta}\right\}\right]^2 + \mathbb{I}\{V_i > 2(\beta_0 + \boldsymbol{x}_i^\top \boldsymbol{\beta})\}\left[\left(\frac{V_i}{2}\right)^2 - \max\left\{c, \beta_0 + \boldsymbol{x}_i^\top \boldsymbol{\beta}\right\}^2\right]\right\}.$$

Both estimators have been shown to be consistent and asymptotically normally distributed for a wide range of distributions. The use of numerical methods of optimization

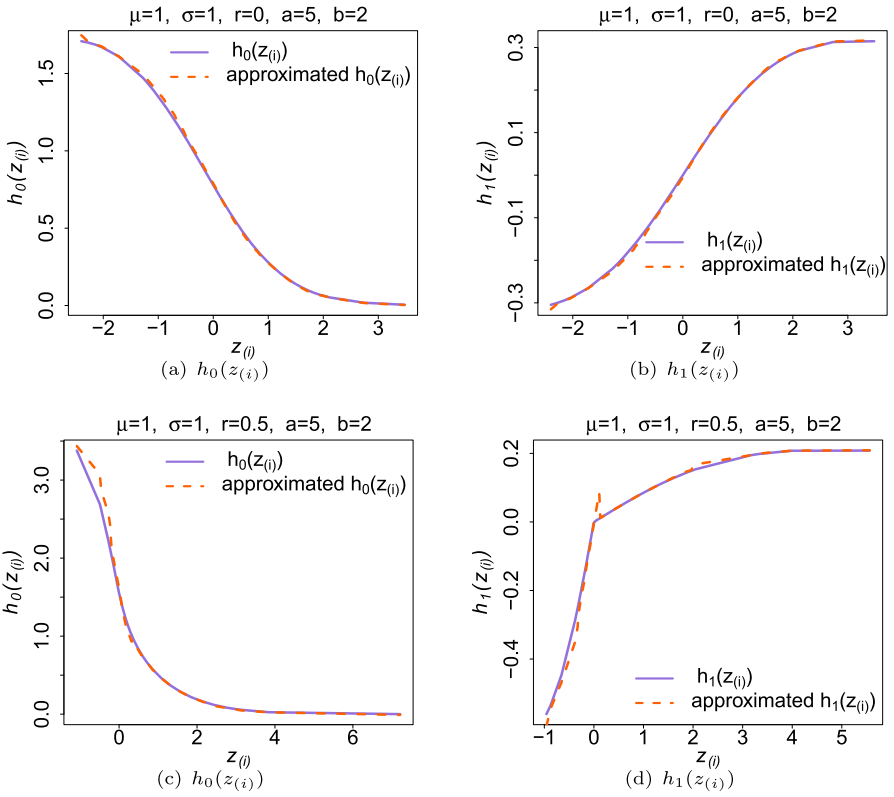


Fig. 2 Plotting of original and the approximated values of $h_0(z_{(i)})$ and $h_1(z_{(i)})$ for TPGT-CR model with $n = 100$ and 10% censoring level

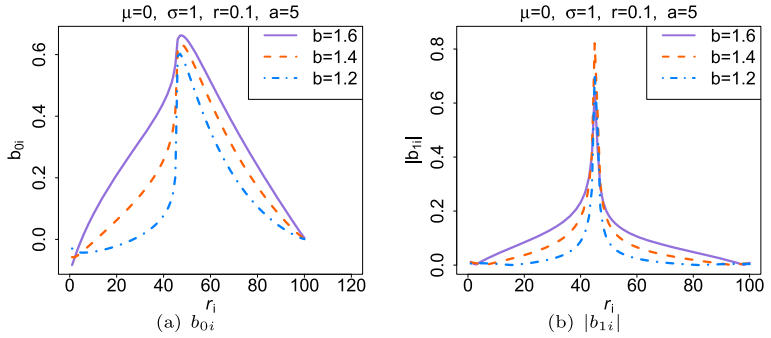


Fig. 3 Line graphs of the weights b_{0i} and $|b_{1i}|$ for TPGT-CR model with fixed sample size $n = 100$

are required to obtain the above estimates. We can use the $Nelder\text{-}Mead$ algorithm in $\mathcal{R}$.

3.4 Implementation details

It is worth noting that the MML method may give numerically unreliable results if all parameters are estimated simultaneously unless the sample size is large enough. As in many applications, the model parameters are what we focus on. The MML method succeeds in estimating the regression and scale parameters but it cannot give a satisfactory estimate for the distribution parameter; see Yalçınkaya et al. (2018). In this case, we have employed the user-specified approach described in Acitas et al. (2013), Arslan and Genc (2009), Lucas (1997). The parameters r, a and b can be fixed in advance or estimated from the data for a robust procedure. The following two strategies can be adopted.

Identification of the skewing factor and shape parameters

As suggested by one anonymous reviewer, it is essential to clarify the estimation for the three parameters r, a and b. As the first and simplest strategy, we can try different fixed values of r, a and b parameters until the largest value of the KS test statistic and the smallest values of model selection criteria have been obtained. The values of r, a and b can alternatively be identified by plotting the QQ-plots of residuals (see Acitas et al. 2013).

In the second strategy, we develop a two-step estimate procedure (comprising the ML step and the MML step, respectively) as follows.

1. Initialize $\hat{\beta}_0$, $\hat{\boldsymbol{\beta}}$, $\hat{\sigma}$ and estimate r, a, b by the ML estimate proposed by Lian et al. (2024);
2. Using the estimated values of parameters r, a, and b obtained in step 1 as initial values, optimize the likelihood function of TPGT-CR. Select the values of $\hat{r}$, $\hat{a}$, and $\hat{b}$ that maximize the likelihood function of TPGT-CR.
3. Calculate the MML estimates $\hat{\beta}_{0MML}$, $\hat{\boldsymbol{\beta}}_{MML}$, $\hat{\sigma}_{MML}$ for given $\hat{r}$, $\hat{a}$ and $\hat{b}$ values from step 2.

We adopted this strategy in this paper to determine the parameter values of r, a and b, in simulation studies and in real data analysis.

Notes on implementation

- Note that the denominator $2J_0$ was replaced by $2\sqrt{J_0(J_0-(p+1))}$ in $\hat{\sigma}_{MML}$ as a bias correction; see Arslan and Senoglu (2018).
- The final MML estimators in real data are obtained by the following steps: (a) For given r, a and b values, first obtain the Tobit estimates as the initial values for parameters β_0, $\boldsymbol{\beta}$ and σ to calculate the values of z_i, $i=1,\cdots,n$. Then the MML estimators are computed using steps 1–4 in Sect. 3.3 based on $\{z_i\}_{i=1}^n$.
  (b) Repeat the process in step (a), with the initial values for the parameters β_0, $\boldsymbol{\beta}$ and σ replaced by the MML estimates obtained in step (a). Note that two iterations are enough to stabilize the estimates (Acitas et al. 2021).

3.5 Standard error approximation

To estimate the standard error of the MML estimates of $\boldsymbol{\theta} = (\beta_0, \boldsymbol{\beta}^\top, \sigma)^\top$, we exploit an information-based method suggested by Arslan and Senoglu (2018), Lin (2010) and Vaughan and Tiku (2000). It has been proved that the likelihood equation is asymptotically equivalent to the modified likelihood equations. Thus, we can obtain the asymptotic covariance matrix of the MML estimators $\hat{\beta}_{0MML}$, $\hat{\boldsymbol{\beta}}_{MML}$ and $\hat{\sigma}_{MML}$ using the second-order derivatives of the modified likelihood equations. Notably, the information matrix $I(\boldsymbol{\theta})$ is symmetric by definition, as indicated by its structure. The explicit expression for the symmetric information matrix $I(\boldsymbol{\theta})$ is presented as follows

$$I(\boldsymbol{\theta}) = \begin{bmatrix} I_{11} & I_{12} & I_{13} \\ I_{21} & I_{22} & I_{23} \\ I_{31} & I_{32} & I_{33} \end{bmatrix} = \begin{bmatrix} -E\left(\frac{\partial^2 \ln L^*}{\partial \beta_0^2}\right) & -E\left(\frac{\partial^2 \ln L^*}{\partial \beta_0 \partial \boldsymbol{\beta}^\top}\right) & -E\left(\frac{\partial^2 \ln L^*}{\partial \beta_0 \partial \sigma}\right) \\ -E\left(\frac{\partial^2 \ln L^*}{\partial \boldsymbol{\beta} \partial \beta_0}\right) & -E\left(\frac{\partial^2 \ln L^*}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}^\top}\right) & -E\left(\frac{\partial^2 \ln L^*}{\partial \boldsymbol{\beta} \partial \sigma}\right) \\ -E\left(\frac{\partial^2 \ln L^*}{\partial \sigma \partial \beta_0}\right) & -E\left(\frac{\partial^2 \ln L^*}{\partial \sigma \partial \boldsymbol{\beta}^\top}\right) & -E\left(\frac{\partial^2 \ln L^*}{\partial \sigma^2}\right) \end{bmatrix},$$

and

$$\begin{aligned}
I_{11} &= \frac{b^*}{\sigma^2}, \quad I_{12} = \frac{b^*}{\sigma^2} \cdot \bar{\boldsymbol{x}}, \\
I_{13} &= \frac{1}{o^2}\left\{\sum_{y_i \le c_i} b_{0i} t_i - \left(a + \frac{1}{b}\right) \sum_{y_i > c_i} b_{1i} t_i\right\}, \\
I_{22} &= \frac{1}{\sigma^2}\left\{\sum_{y_i \le c_i} b_{0i} \cdot \boldsymbol{x}_i \boldsymbol{x}_i^\top - \left(a + \frac{1}{b}\right) \sum_{y_i > c_i} b_{1i} \cdot \boldsymbol{x}_i \boldsymbol{x}_i^\top\right\}, \\
I_{23} &= -\frac{1}{\sigma^2}\left\{\sum_{y_i \le c_i} (a_{0i} - b_{0i} t_i)\, \boldsymbol{x}_i - \left(a + \frac{1}{b}\right) \sum_{y_i > c_i} (a_{1i} - b_{1i} t_i)\, \boldsymbol{x}_i\right\}, \\
&\quad + \frac{1}{\sigma^2}\left\{\sum_{y_i \le c_i} b_{0i} \boldsymbol{x}_i t_i - \left(a + \frac{1}{b}\right) \sum_{y_i > c_i} b_{1i} \boldsymbol{x}_i t_i\right\}, \\
I_{33} &= \frac{-1}{\sigma^2}\left\{\sum_{y_i \le c_i} (2a_{0i} t_i - 3b_{0i} E(z_{(r_i)}^2)) \right. \\
&\quad \left. - \left(a + \frac{1}{b}\right) \sum_{y_i > c_i} (2a_{1i} t_i - 3b_{1i} E(z_{(r_i)}^2)) + n - n_c\right\}.
\end{aligned}$$

Note that $E(z_{(r_i)}^2)$ can be given by Section B. Then, standard error estimates of $\widehat{\boldsymbol{\theta}}$ can be obtained by inverting $I(\boldsymbol{\theta})$ under some regularity conditions. It is important to emphasize that the SE of r, a, and b depends heavily on the computation of expectations which relies on computationally intensive Monte Carlo integrations. In our paper, we focus solely on calculating the SE of $\boldsymbol{\theta}$.

4 Residual analysis

The residual analysis aims to identify outliers observations and to study departures from the error distribution assumption. To investigate departures from the assumption of errors and identify outliers, various residual analyses have been proposed in the literature [see Collett (2003) and Ortega et al. (2003)]. Pescim et al. (2017) and Carrasco et al. (2008) developed the analysis of residuals for a log-location regression model and a log-modified Weibull regression model. Following their works, we consider residual analysis for the TPGT-CR model based on the modified deviance residual defined as follows.

Let $S(y_i;\boldsymbol{\theta})$ denote the survival function of Y_i, then it can be estimated as

$$S(y_i;\widehat{\boldsymbol{\theta}}) = 1 - F_0\left(\frac{y_i - \widehat{\beta}_0 - \boldsymbol{x}_i^\top\widehat{\boldsymbol{\beta}}}{\widehat{\sigma}};r,a,b\right),$$

where $F_0(\cdot)$ is given in Eq. (4). The martingale residual was proposed in counting processes (see [10]). Thus, the martingale residual for the TPGT-CR model takes the form

$$r_{M_i} = \begin{cases} \log[S(y_i;\widehat{\boldsymbol{\theta}})] & \text{if } \rho_i = 1 \text{ (i.e. } y_i \le c_i)\,, \\ 1 + \log[S(y_i;\widehat{\boldsymbol{\theta}})] & \text{if } \rho_i = 0 \text{ (i.e. } y_i > c_i)\,, \end{cases} \quad i = 1,\ldots,n. \tag{14}$$

Due to the skewed form of the distribution of r_{M_i} ($-\infty \le r_{M_i} \le 1$), transformations to produce a new residual symmetrically distributed around zero would be more appropriate for residual analysis. In this case the modified deviance residual (r_{D_i}) introduced by Collett (2003) is given by

$$r_{D_i} = \text{sign}(r_{M_i})\left\{-2[r_{M_i} + (1-\rho_i)\log(1-\rho_i - r_{M_i})]\right\}^{\frac{1}{2}},$$

