

Article

Parametric Quantile Regression Models for Fitting Double Bounded Response with Application to COVID-19 Mortality Rate Data

Diego I. Gallardo ¹, Marcelo Bourguignon ², Yolanda M. Gómez ¹, Christian Caamaño-Carrillo ³
and Osvaldo Venegas ^{4,*}

¹ Department of Mathematics, Faculty of Engineering, University of Atacama, Copiapó 1530000, Chile; diego.gallardo@uda.cl (D.I.G.); yolanda.gomez@uda.cl (Y.M.G.)

² Department of Statistics, Federal University of Rio Grande do Norte, Natal 59078-970, Brazil; marcelo.bourguignon@ufrn.br

³ Department of Statistics, Faculty of Science, University of Bío-Bío, Concepción 4081112, Chile; chcaaman@ubiobio.cl

⁴ Departamento de Ciencias Matemáticas y Físicas, Facultad de Ingeniería, Universidad Católica de Temuco, Temuco 4780000, Chile

* Correspondence: ovenegas@uct.cl; Tel.: +56-45-220-5351

Abstract: In this paper, we develop two fully parametric quantile regression models, based on the power Johnson S_B distribution for modeling unit interval response in different quantiles. In particular, the conditional distribution is modeled by the power Johnson S_B distribution. The maximum likelihood (ML) estimation method is employed to estimate the model parameters. Simulation studies are conducted to evaluate the performance of the ML estimators in finite samples. Furthermore, we discuss influence diagnostic tools and residuals. The effectiveness of our proposals is illustrated with a data set of the mortality rate of COVID-19 in different countries. The results of our models with this data set show the potential of using the new methodology. Thus, we conclude that the results are favorable to the use of proposed quantile regression models for fitting double bounded data.

Keywords: COVID-19; parametric quantile regression; power Johnson S_B distribution; proportion

MSC: 62F10; 62F35



Citation: Gallardo, D.I.; Bourguignon, M.; Gómez Y.M.; Caamaño-Carrillo, C.; Venegas, O. Parametric Quantile Regression Models for Fitting Double Bounded Response with Application to COVID-19 Mortality Rate Data. *Mathematics* **2022**, *10*, 2249. <https://doi.org/10.3390/math10132249>

Academic Editor: Maurizio Brizzi

Received: 13 May 2022

Accepted: 17 June 2022

Published: 27 June 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The most commonly employed two-parameter distribution for modeling double bounded random variables in the unit interval is the beta distribution. In order to accommodate explanatory variables in the modeling, Ferrari and Cribari-Neto [1] introduced the beta regression model based on a parameterization of the beta distribution in terms of the mean and precision parameters.

Several researchers have focused on the use of the mean reparameterized beta distribution as an integral of the model, for example, see [2–4]. However, there are many limitations of the conditional mean models. For example, in an asymmetric distribution, or in the presence of outliers, point estimates of the population mean are typically pulled in the direction of one of the distribution tails.

Quantile regression models, introduced by Koenker and Bassett [5], are an approach for understanding the conditional distribution of a response variable given the values of some covariates at different quantiles, thus providing users with a more complete picture. In particular, several authors [6,7] have highlighted the robustness to outliers connected with quantile regression models. Furthermore, if the conditional dependent variable is skewed, quantiles may be more appropriate when compared with the mean [8].

However, parametric quantile regression models for limited range response variables have not received much attention in the literature. Lemonte and Bazán [9] introduced a new class of distributions named the generalized Johnson S_B with bounded support on the basis of the symmetric family of distributions. In particular, Lemonte and Bazán [9] provided the median reparameterization of the Johnson S_B distribution [10], which facilitates its use in a regression setting. Unlike the beta regression, the median in the reparameterized Johnson S_B distribution is related to a linear predictor. Cancho et al. [11] generalized the Johnson S_B model to a general class of distributions. The authors introduced an additional shape parameter to the Johnson S_B distribution and studied a quantile regression model for unit interval response variables. However, they considered the model only based on the normal distribution. Other quantile regression models for limited range response variables are presented in [6,8,12].

In this paper, we formulate two rich classes of parametric quantile regression models for a bounded response, where the response variable is power Johnson S_B distributed [11] using a new parametrization of this distribution that is indexed by quantile (not only for median regression) and shape parameters. The estimation and inference for the proposed quantile regression models can be carried out based on the likelihood paradigm. Furthermore, we give full diagnostic tools and discuss a type of residual. The main motivations for these new parametric quantile regression models are fourfold: (i) the Johnson S_B and power Johnson S_B regression models are themselves special cases of the proposed quantile models; (ii) the first proposed model has a parameter which controls the shape and skewness of the distribution; (iii) the parameter estimation for the second proposed model has a lower computing cost; and (iv) we considered the model based on several models (logistic, Cauchy, and normal) and several link functions.