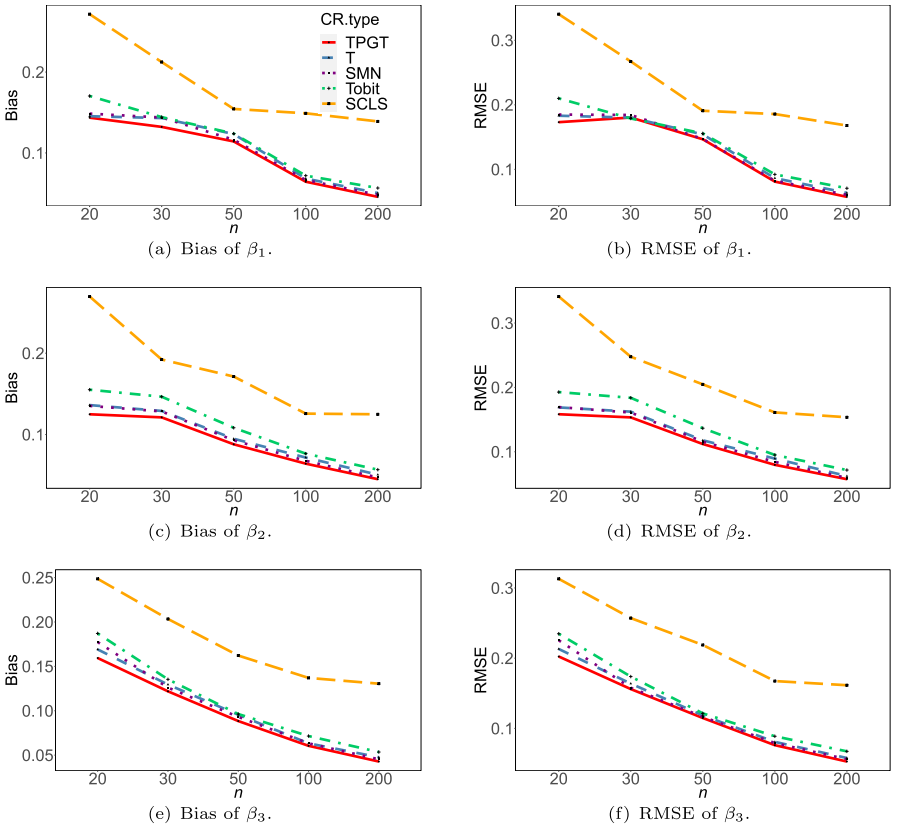
where r_{M_i} is the martingale residual defined in (14). The residual r_{D_i} applied to the TPGT-CR model can be expressed as

$$r_{D_i} = \begin{cases} \text{sign}\,(q_i)\,\{-2q_i\}^{\frac{1}{2}} & \text{if } \rho_i = 1, \\ \text{sign}\,(1+q_i)\,\{-2-2q_i-2\log(-q_i)\}^{\frac{1}{2}} & \text{if } \rho_i = 0, \end{cases} \tag{15}$$

where $q_i = \log[S(y_i;\widehat{\boldsymbol{\theta}})]$.

5 Monte Carlo simulation

In this section, a series of simulation studies are conducted to investigate the performance of the proposed MML estimates in Sect. 3.3. Three studies are designed to demonstrate the effectiveness of the proposed MML estimates, especially in dealing with highly skewed and heavily tailed data, to check the robustness of the MML estimates when there exist plausible deviations from an assumed model as well as its



(a) Bias of β_1. (b) RMSE of β_1.

(c) Bias of β_2. (d) RMSE of β_2.

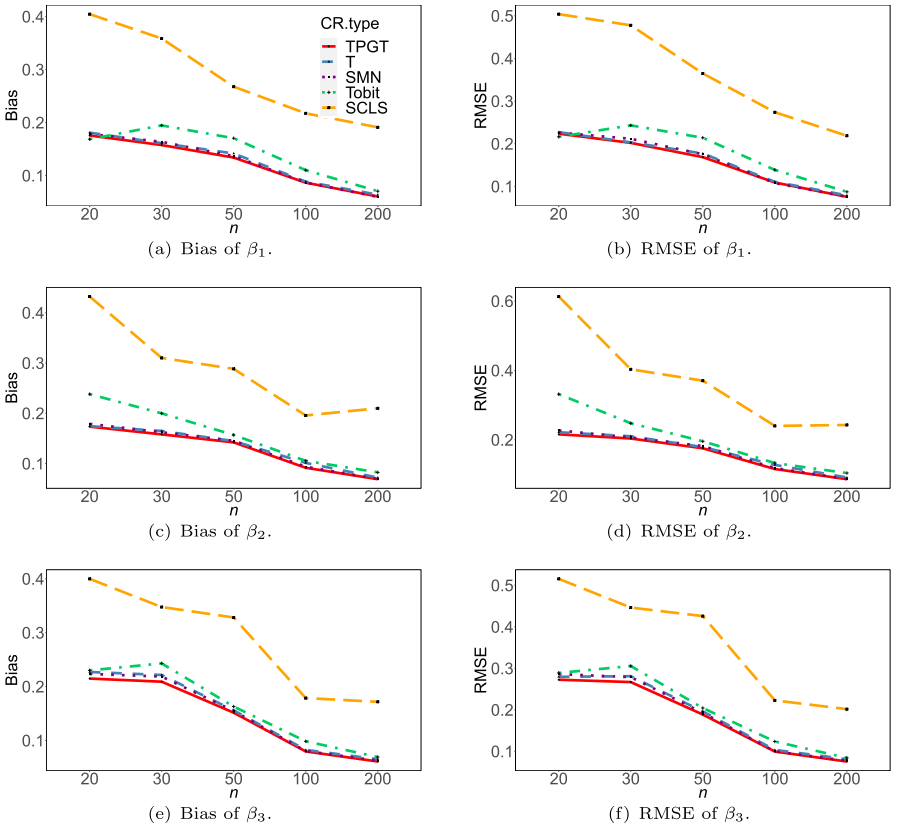
(e) Bias of β_3. (f) RMSE of β_3.

Fig. 4 Design I: Bias and RMSE of different estimators of slope coefficient (true $\beta_i = 1,\ i = 1, 2, 3$) in 1000 random trials for TPGT-CR model with $n = 20, 30, 50, 100, 200$ observations when the level of censoring is 10%

sensitivity in the presence of outliers, and to verify the asymptotic normality of the MML estimates. For the sake of data generation, one of the simplest ways to generate left-censored data is to consider $c_i = \min\{Y_i + U_i^{(2)}, Y_i - U_i^{(1)} + 1\}$ as recommended in Gómez et al. (2009), Mirfarah et al. (2021). Here $U_i^{(1)}$ and $U_i^{(2)}$ are two independent continuous variables followed by standard uniform distribution $\mathcal{U}(0, 1)$. For the random samples $y_1, \cdots, y_n$ generated from the TPGT-CR model (5), the following steps are used to obtain a $k\%$ left-censored dataset.

(1) Compute the number of censored samples $\mathcal{N}_c = [n \times k] + 1$, and then generate an index set $\mathcal{IND}$ of size $\mathcal{N}_c$ from $\{1, 2, \cdots, n\}$. Use $\mathcal{R}$ function $sample(\cdot)$ without replacement for this purpose.

(2) For $i = 1, \ldots, n$, if $i \in \mathcal{N}_c$, we first generate two random samples $u_i^{(1)}$ and $u_i^{(2)}$ independently from $\mathcal{U}(0, 1)$, and then set the thresholds to $c_i = \min\{y_i + u_i^{(2)}, y_i - u_i^{(1)} + 1\}$.



(a) Bias of β_1. (b) RMSE of β_1.

(c) Bias of β_2. (d) RMSE of β_2.

(e) Bias of β_3. (f) RMSE of β_3.

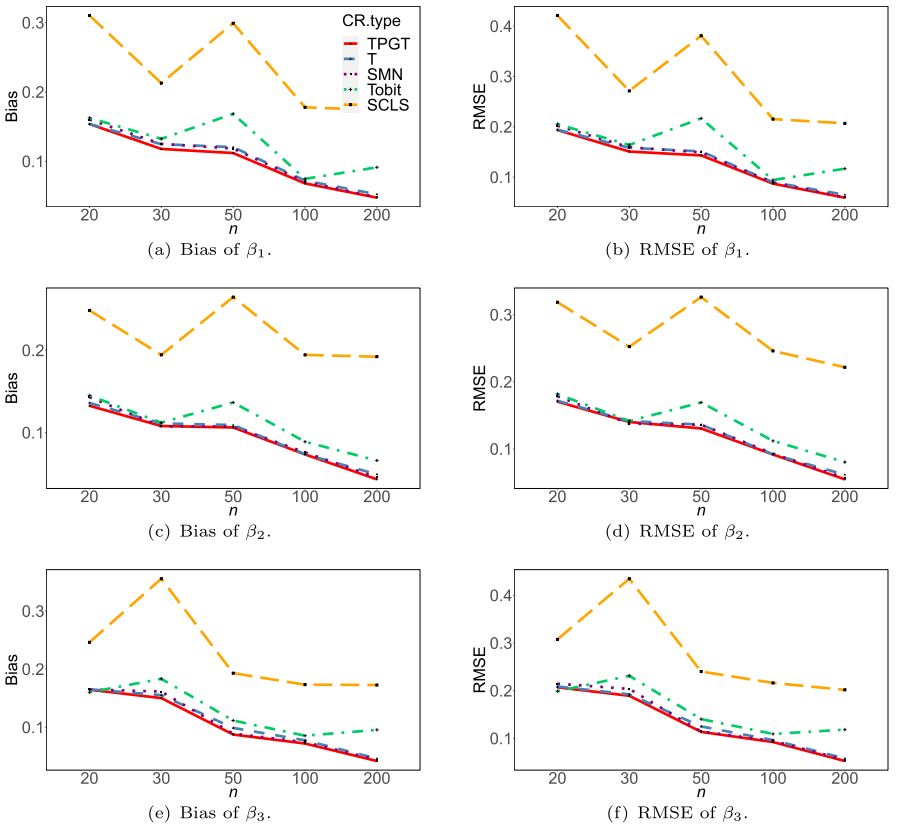
Fig. 5 Design II: Bias and RMSE of different estimators of slope coefficient (true $\beta_i = 1,\ i = 1, 2, 3$) in 1000 random trials for TPGT-CR model with $n = 20, 30, 50, 100, 200$ observations when the level of censoring is 10%

5.1 Finite sample properties of the MML estimates

In this subsection, we present a Monte Carlo study to assess finite-sample properties of the proposed MML estimators of the TPGT-CR model. The proposed TPGT-CR is compared with Student-t censored regression (T-CR) model (Arellano-Valle et al. 2012), SMN-CR model (Garay et al. 2017), Tobit models and SCLS estimates. The Tobit model is implemented using the *tobit* function from the AER package. To evaluate the bias and efficiency of the MML estimators, we compared these models under different levels of censoring and various error distributions. The specific design of the Monte Carlo study is as follows:

(1) We generate data using a multiple censored regression model as follows

$$Y_i = \beta_0 + \boldsymbol{x}_i^\top \boldsymbol{\beta} + \epsilon_i, \quad \epsilon_i \overset{iid}{\sim} \text{TPGT}\,(0, \sigma, r, a, b)\,, \quad i = 1, \cdots, n, \qquad (16)$$



(a) Bias of β_1. (b) RMSE of β_1.

(c) Bias of β_2. (d) RMSE of β_2.

(e) Bias of β_3. (f) RMSE of β_3.

Fig. 6 Design III: Bias and RMSE of different estimators of slope coefficient (true $\beta_i = 1,\ i = 1, 2, 3$) in 1000 random trials for TPGT-CR model with $n = 20, 30, 50, 100, 200$ observations when the level of censoring is 10%

where $\boldsymbol{x}_i = (x_{i1}, x_{i2}, x_{i3})^\top$, $x_{ik} \sim N(0, 1),\ k = 1, 2, 3$. Without loss of generality, in model (5), β_0 and σ are taken to be 1, and $\boldsymbol{\beta} = (1, 1, 1)^\top$.

(2) To explore the finite sample properties of our MML estimators across various error term distributions, ϵ_i is generated based on the following four designs.

- Design I: Symmetry with light tails: $\epsilon_i \sim TPGT(0, 1, 0, 20, 2)$, i.e. $\epsilon_i \sim N(0, 1)$.
- Design II: Lightly skewed with light tails: $\epsilon_i \sim TPGT(0, 1, 0.1, 6, 3)$.
- Design III: Moderately skewed with heavy tails: $\epsilon_i \sim TPGT(0, 1, -0.5, 6, 2)$.
- Design IV: Highly skewed with heavy tails: $\epsilon_i \sim TPGT(0, 1, -1, 20, 3)$.

The designs above encompass symmetric and skewed, light-tailed and heavy-tailed, as well as leptokurtic and platykurtic distributions.

(3) We wish to compare the small sample properties of the four estimators for the proposed censored regression model. Thus, for each model, we replicate the simu-

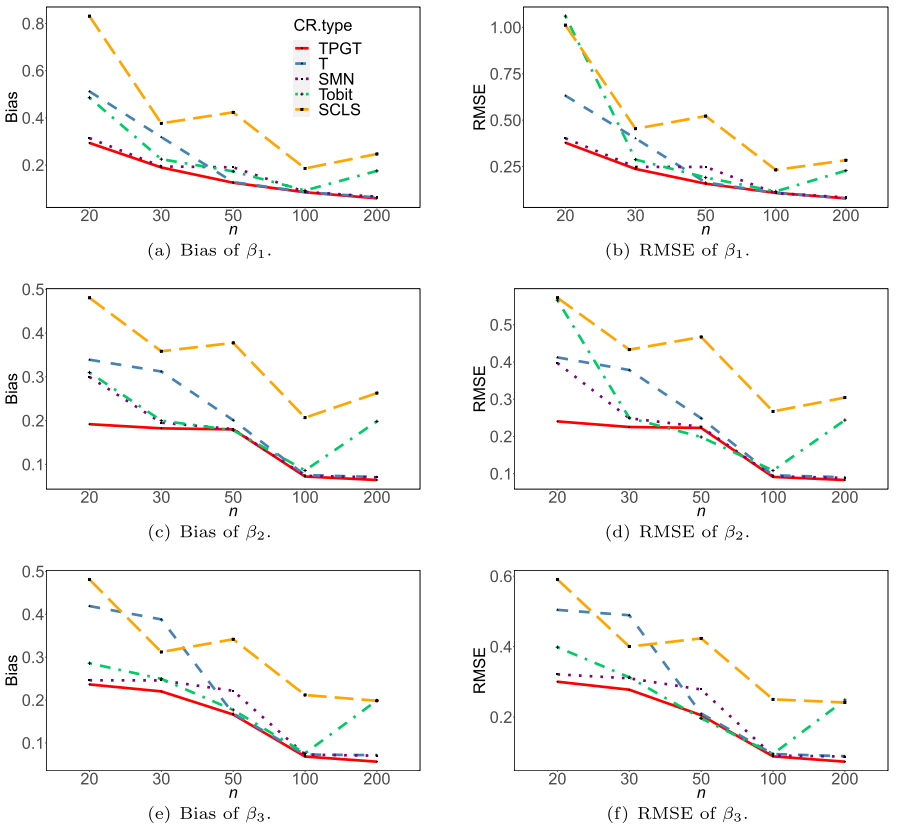


Fig. 7 Design IV: Bias and RMSE of different estimators of slope coefficient (true $\beta_i = 1,\ i = 1, 2, 3$) in 1000 random trials for TPGT-CR model with $n = 20, 30, 50, 100, 200$ observations when the level of censoring is 10%

lation $N = 1000$ times with sample size $n = 20, 30, 50, 100$ and 200. Further, Y_i was subjected to left-censoring. The levels of censoring are taken as low (10%), middle (20%) and high (30%). The left censoring point c_i is chosen such that the censoring rate equals the desired one.

(4) Finally, for each sample under different censoring levels, TPGT-CR, T-CR, SMN-CR, Tobit models and SCLS estimators were applied. To compare the finite sample performance of different methods, we report the absolute bias (Bias) and root mean square error (RMSE) of the parameter estimates of the slope coefficient $\boldsymbol{\beta}$ over the 1000 replicates:

$$\text{Bias}(\theta) = \frac{1}{M}\sum_{i=1}^{M}|\theta_i - \theta| \ \text{ and } \ \text{RMSE}(\theta) = \sqrt{\frac{1}{M}\sum_{i=1}^{M}(\theta_i - \theta)^2}.$$

All computational procedures were implemented using the statistical software $\mathcal{R}$.

In this Section, Fig. 4, 5, 6 and 7 summarize the Bias and RMSE of the estimators of the slope coefficient $\boldsymbol{\beta}$ obtained by TPGT-CR and four competing methods under a 10% censoring level. Results for the symmetric and light-skewed error distributions are reported in Fig. 4, 5, while those for moderate and highly skewed error distributions are shown in Fig. 6, 7. For the simulation results corresponding to censoring levels of 20% and 30%, please refer to the Appendix C.

At 10% censoring, as sample size n increases, the four methods except SCLS show comparable performance, as seen in Fig. 4. Notably, for symmetric distributions, the advantage of MML estimates of TPGT-CR is particularly evident in small samples. Figure 5 shows that for a lightly skewed error distribution, TPGT-CR, T-CR, and SMN-CR perform comparably well, all outperforming Tobit and SCLS. The SCLS estimator does worst and is associated with the largest Bias and RMSE. It is evident from Figs. 6, 7 that the MML estimates of the TPGT-CR perform best in terms of Bias and RMSE when the true distribution is moderately/highly skewed and heavy-tailed, under almost all sample sizes.

From the results shown in Fig. 4, 5, 6 and 7, we observe a decrease in Bias and RMSE of the MML estimators with increasing sample size n. However, Fig. 5 shows that the Bias and RMSE of SCLS no longer strictly increase with increasing n for parameter β_2. For parameters β_1 and β_3, Tobit estimates exhibit similar performance. Figures 6 and 7 indicate that for all slope parameters, the Bias and RMSE of Tobit and SCLS estimators no longer strictly decrease with increasing n. This is because the SCLS estimator's consistency requires a symmetric error term distribution, while Tobit estimation requires a normal error distribution (Tobin 1958; Powell 1986). In general, TPGT-CR exhibits significant advantages in fitting highly skewed and heavy-tailed data. For symmetric and lightly skewed data, TPGT-CR performs comparably to competing models.

In Appendix C, simulation results under censoring rates of 20% and 30% are given. From Fig. 10, it can be observed that the Tobit estimator performs best for a censoring rate of 20% when $n = 20$ and 30. However, in most cases, the TPGT-CR model maintains the best performance. For details, please refer to Appendix C.

5.2 Robustness of the MML estimates

In this subsection, we aim at investigating the robustness of the MML estimate against perturbation in the model (or data). In this regard, 1000 Monte Carlo samples of size $n = 200$ are simulated from the left-censored TPGT-CR model proposed in Sect. 3.1. We set $\beta_0 = 1$, $\boldsymbol{\beta} = (1, 1, 1)$, $\sigma = 1$, and $\boldsymbol{x}_i = (x_{i1}, x_{i2}, x_{i3})^\top$, where x_{ik} are generated from $N(0, 1)$ for $k = 1, 2, 3$. We assume the true error distribution is $TPGT(0, 1, 0.1, 6, 3)$ as Design II in Sect. 5.1. The considered censoring levels are 10%, 20% and 30%. The two aims of the simulation study in this subsection are: (i) to evaluate the impact of the outliers on the MML estimates; (ii) to explore the robustness of the MML against potential model misspecification. Thus, two experiments were carried out as follows.

Experiment 1 In the first experiment, for the generated original censored samples, we add a class of outliers with varying probability ranging from 2% to 6% as suggested

in Mirfarah et al. (2021). To do so, the responses of outliers were set to the minimum observation, namely $y_i^* = y_{min}$. To evaluate the performance of different estimation procedures based on 1000 Monte-Carlo runs, the mean squared error (MSE; Acitas et al. 2021) of the slope parameter vector $\boldsymbol{\beta}$ is calculated as

$$MSE = \frac{1}{1000}\sum_{s=1}^{1000} ||\widehat{\boldsymbol{\beta}}^s - \boldsymbol{\beta}||_2^2,$$

where $\widehat{\boldsymbol{\beta}}^s$ denotes the estimates in sth replication and $||\cdot||$ denotes L_2 norm. Table 1 shows the MSE values obtained using different estimation procedures, including TPGT-CR, T-CR, SMN-CR, Tobit, and SCLS methods under various censoring levels and the percentage of outliers in the data. The values of MSE of MML estimation increase as the percentage of outliers and level of censoring increases. It should be noted that the MML method provides the smallest MSE for all situations. It highlights that the MML estimates are much more robust to outliers due to the small weights assigned to the extreme observations. For other sample sizes n and error distributions, similar conclusions are obtained, so they are not shown to save space but can be provided upon request of the authors.

Experiment 2 In the second experiment, we are interested in examining the robustness of the proposed methodology against possible model misspecification. First, we independently generated 1000 datasets $(\mathbf{x}_i, y_i),\ i = 1, \ldots, n$ with $n = 200$, where ϵ_i are randomly generated from the following alternative models:

Model 1: TPGT model with misspecified shape parameters: $TPGT(0, 1, 0.1, a = 4, b = 2)$;
Model 2: Mixture model: $0.9TPGT(0, 1, 0.1, 6, 3) + 0.1TPGT(0, 2, 0.1, 6, 3)$;
Model 3: Contaminated model: $0.9TPGT(0, 1, 0.1, 6, 3) + 0.1N(-1, 1)$;
Model 4: Contaminated model: $0.9TPGT(0, 1, 0.1, 6, 3) + 0.1\chi_2^2$.

Here, χ_k^2 denotes the chi-squared distribution with k degrees of freedom. For models 2–4, we generated n observations, $0.9 * n$ of which come from the former model and $0.1 * n$ of them come from the latter one. Then, we estimated the parameters of the TPGT-CR model under the assumption that $\epsilon_i \sim TPGT(0, 1, 0.1, 6, 3)$ and computed the MSE of $\boldsymbol{\beta}$. The simulated MSE values are tabulated in Table 2 under the TPGT-CR and alternative models when the level of censoring is 10%, 20%, 30% and $n = 200$.

Table 2 reveals the following findings. Firstly, the MML estimates of the TPGT-CR model has the smallest MSE among the five methods considered. Secondly, while the values of MSEs for the TPGT-CR model increase as the level of censoring increases, TPGT-CR consistently performs reasonably well even under higher censoring levels when there is a plausible deviation from the assumed model. This indicates that the MML estimates under TPGT-CR model is robustness to model misspecification. Lastly, the TPGT-CR model outperforms others for all scenarios, followed by SMN-CR, T-CR and Tobit; SCLS performs the worst, indicating that SCLS depends heavily on the model assumptions.

Table 1 The MSEs for TPGT-CR, T-CR, SMN-CR, Tobit and SCLS estimation procedures for slope parameter $\boldsymbol{\beta}$: $n = 200$, with varying censoring levels and outlier percentages

Methods ↓	10% cens.			20% cens.			30% cens.		
Outliers	2%	4%	6%	2%	4%	6%	2%	4%	6%
TPGT-CR	0.0471	0.0511	0.0524	0.0513	0.0571	0.0624	0.0562	0.0628	0.0687
T-CR	0.0651	0.0688	0.0691	0.0714	0.0710	0.0779	0.0762	0.0787	0.0826
SMN-CR	0.0625	0.0673	0.0674	0.0652	0.0672	0.0739	0.0678	0.0695	0.0759
Tobit	0.0665	0.0683	0.0715	0.0666	0.0714	0.0718	0.0667	0.0718	0.0730
SCLS	0.1536	0.3165	0.3855	0.1556	0.3749	0.4424	0.1614	0.4332	0.5546

Table 2 Simulation results evaluating the robustness of TPGT-CR, T-CR, SMN-CR, Tobit, and SCLS estimation procedures against possible model misspecification under varying censoring levels with $n = 200$

Methods ↓	Model 1			Model 2			Model 3			Model 4		
Cens.Level	10%	20%	30%	10%	20%	30%	10%	20%	30%	10%	20%	30%
TPGT-CR	0.0740	0.0762	0.0766	0.0524	0.0551	0.0555	0.0617	0.0637	0.0646	0.0587	0.0616	0.0621
T-CR	0.0759	0.0822	0.0883	0.0581	0.0649	0.0699	0.0666	0.0679	0.0755	0.0626	0.0686	0.0728
SMN-CR	0.0777	0.0779	0.0793	0.0546	0.0585	0.0613	0.0644	0.0645	0.0690	0.0594	0.0635	0.0638
Tobit	0.0835	0.0850	0.0855	0.0650	0.0674	0.0680	0.0648	0.0658	0.0672	0.0871	0.0879	0.0896
SCLS	0.2083	0.2093	0.2136	0.1534	0.1586	0.1611	0.1857	0.1858	0.1930	0.1576	0.1578	0.1604

5.3 The asymptotic normality of the MML estimators

In this subsection, we conducted a simple simulation study to evaluate the asymptotic normality of the MML estimators. We independently simulated 1000 data sets $(\mathbf{x}_i, y_i),\ i = 1, \ldots, n$ with $n = 100, 200, 500, 700$ from the proposed TPGT-CR model with 10% censoring level, where $\epsilon_i \sim TPGT(0, 1, 0.1, 6, 3)$ and the design considered here is the same as the model discussed in Sect. 5.2. To examine the asymptotic normality of the MML estimators, Fig. 18 in Appendix C shows the QQ-plots for a normal fit to the 1000 standardized MML estimators of β_0, $\boldsymbol{\beta}$, σ, as well as the 95% confidence envelopes. The diagonal represents the line drawn when the theoretical quantiles perfectly match the sample quantiles. All these plots fall within the confidence bands of the normal distribution for the larger sample size $n = 200, 500$ and 700, which clearly supports that MML estimators are asymptotically normally distributed.

It can be seen that some points in QQ-plots of $\beta_1, \beta_2, \beta_3$ with $n = 100$ fall outside the theoretical 95% confidence envelopes. Thus, a comparison is also carried out between the empirical distribution and the theoretical distribution based on the corresponding Kolmogorov-Smirnov (K-S) test statistic to verify the results in Fig. 18. Table 3 lists the p-values of K-S normality tests. Note that the smallest p-value is 0.0823, which indicates that we can accept the null hypothesis of being a normal distribution at the 0.05 significance level. The small p-values of $\beta_1, \beta_2, \beta_3$ for $n = 100$ are observed in Table 3. All results match the conclusion from Fig. 18.

Based on the above findings, we can conveniently construct asymptotic confidence intervals and hypothesis tests for the parameters of interest for the proposed TPGT-CR model.

5.4 Residual analysis

In this simulation study, we investigate the empirical distributions of the modified deviance residuals (r_{D_i}) for different values of n and censoring levels. We consider $n = 50$, 200, and 700 (close to the sample size of real data), and 10%, 20% and 30% censoring levels. We set $\beta_0 = 1$, $\boldsymbol{\beta} = (1, 1, 1)$, $\sigma = 1$, and $\boldsymbol{x}_i = (x_{i1}, x_{i2}, x_{i3})^\top$, where x_{ik} are generated from $N(0, 1)$ for $k = 1, 2, 3$. The data were generated using the TPGT-CR model in (5) and (6), where $\epsilon_i \sim TPGT(0, 1, 0.1, 6, 3)$.

For each configuration of n and censoring level, 1000 samples were generated, and each one was fitted under the TPGT-CR model using the proposed MML method. For each fit, the residuals r_{D_i} were calculated. Figure 19 in Appendix C shows the normal QQ-plots of the the modified deviance residuals. As we can see from Fig. 19, the empirical distribution of the residuals r_{D_i} presents a better agreement with the standard normal distribution for a large sample $n = 200$ and 700. It can be observed that, as the censoring levels decrease and n increases, the empirical distribution of the residual r_{D_i} approaches closer to the normal distribution.