The article is organized as follows. In Section 2, we construct two new quantile regression models for bounded response variables. Estimation, residuals, and diagnostic measures are discussed in Section 3. Section 4 discusses some simulation results for the maximum likelihood (ML) estimation method. The effectiveness of our models is illustrated in Section 5 using the COVID-19 mortality rate in different countries. Final comments are presented in Section 6.

2. The Generalized Johnson S_B Distribution

Lemonte and Bazán [9] introduced a new class of distributions named the generalized Johnson S_B (“GJS” for short) distribution. The class was defined by the transformation $Y = Q^{-1}((X - \gamma)/\delta) \in (0, 1)$, where $\gamma \in \mathbb{R}$, $\delta > 0$, $Q(y) = \log(y/(1 - y))$ is the logit function (also representing the quantile function for the standard logistic distribution), and $X \sim S(0, 1; g)$, i.e., the symmetrical family of distributions with pdf given by $g(w)$, $w \in \mathbb{R}$, where g is a function such that $g : \mathbb{R} \rightarrow [0, \infty)$. Considering the reparameterization $\gamma = -\delta Q(\xi)$, the cdf of the GJS is given by

$$F(y; \xi, \delta) = \int_{-\infty}^{\delta[Q(y) - Q(\xi)]} g(u) du, \quad y, \xi \in (0, 1).$$

As $F(\xi; \xi, \delta) = 1/2$, the parameter ξ represents directly the median of the distribution. Additionally, the authors interpreted δ as a dispersion parameter. Therefore, a regression structure on ξ and δ was studied by the authors, providing a rich class of median regression models with varying dispersion. Cancho et al. [11] considered $g(u) = \phi(u)$ (where $\phi(\cdot)$ denotes the pdf of the standard normal model) and the power model transformation [13,14] to extend this class of models (named PJSB), in which the cdf is given by

$$F(y; \alpha, \gamma, \delta) = [\Phi(\gamma + \delta Q(y))]^\alpha, \quad y \in (0, 1), \alpha, \delta > 0, \gamma \in \mathbb{R}.$$

In addition to the logistic model, the authors also considered $Q(y)$ as the quantile function for the normal, Cauchy, Gumbel, and reverse Gumbel models. Thus, the pdf of the PJSB model is

$$f(y; \gamma, \delta, \alpha) = \delta \alpha [\Phi(\gamma + \delta Q(y))]^{\alpha-1} \phi(\gamma + \delta Q(y)) \left| \frac{dQ(y)}{dy} \right|, \quad y \in (0, 1).$$

Defining $x_q = \Phi^{-1}(q^{1/\alpha})$, the authors considered the reparametrization $\psi = Q^{-1}\left(\frac{x_{0.5}(\alpha) - \gamma}{\delta}\right)$, which represents the median of the PJSB distribution (for any $Q(\cdot)$ quantile function). As $\gamma = x_{0.5}(\alpha) - \delta Q(\psi)$, the pdf of the PJSB can be expressed as

$$f(y; \psi, \delta, \alpha) = \delta \alpha [\Phi(\delta[Q(y) - Q(\psi)] + x_{0.5}(\alpha))]^{\alpha-1} \phi(\delta[Q(y) - Q(\psi)] + x_{0.5}(\alpha)) \left| \frac{dQ(y)}{dy} \right|, \quad y \in (0, 1).$$

The authors proposed a regression model for ψ and δ in this model. However, this model can be restrictive, because it considers only the normal distribution. For this reason, we consider the power model transformation of Lehmann [14] and Durrans [13] for the GJS distribution of Lemonte and Bazán [9], i.e., the power generalized Johnson S_B (PGJSB) distribution, with cdf given by

$$F(y; \xi, \delta, \alpha) = \left(\int_{-\infty}^{\delta[Q(y) - Q(\xi)]} g(u) du \right)^\alpha = [G(\delta[Q(y) - Q(\xi)])]^\alpha = [G(\gamma + \delta Q(y))]^\alpha, \quad y \in (0, 1) \tag{1}$$

and pdf given by

$$f(y; \gamma, \delta, \alpha) = \delta \alpha [G(\gamma + \delta Q(y))]^{\alpha-1} g(\gamma + \delta Q(y)) \left| \frac{dQ(y)}{dy} \right|, \quad y \in (0, 1),$$

where G is the cdf related to g . As will be discussed later, α modifies the shape of the distribution. Evidently, for $G = \Phi$, we recover the model in [11]. However, we are interested in modeling a general quantile, say q , not only the median. In this work, we discuss two ways to model the $100 \times q$ th quantile considering the PGJSB model.