Table 3 P-values from the Kolmogorov-Smirnov normality test under various censoring levels and sample sizes n

MML estimator ↓	10% cens.				30% cens.			
n	100	200	500	700	100	200	500	700
$\hat{\beta_0}$	0.9175	0.5228	0.5932	0.7469	0.9959	0.4974	0.7552	0.9895
$\hat{\beta_1}$	0.0823	0.6057	0.9187	0.2385	0.5178	0.4876	0.7883	0.8269
$\hat{\beta_2}$	0.3057	0.9638	0.6585	0.9691	0.8416	0.9775	0.9026	0.9485
$\hat{\beta_3}$	0.1520	0.7574	0.7055	0.8439	0.7795	0.6863	0.4411	0.9942
$\hat{\sigma}$	0.3538	0.8493	0.7731	0.8870	0.5873	0.6256	0.6069	0.9374

6 An empirical example: the wage rates data

To illustrate the proposed regression methodology, we use a dataset obtained from Mroz (1987), which is available in the $\mathcal{R}$ package $CensRegMod$. The code can be found at https://github.com/chengdi588/Modified-MLE-of-TPGT-CR-model.git. The dataset contains 753 wage rates (hours worked outside the home) and several other characteristics of married white women between the ages of 30 and 60 in 1975, of whom 325 worked zero hours. It is important to emphasize that wage rates are assumed to be zero for wives not working in 1975, (i.e. they are censored or simply unobserved), which means 43.16% degree of censoring. This is a common assumption in economics; see DaVanzo and Lee (1978). This dataset has been previously analyzed by Arellano-Valle et al. (2012) using the T-CR model, by Caudill (2012), Karlsson and Laitila (2014) and Zeller et al. (2019) to illustrate the performance of the finite mixture of censored regression models based on the normal distribution. In this paper, we revisit this dataset in order to expand the inferential results to the TPGT family and to evaluate the performance of MML estimates.

Following the work of Arellano-Valle et al. (2012), we consider the housewife's wage rates as the response variable (y), and the explanatory variables include the wife's age (x_1), the wife's education in years (x_2), the number of kids younger than 6 years old in the household (x_3) and the number of children between the ages of six and nineteen (x_4). Thus, the model is given by

$$y_i = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \epsilon_i, \quad i = 1, \ldots, 753.$$

We consider that $\epsilon_i \sim TPGT(0, \sigma, r, a, b)$. Before employing the MML method to estimate the TPGT-CR model, we first need to identify the plausible values of the skewing factor r, and two shape parameters a and b. The second strategy explained in Sect. 3.4 was adopted, and r, a, b are identified as -0.4822, 1.5 and 1.7, respectively. Based on $\hat{r}$, $\hat{a}$ and $\hat{b}$, model parameters β_0, $\boldsymbol{\beta}$ and σ are estimated by using the MML method.

We compare our results with three comparative models, T-CR (Arellano-Valle et al. 2012), Tobit and skew-normal censored regression (SN-CR) models (Mattos et al. 2018) in terms of log-likelihood ($\ell(\widehat{\boldsymbol{\theta}})$), Akaike information criterion (AIC), Bayesian information criterion (BIC) and Efficient determination criterion (EDC). Table 4 shows the results. In all CR models, the consistent signs of the estimate of $\beta_1, \beta_2, \beta_3, \beta_4$ indicate that the effect of the four covariates on the response variable is consistent across the three models. Specifically, the wife's age (x_1), the number of children younger than 6 years old in the household (x_3), and the number of children between the ages of six and nineteen (x_4) all have negative effects on wage rates. Only the wife's education in years (x_2) has a positive effect on wage rates, with x_3 having the highest negative effect on y. This aligns with practical realities. Table 4 shows that the estimate of scale parameter σ of the SN-CR model is the largest. The values of standard errors of the parameters for the TPGT-CR model are calculated via the empirical information matrix presented in Sect. 3.5. It can be seen that the estimates of the TPGT-CR model have smallest standard errors, indicating that the TPGT-CR model can provide more accurate estimates.

Table 4 summaries the results of model selection criteria of the TPGT-CR, T-CR, Tobit and SN-CR models. It should be noted that the values of AIC, BIC, EDC of the proposed TPGT-CR model is smallest than that for the other two comparison models. This indicates that the TPGT-CR model provides a better fit to the wage rates data compared to the other models.

In addition, we conduct three likelihood ratio test (LRT) to test the following null hypotheses

$$H_0 : \text{error terms follow a normal distribution},$$
$$H_0 : \text{error terms follow a } t \text{ distribution},$$
$$H_0 : \text{error terms follow a SN distribution}$$

against the alternative H_1 : error terms follow a TPGT distribution. The LRT statistics are 110.4614, 27.4413 and 88.3544, respectively, which are highly significant compared to the critical values of the χ^2_2 and χ^2_3 distributions. Figure 8 shows the histogram of y overlaid with the best-fitted TPGT density curves and the corresponding probability-probability (PP) plot for the TPGT fit without covariates. It is clear to see that the fitted TPGT distribution adapts the shape of the histogram satisfactorily from Fig. 8b. Figure 8a indicates an underlying moderately skewed and heavy-tailed distribution, and thus it seems suitable for fitting the TPGT model to the data.

6.1 Residual analysis

To detect possible outliers in fitting the TPGT-CR models for wage rates data, Fig. 9 provides the index plot of the modified deviance residuals r_{D_i}. We observe that very few observations may be outliers, suggesting that the TPGT-CR model provides a good fit to wage rates data.

7 Conclusion and discussion

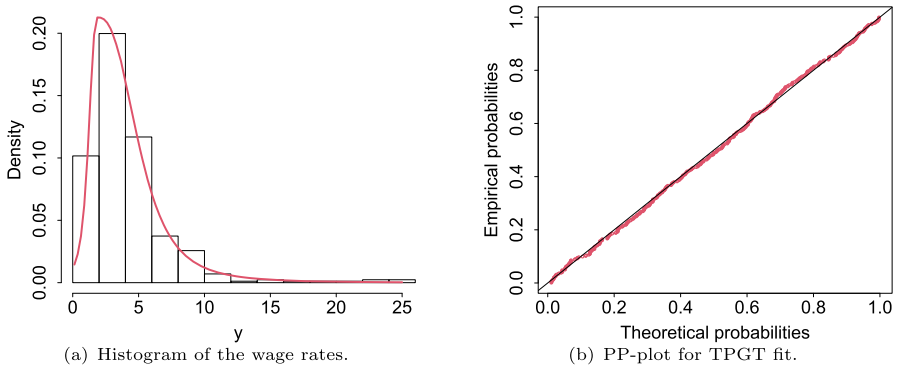
We have proposed a linear censored regression model based on the two-piece generalized t distributions, denoted by TPGT-CR models, as a replacement of symmetric distribution for censored regression models. The TPGT-CR model accommodates censoring to the responses, asymmetry and heavy tails, therefore, can be applied to a wide range of applications. We proposed a robust MML estimation approach in the random censoring setting, which includes more sources of data contamination (Li and Peng 2017; Salah and Yousri 2019). We derived the closed-form expressions of MML estimators based on the TPGT-CR model. Finite sample performance of the proposed MML estimators were evaluated through extensive simulation studies and an application to a wage rates dataset. The results demonstrated the superiority of our proposed method over with the traditional estimators such as Tobit, CLAD and SCLS.

There are diverse possible extensions of the current work. One limitation of the current work lies in that if the measurement errors are interval-censored, the current MML method may not be directly used to estimate such models. This is because the

Table 4 Wage rates data: Parameter estimates with corresponding approximate standard errors (SE) together with their log-likelihood ($\ell(\widehat{\boldsymbol{\theta}})$), AIC, BIC and EDC for fitted different models

Parameter	TPGT-CR	SE	T-CR	SE	Tobit	SE	SN-CR	SE
constant	1.9637	0.0495	−1.0471	1.4403	−2.7510	1.8899	−1.3589	1.7619
β_1	−0.1151	0.0136	−0.1108	0.0232	−0.1046	0.0282	−0.1182	0.0272
β_2	0.6317	0.0509	0.6475	0.0649	0.7281	0.0827	0.6910	0.0808
β_3	−2.0609	0.3332	−3.1637	0.3843	−3.0264	0.4196	−3.2309	0.4345
β_4	−0.3620	0.1015	−0.2964	0.1218	−0.2143	0.1494	−0.2586	0.1432
σ	3.0540	0.1107	3.2616	1.0118	4.5760	0.8106	5.7228	2.0125
r	−0.4822		–	–	–	–	–	–
a/ν	1.5000		4.1995	–	–	–	1.5430	0.4402
b	1.7000		–	–	–	–	–	–
$\ell(\widehat{\boldsymbol{\theta}})$	**−1426.4248**		−1440.1455		−1481.6555		−1470.6020	
AIC	**2864.8496**		2894.2909		2975.3110		2955.2030	
BIC	**2892.5940**		2926.6594		3003.0553		2987.5720	
EDC	**2885.7786**		2918.7081		2996.2400		2979.6210	

Bold font highlights the best model



(a) Histogram of the wage rates. (b) PP-plot for TPGT fit.

Fig. 8 Wage rates data. Histogram of the response variable y overlaid with TPGT density estimate and the PP-plot for the TPGT distribution

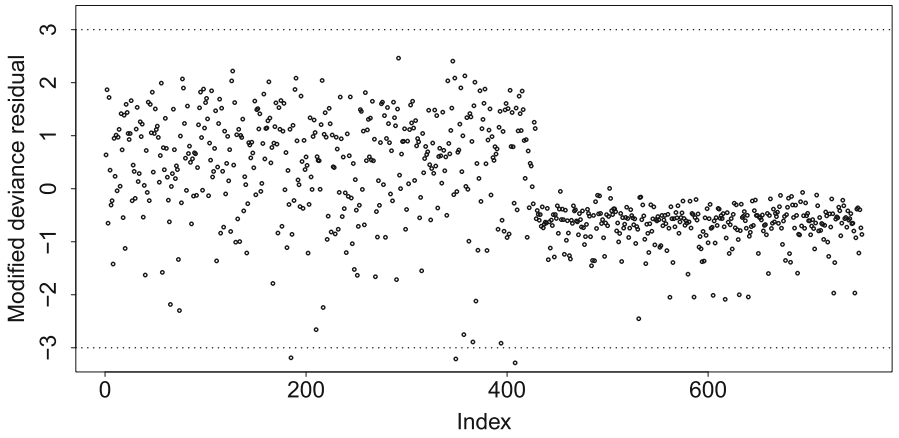


Fig. 9 Wage rates data. Index plot of the modified deviance residuals r_{D_i}

modified likelihood equations for interval-censored data are complicated, which adds to the complexity in implementation. In addition, it would be a worthwhile task to evaluate the qualitative robustness of the parameter estimators of β_0, $\boldsymbol{\beta}$ and σ when r, a and b are known or not. To modelling multivariate data, a further extension of the current work to the multivariate case and mixtures of linear experts model could also be developed via the use of copulas, see for instance (Fernández and Steel 1999; Wang 2023; Zeng et al. 2017).

Appendix A Proofs

A.1 Proof of Theorem 1

Proof *Uncensored case*: The asymptotic equivalence of the ML and MML estimators is established by applying a general result in Hoeffding (1953) regarding the sum of

a function evaluated at the expected values of order statistics. First, note that the absolute values of the following functions $w_1(z) = h_1(z)$, $w_2(z) = z \cdot h_1(z)$, $w_3(z) = h_0(z)$ and $w_4(z) = z \cdot h_0(z)$ are dominated by nonnegative convex functions with finite expectations (with respect to the TPGT distribution), where $h_0(\cdot), h_1(\cdot)$ are defined in (9). Hence, by the result of Hoeffding (1953), $\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} w_j\left(t_{i:n}\right) = E\left[w_j(Z)\right]$, $j = 1, 2, 3, 4$, where $Z \sim TPGT(0, 1, r, a, b)$, $t_{i:n} = E\left[z_{(i)}\right]$, $v_{i:n} = \mathrm{Var}\left(z_{(i)}\right)$ denote the expected value and variance of the i^{th} order statistic from a random sample of size n drawn from Z's distribution. Thus, we have

$$\begin{aligned} \lim_{n\to\infty} \sum_{i=1}^{n} \frac{h_1\left(t_{i:n}\right)}{n} &= E\left[h_1(Z)\right] = 0, \\ \lim_{n\to\infty} \sum_{i=1}^{n} \frac{h_1\left(t_{i:n}\right) t_{i:n}}{n} &= E\left[Z \cdot h_1(Z)\right] = \left(a + \frac{1}{b}\right)^{-1}. \end{aligned} \tag{A1}$$

From these results, it is easy to establish (a_{1i} and b_{1i} defined as before)

$$\begin{aligned} \lim_{n\to\infty} E\left\{\frac{1}{n} \sum_{i=1}^{n} \left(h_1\left(z_{(i)}\right) - a_{1i} + b_{1i} z_{(i)}\right)\right\} &= 0, \\ \lim_{n\to\infty} E\left\{\frac{1}{n} \sum_{i=1}^{n} \left(h_1\left(z_{(i)}\right) z_{(i)} - a_{1i} z_{(i)} + b_{1i} z_{(i)}^2\right)\right\} &= 0, \end{aligned}$$

Asymptotically, therefore, $\frac{1}{n} E\left(\frac{\partial \ln L^*}{\partial \beta_0}\right) = 0$, similarly, $\frac{1}{n} E\left(\frac{\partial \ln L^*}{\partial \boldsymbol{\beta}}\right) = 0$, and $\frac{1}{n} E\left(\frac{\partial \ln L}{\partial \sigma}\right) = 0$. The asymptotic equivalence of the ML and MML information matrices and the asymptotic unbiasedness of the estimators now follow. This is primarily the reason that $\frac{1}{n}\left|\frac{\partial \ln L}{\partial \beta_0} - \frac{\partial \ln L^*}{\partial \beta_0}\right|$, $\frac{1}{n}\left|\frac{\partial \ln L}{\partial \boldsymbol{\beta}} - \frac{\partial \ln L^*}{\partial \boldsymbol{\beta}}\right|$ and $\frac{1}{n}\left|\frac{\partial \ln L}{\partial \sigma} - \frac{\partial \ln L^*}{\partial \sigma}\right|$ tend to zero as n tends to infinity (see Vaughan and Tiku 2000, p. 57).