1. We note that $\psi = Q^{-1}\left(\frac{x_q^*(\alpha) - \gamma}{\delta}\right)$ is the $100 \times q$ th quantile for the PGJSB model, where $x_q^*(\alpha) = G^{-1}(q^{1/\alpha})$. Based on this idea, we also can reparametrize the model defining $\gamma = x_q^*(\alpha) - \delta Q(\psi)$. The pdf for this reparametrization is

$$f(y; \psi, \delta, \alpha) = \delta \alpha [G(\delta[Q(y) - Q(\psi)] + x_q^*(\alpha))]^{\alpha-1} g(\delta[Q(y) - Q(\psi)] + x_q^*(\alpha)) \left| \frac{dQ(y)}{dy} \right|, \quad y \in (0, 1). \tag{2}$$

In this work, we will refer to this specific parametrization as $RPGJSB1_q(\psi, \delta, \alpha)$.

2. Although α is a parameter that needs to be estimated from the sample, we can consider $\alpha(q) = -\log(q) / \log(2)$, $q \in (0, 1)$ as fixed. With this definition, the cdf in (1) evaluated in ξ is given by $F(\xi; \xi, \delta) = (1/2)^{\alpha(q)} = q$. Therefore, fixing $\alpha(q) = -\log(q) / \log(2)$, $q \in (0, 1)$, we have that ξ represents the $100 \times q$ th quantile of the distribution; further, as in the work of Lemonte and Bazán [9], δ also can be interpreted as a dispersion parameter. We will refer to this parametrization as $RPGJSB2_q(\xi, \delta)$.

In both cases, the $RPGJSB1_q$ and $RPGJSB2_q$ models can be used to define a rich class for performing quantile regression for data in the $(0, 1)$ interval (not only for median regression). The advantage of the $RPGJSB1_q$ model is that α , for a fixed quantile ψ , controls the shape of the distribution (different α 's produce different shapes). However, in this parametrization the shape of the model also depends on ψ . As we will perform regression on ψ , this indicates that the shape related to the distribution of the quantile depends on the covariates. A second problem is the computing cost, because evaluating Equation (2) can be hard to compute for some combinations of g and Q . On the other hand, the advantage of $RPGJSB2_q$ is its parsimony (because one of the parameters is not estimated) and the

reduction in the computing costs, because α is considered fixed. However, in the $RPGJSB2_q$ model the shape of the distribution is maintained (because the model belongs to the location-scale family of distributions) because that shape is “fixed”.

Figure 1 shows the density function for the $RPGJSB1_q(\psi, \delta = 1, \alpha)$ model with logit link and $G = \Phi$ under different combinations of q, ψ , and α . From Figure 1, note that the proposed model is very flexible since its density can assume different shapes.

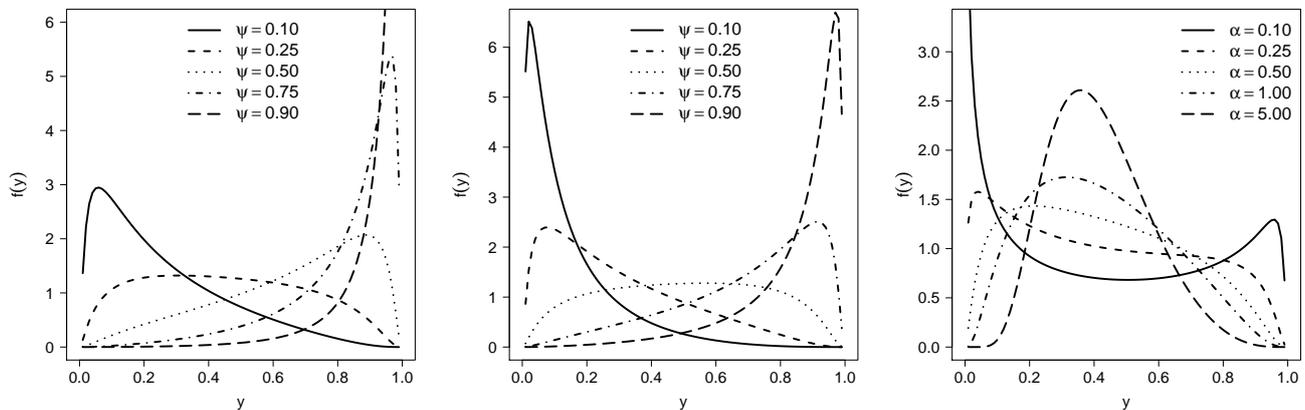


Figure 1. Pdf for $RPGJSB1_q(\psi, \delta = 1, \alpha)$ model with logit link and $G = \Phi$. **Left panel:** $q = 0.25$, $\alpha = 0.5$, and varying ψ ; **center panel:** $q = 0.5$, $\alpha = 0.5$, and varying ψ ; **right panel:** $q = 0.5$, $\psi = 0.4$, and varying α .

3. The Inference and Its Associated Diagnostic Analysis

In this section, we discuss some aspects related to the inference, residuals, and diagnostic analysis of the $RPGJSB1_q$ and $RPGJSB2_q$ quantile regression models.

3.1. Inference

Let y_1, \dots, y_n be an independent random variable such that $y_i \sim RPGJSB1_q(\psi_i, \delta_i, \alpha)$ or $y_i \sim RPGJSB2_q(\xi_i, \delta_i)$. Suppose the $100 \times q$ th quantile ψ for the $RPGJSB1_q$ model and the dispersion parameter δ satisfies the following functional relations

$$Q(\psi_i) = \eta_{1i} = \mathbf{x}_i^\top \boldsymbol{\beta} \quad \text{and} \quad \log(\delta_i) = \eta_{2i} = \mathbf{z}_i^\top \boldsymbol{\nu} \tag{3}$$

or

$$Q(\xi_i) = \eta_{1i} = \mathbf{x}_i^\top \boldsymbol{\beta} \quad \text{and} \quad \log(\delta_i) = \eta_{2i} = \mathbf{z}_i^\top \boldsymbol{\nu}, \tag{4}$$

for the $RPGJSB2_q$ model, where $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)^\top$, and $\boldsymbol{\nu} = (\nu_1, \dots, \nu_r)^\top$ are vectors of unknown regression coefficients, which are assumed to be functionally independent, $\boldsymbol{\beta} \in \mathbb{R}^p$ and $\boldsymbol{\nu} \in \mathbb{R}^r$, with $p + r < n$, η_{1i} and η_{2i} are the linear predictors, and $\mathbf{x}_i = (x_{i1}, \dots, x_{ip})^\top$ and $\mathbf{z}_i = (z_{i1}, \dots, z_{ir})^\top$ are observations on p and r , known regressors, for $i = 1, \dots, n$. Furthermore, we assume that the covariate matrices $\mathbf{X} = (\mathbf{x}_1, \dots, \mathbf{x}_n)^\top$ and $\mathbf{Z} = (\mathbf{z}_1, \dots, \mathbf{z}_n)^\top$ have ranks p and r , respectively. The log-likelihood function for the $RPGJSB1_q$ model is given by

$$\ell_1(\boldsymbol{\theta}) = \sum_{i=1}^n \left\{ \log(\delta_i) + \log(\alpha) + (\alpha - 1) \log \left\{ G \left(\delta_i [Q(y_i) - Q(\psi_i)] + x_q^*(\alpha) \right) \right\} \right. \\ \left. \log \left\{ g \left(\delta_i [Q(y_i) - Q(\psi_i)] + x_q^*(\alpha) \right) \right\} + \log \left| \frac{dQ(y_i)}{dy_i} \right| \right\}; \tag{5}$$

whereas that for the RPGJSB_{2q} model is given by

$$\ell_2(\boldsymbol{\theta}) = \sum_{i=1}^n \left\{ \log(\delta_i) + \log(\alpha) + (\alpha - 1) \log[G(\delta_i[Q(y_i) - Q(\xi_i)])] \right. \\ \left. \log\{g(\delta_i[Q(y_i) - Q(\xi_i)])\} + \log\left|\frac{dQ(y_i)}{dy_i}\right| \right\}. \tag{6}$$

Note that $\boldsymbol{\theta} = (\boldsymbol{\beta}^\top, \boldsymbol{v}^\top, \alpha)$ and $\boldsymbol{\theta} = (\boldsymbol{\beta}^\top, \boldsymbol{v}^\top)$ are the parameter vectors for the RPGJSB_{1q} and RPGJSB_{2q} models, respectively. The ML estimator of $\boldsymbol{\theta}$, say $\hat{\boldsymbol{\theta}}$, is obtained by maximizing Equation (5) or Equation (6), depending on the model considered (RPGJSB_{1q} or RPGJSB_{2q}). We considered the maximization procedure based on the Broyden–Fletcher–Goldfarb–Shanno (BFGS) method (see details in [15]) initialized with a vector of zeros. To validate a solution, we examined: (i) Whether convergence was achieved; and (ii) whether the determinant of the Hessian matrix evaluated at the point estimated was negative. If the two conditions were not satisfied, we re-ran the procedure based on initialization with a random vector generated by independent standard normal variables until (i) and (ii) were satisfied. Under the usual regularity conditions (see Cox and Hinkley [16]) $\boldsymbol{\theta}$ is consistent. Moreover,

$$\boldsymbol{\iota}^{-1}(\hat{\boldsymbol{\theta}}) [\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}] \xrightarrow{D} N_{p+r}(\mathbf{0}_{p+r}, \mathbf{I}_{p+r}), \quad \text{as } n \rightarrow +\infty,$$

where $\boldsymbol{\iota}(\hat{\boldsymbol{\theta}}) = -\partial^2 \ell_l(\boldsymbol{\theta}) / \partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^\top |_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}}$ is minus the estimated Hessian matrix for the RPGJSB_{1q} ($l = 1$) and RPGJSB_{2q} ($l = 2$) models, respectively.