Censored case: MML estimators are known to be asymptotically equivalent to ML estimators under some general regularity conditions. A rigorous proof of this is provided in Bhattacharyya (1985) for censored samples (see Bhattacharyya 1985, p. 404). For the model in this paper, it suffices to focus on the following results. For a fixed censoring ratio γ, as $n \to \infty$, we have $n_c = [n\gamma]$(the integer part of $n\gamma$) tends to ∞ and

$$\begin{aligned} \lim_{n_c\to\infty} \sum_{i=1}^{n_c} \frac{h_0\left(t_{i:n}\right)}{n_c} &= E\left[h_0(Z)\right] < \infty, \\ \lim_{n_c\to\infty} \sum_{i=1}^{n_c} \frac{h_0\left(t_{i:n}\right) t_{i:n}}{n_c} &= E\left[Z \cdot h_0(Z)\right] < \infty, \end{aligned}$$

where

$$E\left[h_0(Z)\right] = \frac{b}{2(2a)^{\frac{1}{b}}\left[B\left(a,\frac{1}{b}\right)\right]^2} \times \left\{ \int_0^1 t^{2a+\frac{1}{b}-1}(1-t)^{\frac{1}{b}-1} I^{-1}\left(t;a,\frac{1}{b}\right) \mathrm{d}t + \frac{1+r}{2}\int_0^1 t^{2a+\frac{1}{b}-1}(1-t)^{\frac{1}{b}-1}\left[1-\frac{1+r}{2}I\left(t;a,\frac{1}{b}\right)\right]^{-1}\mathrm{d}t \right\},$$

$$E\left[Z\cdot h_0(Z)\right] = \frac{b}{\left[B\left(a,\frac{1}{b}\right)\right]^2} \times \left\{ -\frac{(1-r)}{2}\int_0^1 t^{2a-1}(1-t)^{\frac{2}{b}-1} I^{-1}\left(t;a,\frac{1}{b}\right)\mathrm{d}t + \frac{(1+r)^2}{4}\int_0^1 \frac{t^{2a-1}(1-t)^{\frac{2}{b}-1}}{1-\frac{1+r}{2}I\left(t;a,\frac{1}{b}\right)}\mathrm{d}t \right\}.$$

The above expectations are all finite, as verified by numerical computation. Thus, we have

$$\lim_{n_c\to\infty} E\left\{\frac{1}{n_c}\sum_{i=1}^{n_c}\left(h_0\left(z_{(i)}\right) - a_{0i} + b_{0i}z_{(i)}\right)\right\} = 0,$$

$$\lim_{n_c\to\infty} E\left\{\frac{1}{n_c}\sum_{i=1}^{n_c}\left(h_0\left(z_{(i)}\right)z_{(i)} - a_{0i}z_{(i)} + b_{0i}z_{(i)}^2\right)\right\} = 0.$$

Following Sect. 6 of Bhattacharyya (1985) will complete the proof [see Tiku and Sürücü (2009), Bhattacharyya (1985)]. □

A.2 Proof of Theorem 2

Proof Asymptotic normality of the MML estimators can be obtained from the basic tools developed in Sect. 2 of Bhattacharyya (1985). The main result of Bhattacharyya (1985) is the utilization of a well-known and widely applied property of order statistics.

First, to apply Theorem 1 of Bhattacharyya (1985) to obtain asymptotic normality results for the ML estimators, we introduce some notation. Let $\mathbf{g}(y_i;\boldsymbol{\theta}) = (a+1/b)\cdot\left(h_1(z_i), \mathbf{x}_i^\top h_1(z_i), z_i h_1(z_i) - \frac{1}{a+1/b}\right)^\top$, $\mathbf{m}(y_i;\boldsymbol{\theta}) = (h_0(z_i), \mathbf{x}_i^\top h_0(z_i), z_i h_0(z_i))^\top$, $z_i = \frac{v_i-\beta_0-\boldsymbol{x}_i^\top\boldsymbol{\beta}}{\sigma}$, $h_0(\cdot), h_1(\cdot)$ are defined in (9), and $f_{y_i}(\cdot)$ denotes the pdf of y_i. For a fixed censoring ratio γ, ζ denote the γ-quantile of the distribution of y_i. To establish the asymptotic normality of the ML estimators, we need to verify the following conditions. Let g_l and m_l denote the lth coordinate of $\mathbf{g}$ and $\mathbf{m}$ respectively, where $l = 1, \cdots, p+2$, then we have: (a) $m_l'(y_i) = \mathrm{d}m_l(y_i)/\mathrm{d}y_i = \frac{1}{\sigma}\mathrm{d}m_l(y_i)/\mathrm{d}z_i$ exists at $y_i = \zeta$, (b) $g_l(y_i)$ is continuous at ζ, and (c) $\int_{-\infty}^{\zeta} g_l^2(y_i) f_{y_i}(y_i)\mathrm{d}y_i < \infty$. This last follows from the fact that $E\left[h_1^2(Z)\right] = \frac{b^2 B(a+2/b,2-1/b)}{(2a)^{2/b}B(a,1/b)(1-r^2)} < \infty$ and $E\left[Z^2 h_1(Z)\right] = \frac{b^2 B(a,2+1/b)}{B(a,1/b)} < \infty$, where $Z \sim TPGT(0,1,r,a,b)$. To obtain explicit expressions for the mean and

covariance, using these results for $y_i = \zeta$, the expressions (2.3) in Bhattacharyya (1985) for the present case simplify to

$$\boldsymbol{\mu} = \mathbf{0},$$
$$\mathbf{J}(\boldsymbol{\theta}) = -\left\{\int_{\zeta}^{+\infty} \frac{\partial \mathbf{g}(y_i;\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} f_{y_i}(y_i)\mathrm{d}y_i + \gamma \frac{\partial \mathbf{m}(y_i;\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\right\}.$$

It follows that $\mathbf{J}(\boldsymbol{\theta}) = \frac{I(\boldsymbol{\theta})}{n}$, if the censoring ratio $\gamma = 0$, where $I(\boldsymbol{\theta})$ is defined in Sect. 3.5.

Then, by using results from Section 6 of Bhattacharyya (1985), we can establish the asymptotic normality result of the MML estimators for the TPGT-CR model as follows

$$\sqrt{n}\left(\widehat{\boldsymbol{\theta}}_{MML} - \boldsymbol{\theta}\right) \to N_{p+2}\left(\mathbf{0}, \mathbf{J}^{-1}(\boldsymbol{\theta})\right).$$

□

Appendix B The Moments of Order Statistics

Since t_i values have an important effect on the implementation of MML estimation. In this subsection, we derive the explicit expression for the expected values of the standardized order statistics for TPGT distribution under the independent identically-distributed case.

The generalized Kampe de Feriet function from Exton (1978) is defined by

$$F_{C:D}^{A:B}\left((a):(b_1);\cdots;(b_n);(c):(d_1);\cdots;(d_n);x_1,\cdots,x_n\right)$$
$$= \sum_{m_1=0}^{\infty}\cdots\sum_{m_n=0}^{\infty} \frac{((a))_{m_1+\cdots+m_n}\left((b_1)\right)_{m_1}\cdots\left((b_n)\right)_{m_n}}{((c))_{m_1+\cdots+m_n}\left((d_1)\right)_{m_1}\cdots\left((d_n)\right)_{m_n}} \times \frac{x_1^{m_1}\cdots x_n^{m_n}}{m_1!\cdots m_n!},$$

where $a = (a_1, a_2, \cdots, a_A), b_i = \left(b_{i,1}, b_{i,2}, \cdots, b_{i,B}\right)$, $c = (c_1, \cdots, c_C)$, $d_i = \left(d_{i,1}, \cdots, d_{i,D}\right)$ for $i = 1, 2, \cdots, n$ and $((f))_k = \left(\left(f_1, f_2, \cdots, f_p\right)\right)_k = (f_1)_k \cdots \left(f_p\right)_k, (f_i)_k = f_i\left(f_i+1\right)\cdots\left(f_i+k-1\right)$. By using the above special function, we derive the following theorem:

Theorem 3 *The kth moment of order statistics $X_{i:n}$ from $TPGT(0, 1, r, a, b)$ can be calculated by the following convergent expression*

$$E\left(X_{i:n}^k\right) = C_{i,n}\left[A(k,i,n) + B(k,i,n)\right],$$

where $C_{i:n} = \frac{n!}{(i-1)!(n-i)!}$,

$$A(k,i,n) = \frac{(2a)^{\frac{k}{b}}(1+r)^{k+1}}{2\mathrm{B}\left(a,\frac{1}{b}\right)} \sum_{j=0}^{i-1}\binom{i-1}{j}(-1)^j\left(\frac{1+r}{2}\right)^{n-i+j} A_1(j), \quad \text{(B2)}$$

$$B(k,i,n) = \frac{(-1)^k (2a)^{\frac{k}{b}} (1-r)^{k+1}}{2\mathrm{B}\left(a, \frac{1}{b}\right)} \sum_{j=0}^{n-i} \binom{n-i}{j} (-1)^j \left(\frac{1-r}{2}\right)^{i-1+j} B_1(j), \tag{B3}$$

and

$$A_1(j) = \frac{\mathrm{B}\left((n-i+j+1)a - \frac{k}{b}, \frac{k+1}{b}\right)}{\left(\mathrm{B}\left(a, \frac{1}{b}\right) a\right)^{n-i+j}} \times F_{1:1}^{1:2}\left((n-i+j+1)a - \frac{k}{b} : \left(1-\frac{1}{b}, a\right); \right.$$
$$\left. \cdots ; \left(1-\frac{1}{b}, a\right); (n-i+j+1)a + \frac{1}{b} : (a+1); \cdots ; (a+1); 1, \cdots, 1\right), \tag{B4}$$

$$B_1(j) = \frac{\mathrm{B}\left((i+j)a - \frac{k}{b}, \frac{k+1}{b}\right)}{\left(\mathrm{B}\left(a, \frac{1}{b}\right) a\right)^{i+j-1}} \times F_{1:1}^{1:2}\left((i+j)a - \frac{k}{b} : \left(1-\frac{1}{b}, a\right); \right.$$
$$\left. \cdots ; \left(1-\frac{1}{b}, a\right); (i+j)a + \frac{1}{b} : (a+1); \cdots ; (a+1); 1, \cdots, 1\right). \tag{B5}$$

Proof From the formula in Eq. (1), we readily have the integration form for $E\left(X_{i:n}^k\right)$ as

$$\begin{aligned} E\left(X_{i:n}^k\right) &= C_{i,n}\left\{ \int_0^{+\infty} x^k [F_0(x)]^{i-1} [1-F_0(x)]^{n-i} f_0(x) \mathrm{d}x \right. \\ &\quad \left. + \int_{-\infty}^{0} x^k [F_0(x)]^{i-1} [1-F_0(x)]^{n-i} f_0(x) \mathrm{d}x \right\} \\ &\triangleq C_{i,n} \left(A(k,i,n) + B(k,i,n)\right), \end{aligned}$$

where

$$\begin{aligned} A(k,i,n) &= \int_0^{\infty} x^k \left[1 - \frac{1+r}{2} I_{u(x)}(a, 1/b)\right]^{i-1} \left[\frac{1+r}{2} I_{u(x)}(a, 1/b)\right]^{n-i} f_0(x) \mathrm{d}x \\ &= \frac{(2a)^{\frac{k}{b}} (1+r)^{k+1}}{2\mathrm{B}\left(a, \frac{1}{b}\right)} \int_0^1 \left[1 - \frac{1+r}{2} I_u(a, 1/b)\right]^{i-1} \left[\frac{1+r}{2} I_u(a, 1/b)\right]^{n-i} \\ &\quad \times (1-u)^{\frac{k+1}{b}-1} u^{a-\frac{k}{b}-1} \mathrm{d}u \\ &= \frac{(2a)^{\frac{k}{b}} (1+r)^{k+1}}{2\mathrm{B}\left(a, \frac{1}{b}\right)} \sum_{j=0}^{i-1} \binom{i-1}{j} (-1)^j \left(\frac{1+r}{2}\right)^{n-i+j} A_1(j). \end{aligned}$$

Now, by using the known generalized multinomial theorem, we obtain

$$I_u\left(a, \frac{1}{b}\right) = \frac{u^a}{\mathrm{B}\left(a, \frac{1}{b}\right)} \sum_{k=0}^{\infty} \frac{(1-\frac{1}{b})_k u^k}{(a+k)\, k!},$$

and

$$\left[I_u\left(a,\frac{1}{b}\right)\right]^{n_0}=\left[\frac{u^a}{\left[\mathrm{B}\left(a,\frac{1}{b}\right)\right]}\right]^{n_0}\sum_{m_1+\cdots+m_{n_0}=0}^{\infty}\frac{(1-\frac{1}{b})_{m_1}\cdots(1-\frac{1}{b})_{m_{n_0}}}{(a+m_1)\cdots\left(a+m_{n_0}\right)}\times\frac{u^{m_1+\cdots+m_{n_0}}}{m_1!\cdots m_{n_0}!},\tag{B6}$$

where $n_0=n-i+j$, the relation in (B6) can be used to produce expressions in Eq. (B4) for real non-integer values of a. We can proceed similarly to derive explicit expressions for $B(k,i,n)$ and $B_1(j)$. □

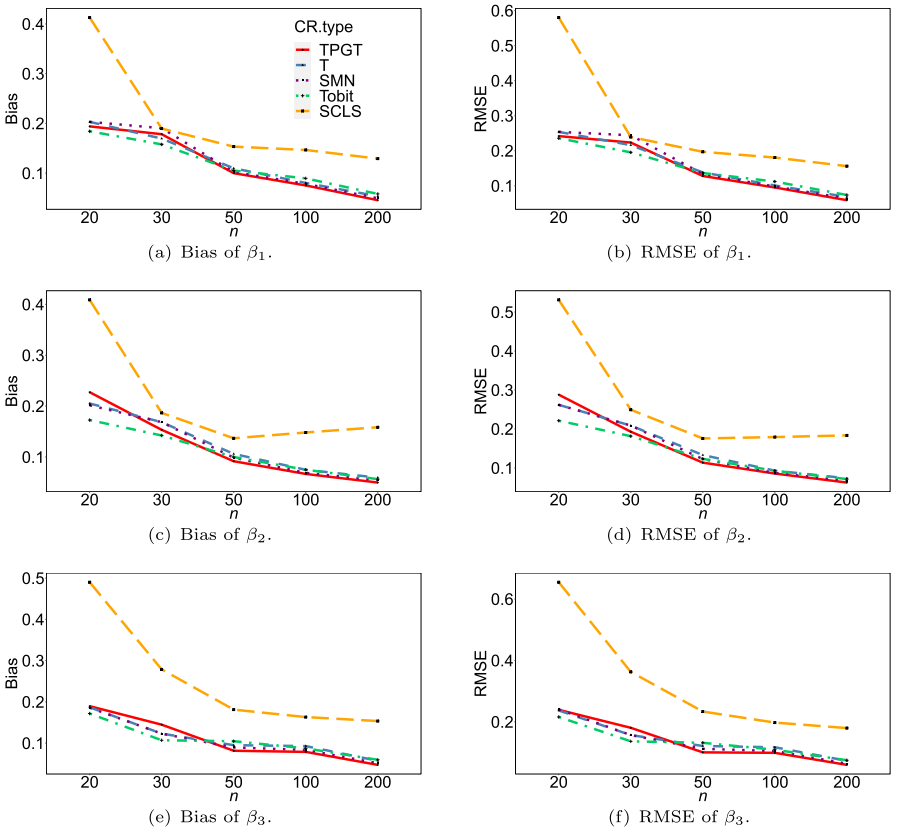
Appendix C Simulation results

C.1 Further plots of the first simulation

In this section, we analyze simulation results under censoring rates of 20% and 30%. The results are presented in Figs. 10, 11, 12 and 13 for a 20% censoring rate and Figs. 14, 15, 16 and 17 for a 30% censoring rate. According to the results in Figs. 10, 11 and 14, 15, the relative performance of the five methods agrees with that at a 10% censoring rate for symmetric and lightly skewed cases. In terms of Bias and RMSE, Fig. 10 shows that the Tobit estimator performs best for $n=20,30$ when the error distribution is normal. The TPGT-CR model is best for $n=50,100,200$, while SMN-CR, T-CR, and Tobit models are comparable. Again, the SCLS estimators exhibit the largest biases and RMSEs across all sample sizes. Figure 11 shows that TPGT-CR exhibits a greater relative advantage in Design II compared to Design I.