3.2. Residuals

In order to assess whether the postulated model was correct, we considered the randomized quantile residuals (RQRs) proposed by Dunn and Smyth [17]. For the RPGJSB_{1q} model, these residuals are given by

$$\hat{r}_i = \Phi^{-1} \left([G(\hat{\delta}_i[Q(y_i) - Q(\hat{\psi}_i)] + x_q^*(\hat{\alpha}))] \hat{\alpha} \right), \quad i = 1, \dots, n;$$

whereas for the RPGJSB_{2q} model, the RQRs are given by

$$\hat{r}_i = \Phi^{-1} \left([G(\hat{\delta}_i[Q(y_i) - Q(\hat{\xi}_i)])]^\alpha(q) \right), \quad i = 1, \dots, n.$$

$\hat{\delta}_i, \hat{\xi}_i$ and $\hat{\psi}_i, i = 1, \dots, n$, corresponding to the expressions in Equations (3) and (4) were evaluated in $\hat{\boldsymbol{\beta}}$ and $\hat{\boldsymbol{v}}$, for each model, respectively. If the model was correctly specified, the distribution of $\hat{r}_1, \dots, \hat{r}_n$ would be standard normal, which can be validated by different normality tests, such as Kolmogorov–Smirnov (KS), Shapiro–Wilks (SW), Anderson–Darling (AD) and the Cramér–Von-Mises (CVM) tests. See [18] for a discussion of these tests.

3.3. Local Influence

The local influence method suggested by Cook [19] evaluates the simultaneous effect of observations on the ML estimator without removing it from the data set, based on the curvature of the plane of the log-likelihood function. Consider $\ell_1(\boldsymbol{\theta}_1; \boldsymbol{w})$ and $\ell_2(\boldsymbol{\theta}_2; \boldsymbol{w})$ the log-likelihood functions corresponding to the RPGJSB_{1q} and RPGJSB_{2q} models, respectively, but now perturbed by \boldsymbol{w} , a vector of perturbations. \boldsymbol{w} belongs to a subset $\Omega \in \mathbb{R}^n$, and \boldsymbol{w}_0 is a nonperturbed $n \times 1$ vector, such that $\ell_l(\boldsymbol{\theta}; \boldsymbol{w}_0) = \ell_l(\boldsymbol{\theta})$, for all $\boldsymbol{\theta}, l = 1, 2$. In this case, the likelihood displacement (LD) is $LD(\boldsymbol{\theta}) = 2(\ell_l(\hat{\boldsymbol{\theta}}) - \ell_l(\hat{\boldsymbol{\theta}}_{\boldsymbol{w}}))$, where $\hat{\boldsymbol{\theta}}_{\boldsymbol{w}}$ denotes the ML estimate of $\boldsymbol{\theta}$ on the perturbed regression models; that is, $\hat{\boldsymbol{\theta}}_{\boldsymbol{w}}$ is obtained from $\ell_l(\boldsymbol{\theta}; \boldsymbol{w})$. Note that $\ell_l(\boldsymbol{\theta}; \boldsymbol{w})$ can be used to assess the influence of the perturbation of the ML estimate. Cook [19] showed that the normal curvature for $\hat{\boldsymbol{\theta}}$ in direction \boldsymbol{d} , with $\|\boldsymbol{d}\| = 1$, is expressed as $C_d(\hat{\boldsymbol{\theta}}) = 2|\boldsymbol{d}^\top \nabla^\top \Sigma(\hat{\boldsymbol{\theta}})^{-1} \nabla \boldsymbol{d}|$, where ∇ is a $(p + r) \times n$ matrix of perturbations with

elements $\nabla_{ji} = \partial^2 \ell_l(\boldsymbol{\theta}; \boldsymbol{w}) / \partial \theta_j \partial w_i$, evaluated at $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}$ and $\boldsymbol{w} = \boldsymbol{w}_0$, for $j = 1, \dots, p + r$ and $i = 1, \dots, n$. A local influence diagnostic is generally based on index plots. For example, denoting $\Sigma(\boldsymbol{\theta})$ the observed Fisher information matrix, the index graph of the eigenvector \boldsymbol{d}_{max} corresponding to the maximum eigenvalue of $\boldsymbol{B}(\boldsymbol{\theta}) = -\nabla^\top \Sigma(\boldsymbol{\theta})^{-1} \nabla$, say $C_{\boldsymbol{d}_{max}}(\boldsymbol{\theta})$, evaluated at $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}$, can detect those cases that, under small perturbations, exert a strong influence on $LD(\boldsymbol{\theta})$. Another important direction of interest is $\boldsymbol{d}_i = \boldsymbol{e}_{in}$, which corresponds to the direction of the case i , where \boldsymbol{e}_{in} is an $n \times 1$ vector of zeros with value equal to one at the i th position; that is, $\{\boldsymbol{e}_{in}, 1 \leq i \leq n\}$ is the canonical basis of \mathbb{R}^n . In this case, the normal curvature is $C_i(\boldsymbol{\theta}) = 2|b_{ii}|$, where b_{ii} is the i th diagonal element of $\boldsymbol{B}(\boldsymbol{\theta})$ given above, for $i = 1, \dots, n$, evaluated $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}$. If $C_i(\hat{\boldsymbol{\theta}}) > 2\bar{C}(\hat{\boldsymbol{\theta}})$, where $\bar{C}(\hat{\boldsymbol{\theta}}) = \sum_{i=1}^n C_i(\hat{\boldsymbol{\theta}}) / n$, it indicates case i as potentially influential. This procedure is called the total local influence of the case i and can be carried out for $\boldsymbol{\theta}$, $\boldsymbol{\beta}$, or \boldsymbol{v} , denoted by $C_i(\boldsymbol{\theta})$, $C_i(\boldsymbol{\beta})$, and $C_i(\boldsymbol{v})$, respectively. We calculated the matrix ∇ for three different perturbation schemes, namely: case weighting perturbation, response perturbation, and explanatory variable perturbation.

3.3.1. Perturbation of the Case Weights

In this case, the perturbed log-likelihood function is given by $\ell_l(\boldsymbol{\theta}; \boldsymbol{w}) = \sum_{i=1}^n w_i \ell_l(\boldsymbol{\theta})$ for $RPGJSB1_q$ ($l = 1$) and $RPGJSB2_q$ ($l = 2$), respectively, with $0 \leq w_i \leq 1$, for $i = 1, \dots, n$, and $\boldsymbol{w}_0 = \mathbf{1}^\top$ (all-ones vector). Hence, the perturbation matrices for the $RPGJSB1_q$ and $RPGJSB2_q$ models are given by

$$\hat{\nabla}_1 = \begin{pmatrix} \boldsymbol{X}^\top \hat{\boldsymbol{D}}_1 \hat{\boldsymbol{D}}_3 \\ \boldsymbol{Z}^\top \hat{\boldsymbol{D}}_2 \hat{\boldsymbol{D}}_4 \end{pmatrix} \quad \text{and} \quad \hat{\nabla}_2 = \begin{pmatrix} \boldsymbol{X}^\top \hat{\boldsymbol{D}}_5 \hat{\boldsymbol{D}}_7 \hat{\boldsymbol{D}}_9 \\ \boldsymbol{Z}^\top \hat{\boldsymbol{D}}_6 \hat{\boldsymbol{D}}_8 \hat{\boldsymbol{D}}_9 \end{pmatrix},$$

respectively, with $\boldsymbol{D}_1 = [a_i \iota_{ij}]$, $\boldsymbol{D}_2 = [b_i \iota_{ij}]$, $\boldsymbol{D}_3 = [\dot{d}_{\psi} \iota_{ij}]$, and $\boldsymbol{D}_4 = [\dot{d}_{\delta} \iota_{ij}]$, where $a_i = \partial \psi_i / \partial \eta_{i1}$ and $b_i = \partial \delta_i / \partial \eta_{i2}$, as defined from (3); $\dot{d}_{\psi} = \partial \ell_1(\psi_i, \delta_i) / \partial \psi_i$ and $\dot{d}_{\delta} = \partial \ell_1(\psi_i, \delta_i) / \partial \delta_i$, as defined from the $RPGJSB1_q$ model, and ι_{ij} is the Kronecker delta for $i, j = 1, 2, \dots, n$. Similarly, $\boldsymbol{D}_5 = [c_i \iota_{ij}]$, $\boldsymbol{D}_6 = [d_i \iota_{ij}]$, $\boldsymbol{D}_7 = [\dot{d}_{\xi} \iota_{ij}]$, $\boldsymbol{D}_8 = [\dot{d}_{\delta} \iota_{ij}]$, and $\boldsymbol{D}_9 = [\dot{d}_{\alpha} \iota_{ij}]$, where $c_i = \partial \xi_i / \partial \eta_{i1}$ and $d_i = \partial \delta_i / \partial \eta_{i2}$, as defined from (4); $\dot{d}_{\xi} = \partial \ell_2(\xi_i, \delta_i, \alpha) / \partial \xi_i$, $\dot{d}_{\delta} = \partial \ell_2(\xi_i, \delta_i, \alpha) / \partial \delta_i$, and $\dot{d}_{\alpha} = \partial \ell_2(\xi_i, \delta_i, \alpha) / \partial \alpha$, as defined from the $RPGJSB2_q$ model.