Figures 12 and 13 summarize the simulation results obtained for designs III and IV at a 20% censoring rate. For moderately skewed and heavy-tailed distributions, the TPGT-CR is generally associated with better results than SMN-CR and T-CR in terms of Bias and RMSE. It can be seen that the advantage of TPGT-CR becomes more pronounced for the censoring level of 30%. From Fig. 12, it is shown that as the sample size n increases from 50 to 200, the bias and RMSE of Tobit and SCLS also increase, indicating a lack of consistency in the Tobit and SCLS estimators for skewed distribution. From Fig. 13, the TPGT-CR is still the best, followed by SMN-CR model when error distribution has strong asymmetry and heavy tails. Because the MML method relies on order statistics rather than original data and all other estimators directly or indirectly employ trimming of observations.

For highly skewed and heavy-tailed distributions, Tobit and SCLS exhibit the worst performance for $n=20,30,200$. This indicates that the symmetry and tail behavior of the data significantly affect the bias and RMSE of SCLS (Powell 1986). Figure 17 shows that only for $n=50,100$, the performance of Tobit is comparable to TPGT-CR and SMN-CR. Furthermore, when comparing Figs. 16 and 17, we can observe that as the degree of skewness increases, the performance of T-CR gradually worsens. This result is reasonable because the T-CR method is better suited for modeling symmetric



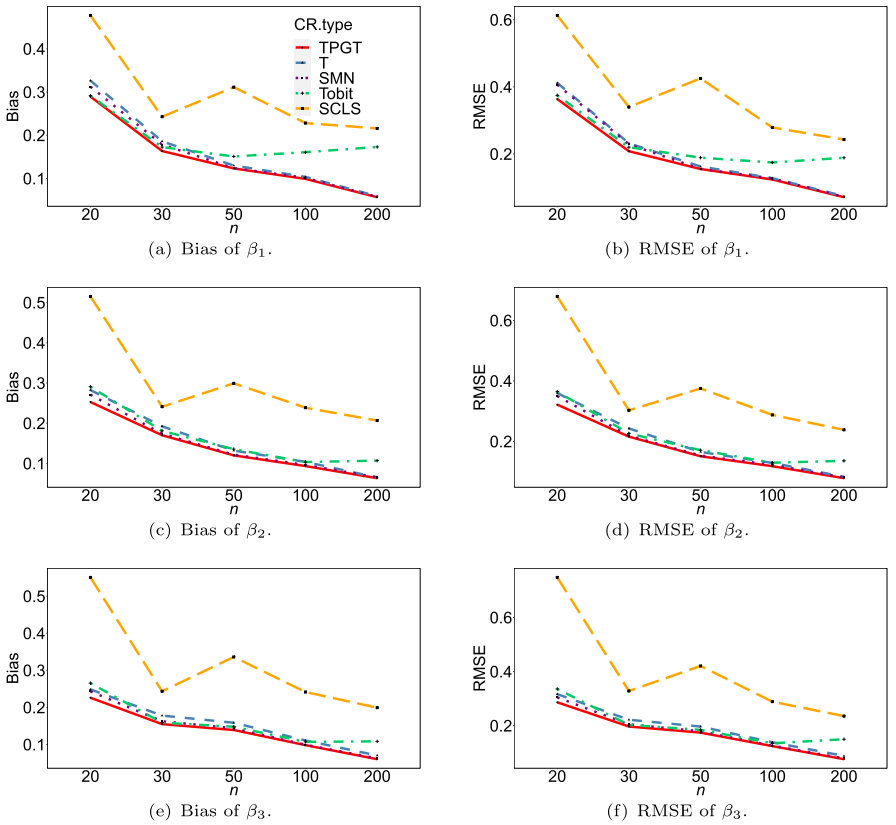
(a) Bias of β_1. (b) RMSE of β_1.

(c) Bias of β_2. (d) RMSE of β_2.

(e) Bias of β_3. (f) RMSE of β_3.

Fig. 10 Design I: Bias and RMSE of different estimators of slope coefficient (true $\beta_i = 1,\ i = 1, 2, 3$) in 1000 random trials for TPGT-CR model with $n = 20, 30, 50, 100, 200$ observations when the level of censoring is 20%

data. Finally, Figs. 13 and 17 demonstrate the superiority of the TPGT-CR when the error distribution is highly skewed, even under high censoring levels.

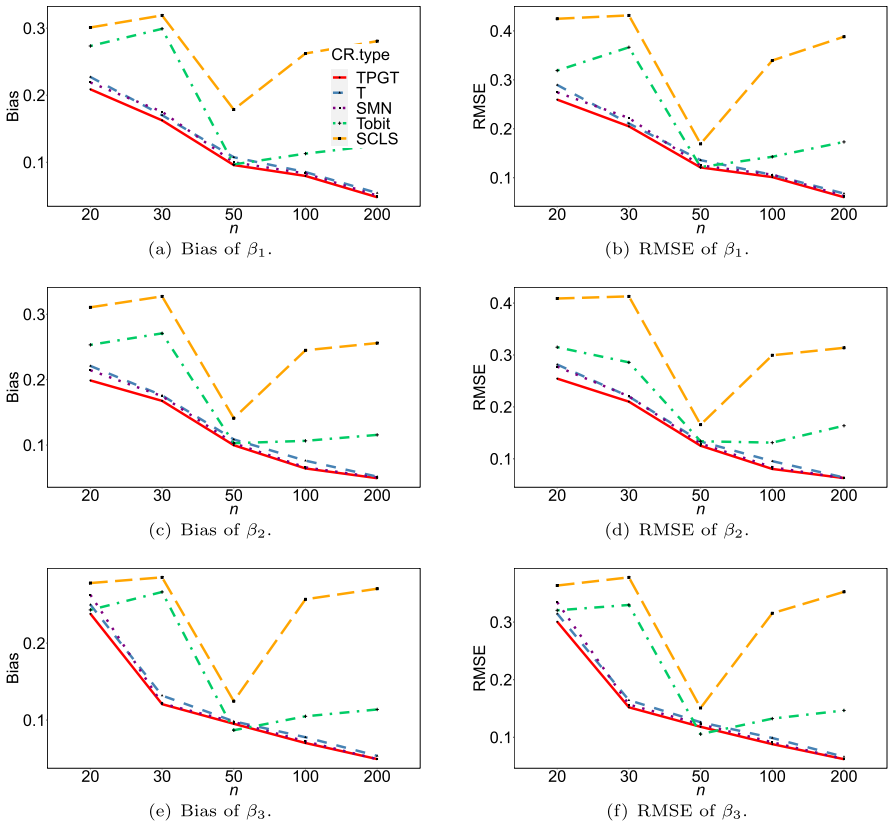


(a) Bias of β_1. (b) RMSE of β_1.

(c) Bias of β_2. (d) RMSE of β_2.

(e) Bias of β_3. (f) RMSE of β_3.

Fig. 11 Design II: Bias and RMSE of different estimators of slope coefficient (true $\beta_i = 1,\ i = 1, 2, 3$) in 1000 random trials for TPGT-CR model with $n = 20, 30, 50, 100, 200$ observations when the level of censoring is 20%

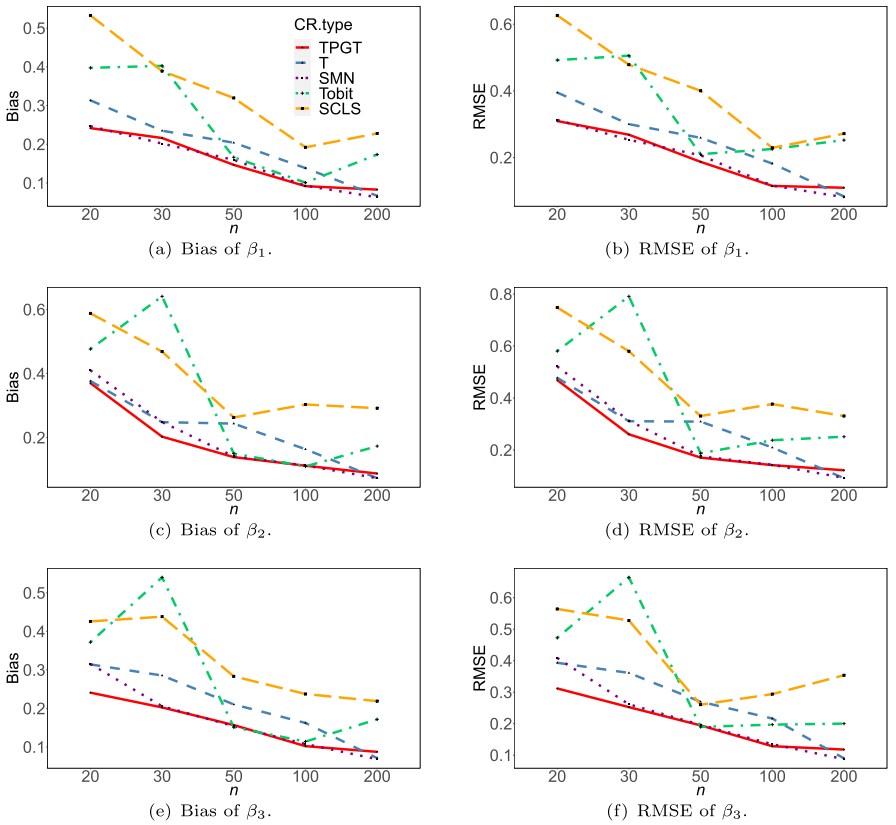


(a) Bias of β_1. (b) RMSE of β_1.

(c) Bias of β_2. (d) RMSE of β_2.

(e) Bias of β_3. (f) RMSE of β_3.

Fig. 12 Design III: Bias and RMSE of different estimators of slope coefficient (true $\beta_i = 1,\ i = 1, 2, 3$) in 1000 random trials for TPGT-CR model with $n = 20, 30, 50, 100, 200$ observations when the level of censoring is 20%

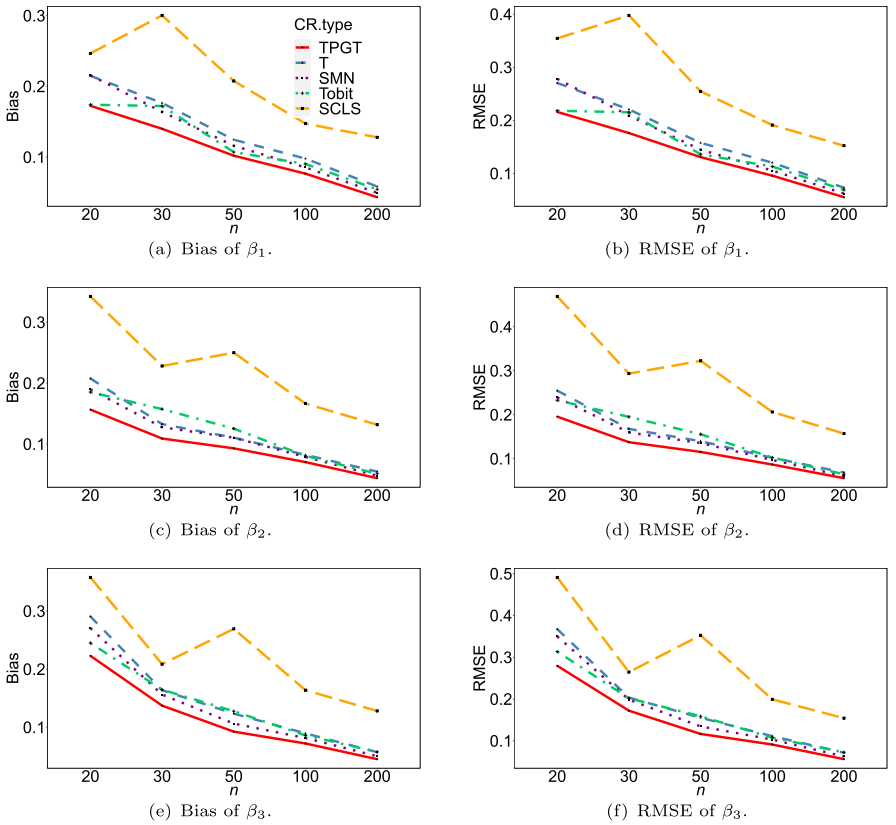


(a) Bias of β_1. (b) RMSE of β_1.