3.3.2. Perturbation of the Response

Now, consider a multiplicative perturbation of the i th response by making $y_i(w_i) = y_i w_i s_y$, where s_y represents a scale factor, and $w_i \in \mathbb{R}$, for $i = 1, \dots, n$. Then, under the scheme of response perturbation, the log-likelihood function is given by $\ell_1(\boldsymbol{\theta}; \boldsymbol{w}) = \sum_{i=1}^n \ell_1(\psi_i, \delta_i, \alpha; \boldsymbol{w})$ for the $RPGJSB1_q$ model and $\ell_2(\boldsymbol{\theta}; \boldsymbol{w}) = \sum_{i=1}^n \ell_2(\xi_i, \delta_i; \boldsymbol{w})$ for the $RPGJSB2_q$ model, where

$$\begin{aligned} \ell_1(\psi_i, \delta_i, \alpha; \boldsymbol{w}) &= (\alpha - 1) \log(G(\tau_{1i})) + \log(\alpha \delta_i) + \log(g(\tau_{1i})) + \log(|w_i s_y \dot{Q}_y(y_i w_i s_y)|) \\ \ell_2(\xi_i, \delta_i; \boldsymbol{w}) &= (\alpha - 1) \log(G(\tau_{2i})) + \log(\alpha \delta_i) + \log(g(\tau_{2i})) + \log(|w_i s_y \dot{Q}_y(y_i w_i s_y)|) \end{aligned}$$

with $\tau_{1i} = \delta_i(Q(y_i w_i s_y) - Q(\psi_i))$ and $\tau_{2i} = \delta_i(Q(y_i w_i s_y) - Q(\xi_i)) + x_q^*(\alpha)$.

The disturbance matrices of the $RPGJSB1_q$ and $RPGJSB2_q$ models, here, take the form

$$\hat{\nabla}_1 = \begin{pmatrix} \boldsymbol{X}^\top \hat{\boldsymbol{D}}_1 \hat{\boldsymbol{D}}_{10} \boldsymbol{S} \\ \boldsymbol{Z}^\top \hat{\boldsymbol{D}}_2 \hat{\boldsymbol{D}}_{11} \boldsymbol{S} \end{pmatrix} \quad \text{and} \quad \hat{\nabla}_2 = \begin{pmatrix} \boldsymbol{X}^\top \hat{\boldsymbol{D}}_5 \hat{\boldsymbol{D}}_{12} \hat{\boldsymbol{D}}_{14} \boldsymbol{S} \\ \boldsymbol{Z}^\top \hat{\boldsymbol{D}}_6 \hat{\boldsymbol{D}}_{13} \hat{\boldsymbol{D}}_{14} \boldsymbol{S} \end{pmatrix}$$

where $\boldsymbol{S} = [s_y \iota_{ij}]$, the i th element of matrices \boldsymbol{D}_{10} and \boldsymbol{D}_{11} for model $RPGJSB1_q$ and matrices \boldsymbol{D}_{12} , \boldsymbol{D}_{13} , and \boldsymbol{D}_{14} for model $RPGJSB2_q$ are detailed in Appendix A.1.

3.3.3. Perturbation of the Predictor

Now, consider a multiplicative perturbation of the i th predictor by making $x_i(w_i) = \boldsymbol{x}_i^\top w_i$ and $z_i(w_i) = \boldsymbol{z}_i^\top w_i$, for $w_i \in \mathbb{R}$, $i = 1, \dots, n$. Then, under the scheme of prediction perturbation, the log-likelihood function is given by $\ell_1(\boldsymbol{\theta}; \boldsymbol{w}) = \sum_{i=1}^n \ell_1(\psi_i^*, \delta_i^*)$ for the $RPGJSB1_q$ model and $\ell_2(\boldsymbol{\theta}; \boldsymbol{w}) = \sum_{i=1}^n \ell_2(\xi_i^*, \delta_i^*, \alpha)$ for the $RPGJSB2_q$ model, where $Q(\psi_i^*) =$

$x_i^\top \beta w_i$ and $\delta_i^* = \exp\{z_i^\top \nu w_i\}$ for the RPGJSB1 $_q$ model and $Q(\xi_i^*) = x_i^\top \beta w_i$ and $\delta_i^* = \exp\{z_i^\top \nu w_i\}$ for the RPGJSB2 $_q$ model.

The disturbance matrices of RPGJSB1 $_q$ and RPGJSB2 $_q$ models, here, take the form

$$\widehat{V}_1 = \begin{pmatrix} X^\top \widehat{D}_{15} \\ Z^\top \widehat{D}_{16} \end{pmatrix} \quad \text{and} \quad \widehat{V}_2 = \begin{pmatrix} X^\top \widehat{D}_{17} \widehat{D}_{19} \\ Z^\top \widehat{D}_{18} \widehat{D}_{19} \end{pmatrix}$$

where the i th elements of matrices D_{15} and D_{16} for RPGJSB1 $_q$ model and matrices D_{17} , D_{18} , and D_{19} for RPGJSB2 $_q$ model are detailed in Appendix A.2.