(c) Bias of β_2. (d) RMSE of β_2.

(e) Bias of β_3. (f) RMSE of β_3.

Fig. 13 Design IV: Bias and RMSE of different estimators of slope coefficient (true $\beta_i = 1,\ i = 1, 2, 3$) in 1000 random trials for TPGT-CR model with $n = 20, 30, 50, 100, 200$ observations when the level of censoring is 20%

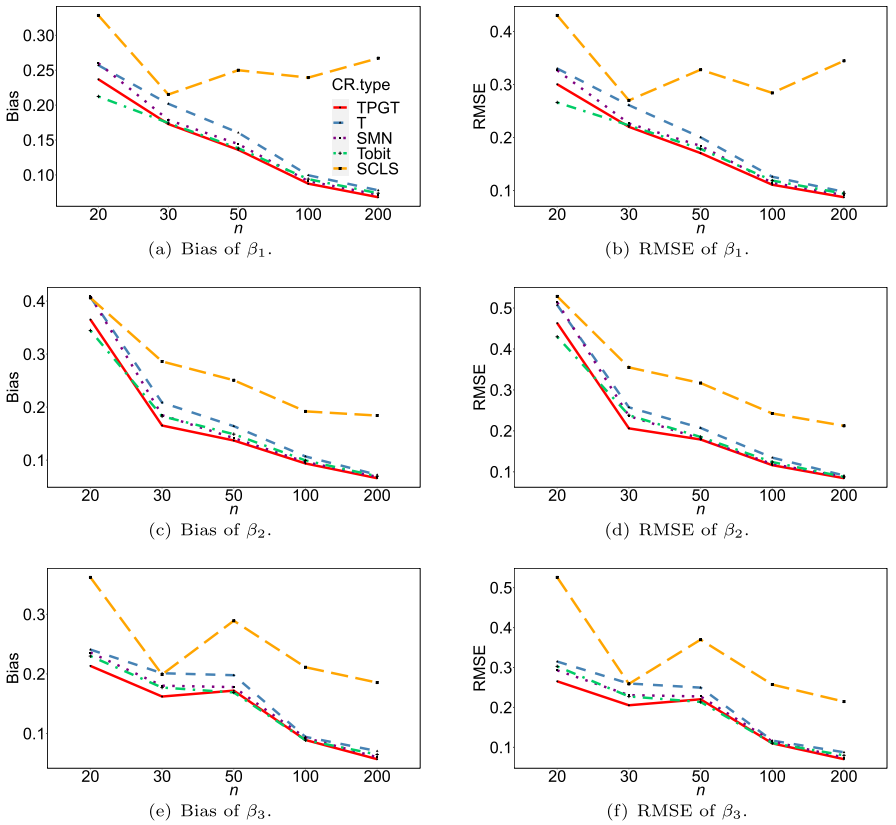


(a) Bias of β_1. (b) RMSE of β_1.

(c) Bias of β_2. (d) RMSE of β_2.

(e) Bias of β_3. (f) RMSE of β_3.

Fig. 14 Design I: Bias and RMSE of different estimators of slope coefficient (true $\beta_i = 1,\ i = 1, 2, 3$) in 1000 random trials for TPGT-CR model with $n = 20, 30, 50, 100, 200$ observations when the level of censoring is 30%

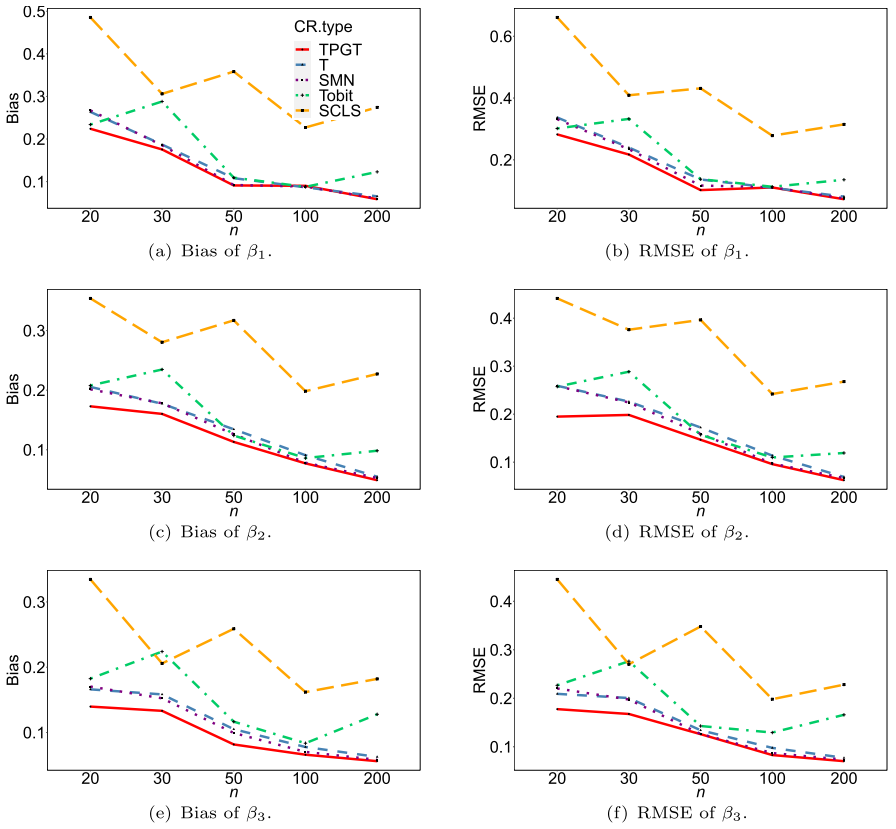


(a) Bias of β_1. (b) RMSE of β_1.

(c) Bias of β_2. (d) RMSE of β_2.

(e) Bias of β_3. (f) RMSE of β_3.

Fig. 15 Design II: Bias and RMSE of different estimators of slope coefficient (true $\beta_i = 1,\ i = 1, 2, 3$) in 1000 random trials for TPGT-CR model with $n = 20, 30, 50, 100, 200$ observations when the level of censoring is 30%

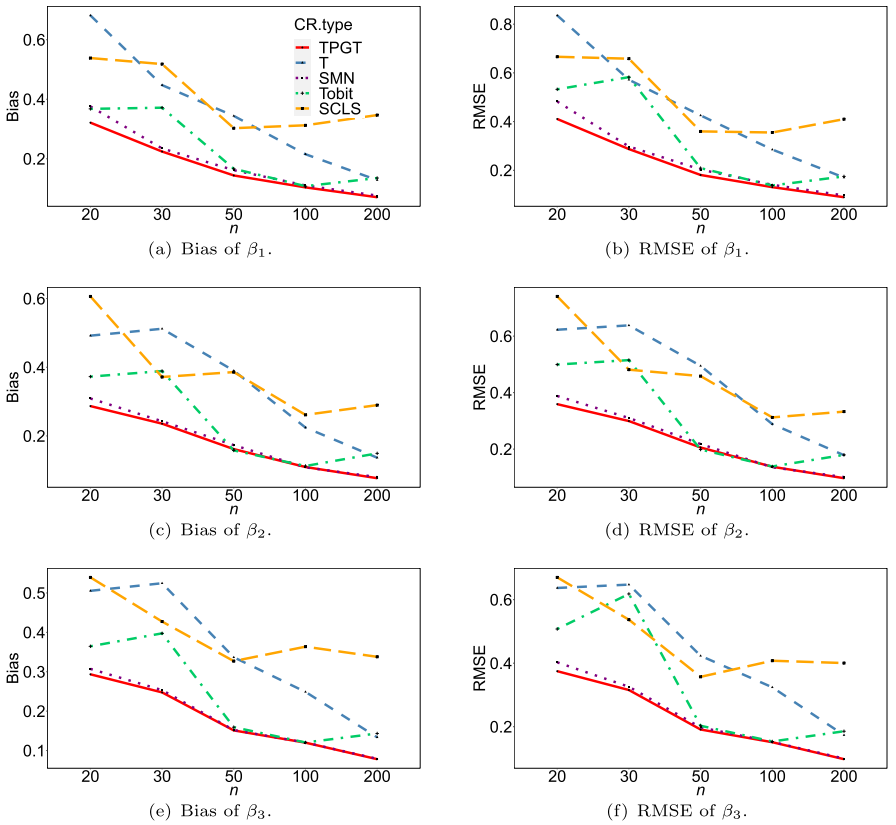


(a) Bias of β_1. (b) RMSE of β_1.

(c) Bias of β_2. (d) RMSE of β_2.

(e) Bias of β_3. (f) RMSE of β_3.

Fig. 16 Design *III*: Bias and RMSE of different estimators of slope coefficient (true $\beta_i = 1,\ i = 1, 2, 3$) in 1000 random trials for TPGT-CR model with $n = 20, 30, 50, 100, 200$ observations when the level of censoring is 30%



(a) Bias of β_1. (b) RMSE of β_1.

(c) Bias of β_2. (d) RMSE of β_2.

(e) Bias of β_3. (f) RMSE of β_3.

Fig. 17 Design IV: Bias and RMSE of different estimators of slope coefficient (true $\beta_i = 1,\ i = 1, 2, 3$) in 1000 random trials for TPGT-CR model with $n = 20, 30, 50, 100, 200$ observations when the level of censoring is 30%

C.2 Further plot of the third simulation

See 18.

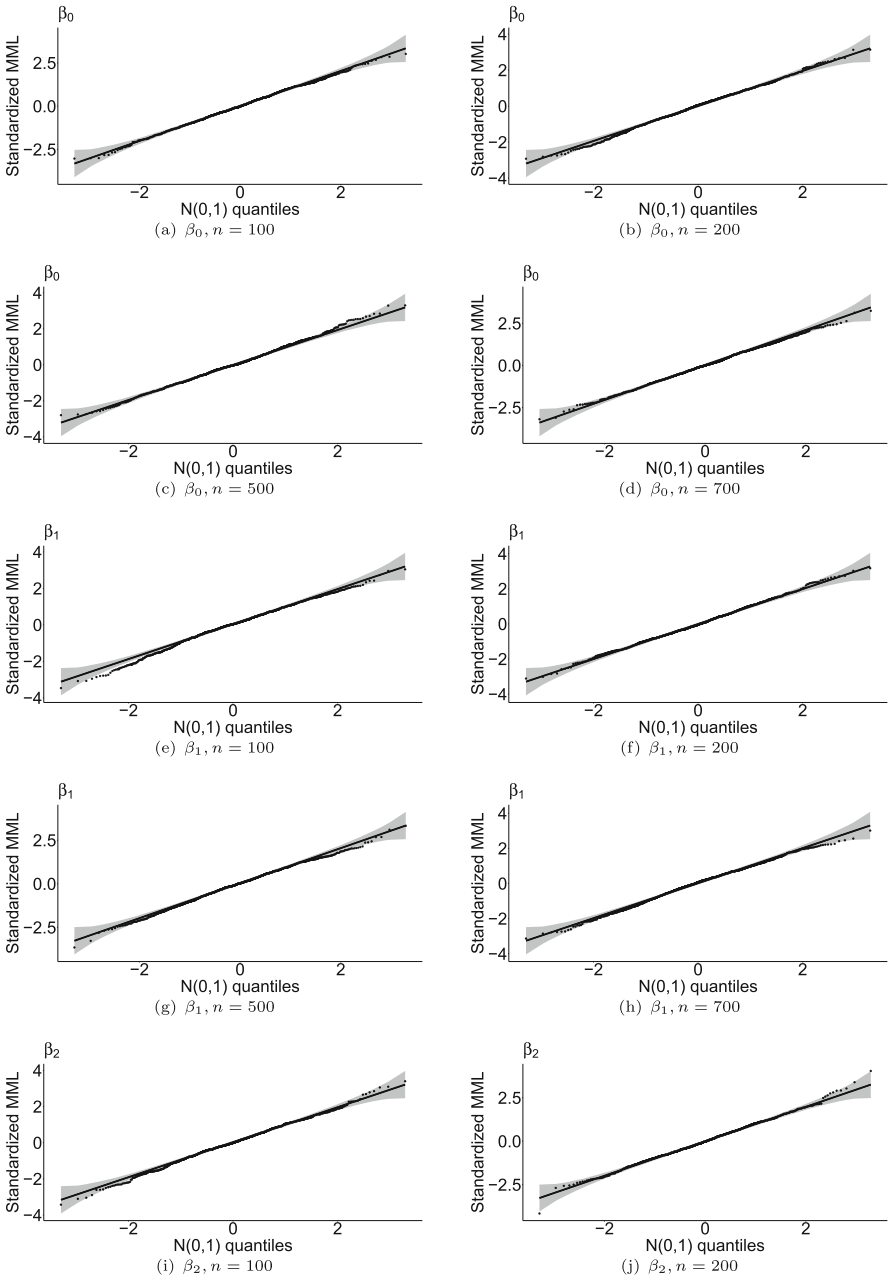
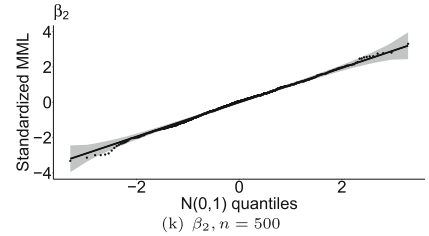
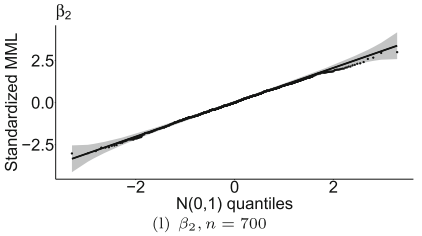


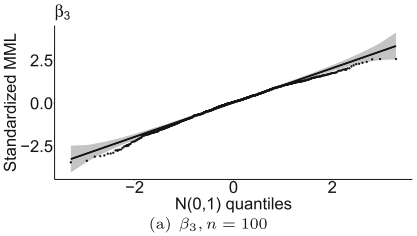
Fig. 18 QQ-plots for normal distribution based on 1000 MML estimates. Sample sizes $n = 100, 200, 500, 700$, and 10% censoring levels



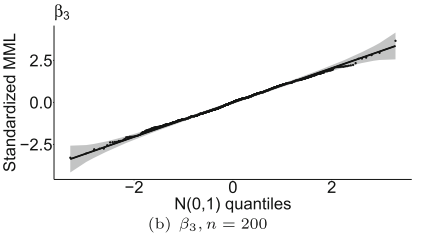
(k) $\beta_2, n = 500$



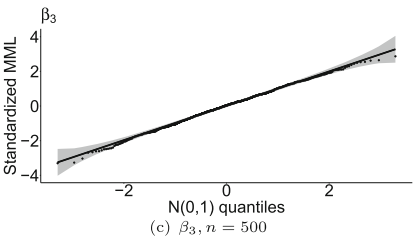
(l) $\beta_2, n = 700$



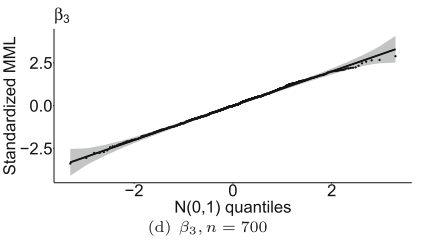
(a) $\beta_3, n = 100$



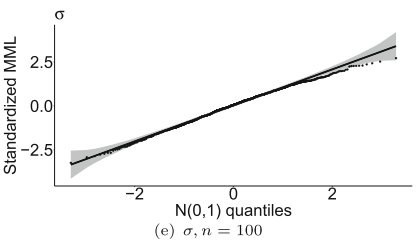
(b) $\beta_3, n = 200$



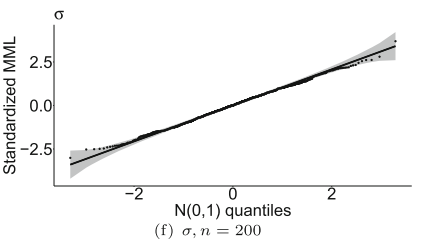
(c) $\beta_3, n = 500$



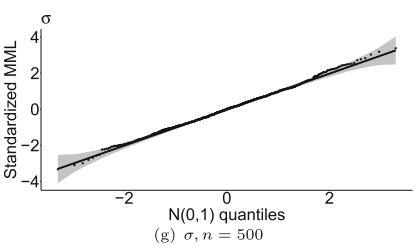
(d) $\beta_3, n = 700$



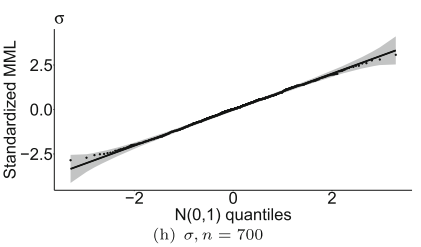
(e) $\sigma, n = 100$



(f) $\sigma, n = 200$



(g) $\sigma, n = 500$

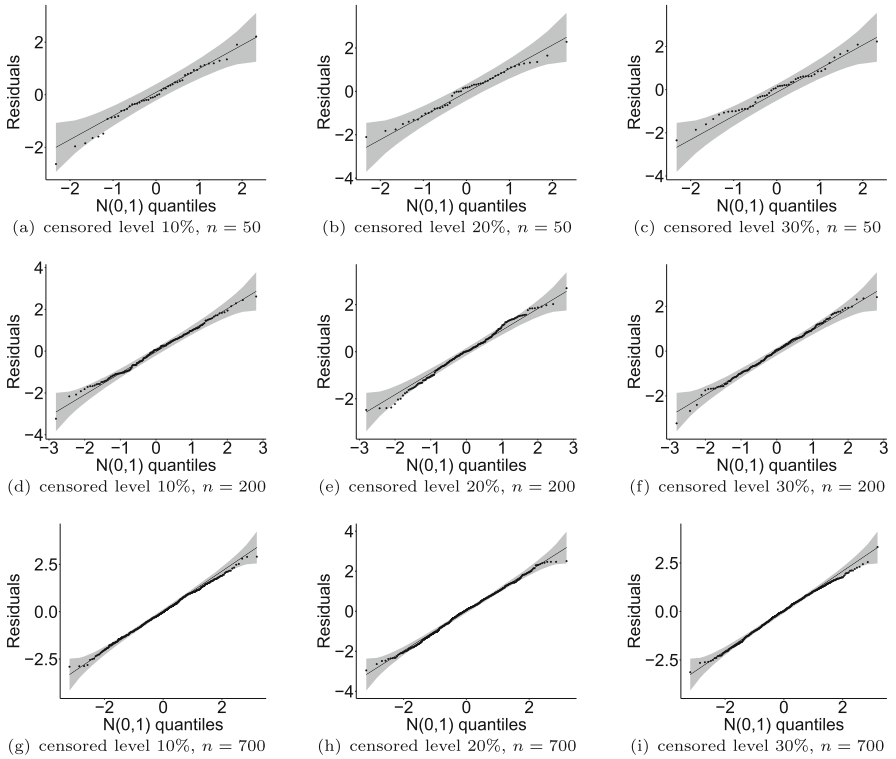


(h) $\sigma, n = 700$

Fig. 18 continued

C.3 Further plot of the fourth simulation

See 19.



(a) censored level 10%, $n = 50$ (b) censored level 20%, $n = 50$ (c) censored level 30%, $n = 50$

(d) censored level 10%, $n = 200$ (e) censored level 20%, $n = 200$ (f) censored level 30%, $n = 200$

(g) censored level 10%, $n = 700$ (h) censored level 20%, $n = 700$ (i) censored level 30%, $n = 700$

Fig. 19 QQ-plots of the modified deviance residual r_{D_i}. Sample sizes $n = 50, 200, 700$, and levels of censoring = 10%, 20% and 30%

Author contributions All authors certify that they have no affiliations with or involvement in any organization or entity with any financial interest or non-financial interest in the subject matter or materials discussed in this manuscript. The authors have no financial or proprietary interests in any material discussed in this article.

Funding No funding was received for conducting this study.

Declarations

Conflict of interest The authors have no conflict of interest to declare that are relevant to the content of this article.

References

Acitas S, Kasap P, Senoglu B et al (2013) One-step m-estimators: Jones and Faddy's skewed t-distribution. J Appl Stat 40(7):1545–1560

Acitas S, Filzmoser P, Senoglu B (2021) A robust adaptive modified maximum likelihood estimator for the linear regression model. J Stat Comput Simul 91(7):1394–1414

Amemiya T (1985) Adv econometrics. Harvard University Press, Cambridge

Arellano-Valle RB, Gómez HW, Quintana FA (2005) Statistical inference for a general class of asymmetric distributions. J Stat Plan Inference 128(2):427–443

Arellano-Valle RB, Castro LM, González-Farías G et al (2012) Student-t censored regression model: properties and inference. Stat Methods Appl 21(4):453–473

Arellano-Valle RB, Azzalini A, Ferreira CS et al (2020) A two-piece normal measurement error model. Comput Stat Data Anal 144:106863

Arslan MT, Senoglu B (2018) Type ii censored samples in experimental design under jones and Faddy's skew t distribution. Iran J Sci Technol Trans A 42:2145–2157

Arslan O, Genc AI (2009) The skew generalized t distribution as the scale mixture of a skew exponential power distribution and its applications in robust estimation. Statistics 43(5):481–498

Azzalini A, Capitanio A (2003) Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t-distribution. J R Stat Soc Ser B 65(2):367–389

Balci S, Akkaya AD, Ulgen BE (2013) Modified maximum likelihood estimators using ranked set sampling. J Comput Appl Math 238:171–179

Bhattacharyya GK (1985) The asymptotics of maximum likelihood and related estimators based on type ii censored data. J Am Stat Assoc 80(390):398–404

Cameron AC, Trivedi PK (2005) Microeconometrics: methods and applications. Cambridge University Press, Cambridge

Carrasco JM, Ortega EM, Paula GA (2008) Log-modified weibull regression models with censored data: sensitivity and residual analysis. Comput Stat Data Anal 52(8):4021–4039

Caudill SB (2012) A partially adaptive estimator for the censored regression model based on a mixture of normal distributions. Stat Methods Appl 21:121–137

Chen S, Khan S (2000) Estimating censored regression models in the presence of nonparametric multiplicative heteroskedasticity. J Econometr 98(2):283–316

Chen Y, Oliver DS (2012) Ensemble randomized maximum likelihood method as an iterative ensemble smoother. Math Geosci 44:1–26

Collett D (2003) Modelling survival data in medical research. Chapman & Hall/CRC, Boca Raton.

DaVanzo J, Lee D (1978) The compatibility of child care with labor force participation and non-market activities. Rand Corporation, Santa Monica

Exton H (1978) Handbook of hypergeometric integrals: theory, applications, tables, computer programs. Halsted Press, New York

Fernández C, Steel MF (1999) Multivariate student-t regression models: pitfalls and inference. Biometrika 86(1):153–167

Garay AM, Bolfarine H, Lachos VH et al (2015) Bayesian analysis of censored linear regression models with scale mixtures of normal distributions. J Appl Stat 42(12):2694–2714

Garay AM, Lachos VH, Bolfarine H et al (2017) Linear censored regression models with scale mixtures of normal distributions. Stat Pap 58:247–278

Gómez G, Calle ML, Oller R et al (2009) Tutorial on methods for interval-censored data and their implementation in r. Stat Model 9(4):259–297

Guzmán DC, Ferreira CS, Zeller CB (2021) Linear censored regression models with skew scale mixtures of normal distributions. J Appl Stat 48(16):3060–3085

Hoeffding W (1953) On the distribution of the expected values of the order statistics. Ann Math Stat 24(1):93–100

Karlsson M, Laitila T (2014) Finite mixture modeling of censored regression models. Stat Pap 55:627–642

Khan S, Powell JL (2001) Two-step estimation of semiparametric censored regression models. J Econometr 103(1–2):73–110

Khan S, Tamer E (2009) Inference on endogenously censored regression models using conditional moment inequalities. J Econometr 152(2):104–119

Lee SY, Zhu HT (2002) Maximum likelihood estimation of nonlinear structural equation models. Psychometrika 67(2):189–210

Lewis RA, McDonald JB (2014) Partially adaptive estimation of the censored regression model. Econometr Rev 33(7):732–750
Li R, Peng L (2017) Assessing quantile prediction with censored quantile regression models. Biometrics 73(2):517–528
Lian C, Rong Y, Cheng W (2024) On a novel skewed generalized t distribution: Properties, estimations and its applications. Publishing Taylor & Francis Online https://www.tandfonline.com/eprint/9HNTXT46S9THSQEYBEN3/full?target=10.1080/03610926.2024.2313034. Accessed 21 Feb 2024
Lin TI (2010) Robust mixture modeling using ate skew t distributions. Stat Comput 20:343–356
Lucas A (1997) Robustness of the student t based m-estimator. Commun Stat—Theory Methods 26(5):1165–1182
Massuia MB, Garay AM, Cabral CR et al (2017) Bayesian analysis of censored linear regression models with scale mixtures of skew-normal distributions. Stat Interface 10(3):425–439
Mattos TdB, Garay AM, Lachos VH (2018) Likelihood-based inference for censored linear regression models with scale mixtures of skew-normal distributions. J Appl Stat 45(11):2039–2066
Mirfarah E, Naderi M, Chen DG (2021) Mixture of linear experts model for censored data: a novel approach with scale-mixture of normal distributions. Comput Stat Data Anal 158:107182
Mroz TA (1987) The sensitivity of an empirical model of married women's hours of work to economic and statistical assumptions. Econometrica 55:765–799
Mudholkar GS, Hutson AD (2000) The epsilon-skew-normal distribution for analyzing near-normal data. J Stat Plan Inference 83(2):291–309
Ortega EM, Bolfarine H, Paula GA (2003) Influence diagnostics in generalized log-gamma regression models. Comput Stat Data Anal 42(1–2):165–186
Pescim RR, Ortega EM, Cordeiro GM et al (2017) A new log-location regression model: estimation, influence diagnostics and residual analysis. J Appl Stat 44(2):233–252
Powell JL (1984) Least absolute deviations estimation for the censored regression model. J Econometr 25(3):303–325
Powell JL (1986) Symmetrically trimmed least squares estimation for tobit models. Econometrica: J Econometr Soc 54:1435–1460
Salah K, Yousri S (2019) Nonparametric relative regression under random censorship model. Stat Probab Lett 151:116–122
Tiku M (1967) Estimating the mean and standard deviation from a censored normal sample. Biometrika 54(1–2):155–165
Tiku M, Suresh R (1992) A new method of estimation for location and scale parameters. J Stat Plan Inference 30(2):281–292
Tiku ML, Sürücü B (2009) Mmles are as good as m-estimators or better. Stat Probab Lett 79(7):984–989
Tobin J (1958) Estimation of relationships for limited dependent variables. Econometrica: J Econometr Soc 26:24–36
Vaughan D, Tiku M (2000) Estimation and hypothesis testing for a nonnormal bivariate distribution with applications. Math Comput Model 32(1–2):53–67
Wang L, McMahan CS, Hudgens MG et al (2016) A flexible, computationally efficient method for fitting the proportional hazards model to interval-censored data. Biometrics 72(1):222–231
Wang WL (2023) Multivariate contaminated normal censored regression model: properties and maximum likelihood inference. J Comput Graphical Stat 32(4):1671–1684
Yalçınkaya A, Şenoğlu B, Yolcu U (2018) Maximum likelihood estimation for the parameters of skew normal distribution using genetic algorithm. Swarm Evolut Comput 38:127–138
Zeller CB, Cabral CRB, Lachos VH et al (2019) Finite mixture of regression models for censored data based on scale mixtures of normal distributions. Adv Data Anal Classif 13:89–116
Zeng D, Gao F, Lin D (2017) Maximum likelihood estimation for semiparametric regression models with multivariate interval-censored data. Biometrika 104(3):505–525

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.