4. Simulation Study

In this section, we present a simulation study to assess the performance of the ML estimators of $\theta = (\beta, \nu, \alpha)^\top$ under different scenarios. All the codes were developed in R [20] version 4.0.2 and are available upon request. This study was performed based on the RPGJSB1 $_q$ model. First, we assumed that G and the link function were correctly specified. We considered $x_i = z_i$, where both matrices included an intercept and a covariate. The covariates were drawn from the $U(-5.478, -2.305)$ distribution. We considered the logistic and normal models for G and the logit and loglog link functions. The true values for the parameters were considered as the estimated parameters for three values for $q = \{0.1, 0.5, 0.9\}$. Table 1 presents the values considered for each combination of link function and quantile. We also considered three sample sizes: 100, 200, and 500.

Table 1. True parameters used for simulation studies.

| Link | q | Logistic | | | | | Normal | | | | |
|--------|-----|-----------|-----------|---------|---------|----------------|-----------|-----------|---------|---------|----------------|
| | | β_0 | β_1 | ν_0 | ν_1 | $\log(\alpha)$ | β_0 | β_1 | ν_0 | ν_1 | $\log(\alpha)$ |
| logit | 0.1 | 4.9 | 2.6 | 2.2 | 0.4 | -0.7 | 4.4 | 2.4 | 1.5 | 0.3 | -1.4 |
| | 0.5 | 4.8 | 2.1 | 2.2 | 0.4 | -0.7 | 4.6 | 2.1 | 1.5 | 0.3 | -1.4 |
| | 0.9 | 4.7 | 1.8 | 2.2 | 0.4 | -0.7 | 4.8 | 1.9 | 1.5 | 0.3 | -1.4 |
| loglog | 0.1 | 1.3 | 0.8 | 0.8 | -0.3 | 0.1 | 1.2 | 0.7 | -0.1 | -0.3 | 1.1 |
| | 0.5 | 2.1 | 0.9 | 1.0 | -0.2 | 0.1 | 2.0 | 0.9 | 0.0 | -0.3 | 1.0 |
| | 0.9 | 2.8 | 1.0 | 1.1 | -0.2 | 0.1 | 2.8 | 1.0 | 0.1 | -0.2 | 1.0 |

As mentioned previously, to validate a solution we examined whether convergence was achieved and whether the determinant of the Hessian matrix was positive. If the two conditions were not satisfied, we re-ran the procedure initialized with a random vector generated by independent standard normal variables until both conditions were satisfied. For each combination of G , link, q , and sample size, we considered 10,000 replicates, and in each case, the estimation was performed based on the same G and link function. Based on the 10,000 replicates, we report the bias for each estimator, the standard error of the estimated parameters (SE_1), the mean of the estimated standard errors (SE_2), and the 95% coverage probabilities (CP). Tables 2 and 3 summarize the results. Note that the bias of the parameters was reduced, and the terms SE_1 and SE_2 were closer when n was increased, suggesting that the estimators were asymptotically consistent. Additionally, when the sample size was increased, the CP were closer to the nominal value used. Finally, Table 4 presents the percentage of times where the algorithm converged when it was initialized with a vector of zeros. Note that the maximization procedure converged in at least 89.43% of the samples generated, and this percentage increased with the sample size.

Table 2. Estimated bias, standard error of the estimated parameters (SE_1), the mean of the estimated standard errors (SE_2), and 95% coverage probabilities (CP), when G and the link are correctly specified (case G is the cdf of the logistic distribution).

| G | Link | q | Parameter | $n = 100$ | | | | $n = 200$ | | | | $n = 500$ | | | |
|----------|-------|-----|----------------|-----------|--------|--------|-------|-----------|--------|--------|-------|-----------|--------|--------|-------|
| | | | | Bias | SE_1 | SE_2 | CP | Bias | SE_1 | SE_2 | CP | Bias | SE_1 | SE_2 | CP |
| logistic | logit | 0.1 | β_0 | -0.034 | 0.753 | 0.728 | 0.938 | -0.017 | 0.538 | 0.529 | 0.946 | -0.007 | 0.345 | 0.339 | 0.946 |
| | | | β_1 | -0.015 | 0.238 | 0.229 | 0.934 | -0.007 | 0.166 | 0.163 | 0.942 | -0.003 | 0.104 | 0.102 | 0.947 |
| | | | ν_0 | 0.041 | 0.381 | 0.367 | 0.935 | 0.020 | 0.269 | 0.263 | 0.942 | 0.009 | 0.171 | 0.170 | 0.947 |
| | | | ν_1 | -0.001 | 0.088 | 0.085 | 0.939 | 0.000 | 0.061 | 0.060 | 0.946 | 0.000 | 0.039 | 0.038 | 0.946 |
| | | | $\log(\alpha)$ | -0.004 | 0.355 | 0.331 | 0.947 | -0.002 | 0.232 | 0.224 | 0.946 | -0.002 | 0.140 | 0.138 | 0.948 |
| | | 0.5 | β_0 | -0.017 | 0.485 | 0.472 | 0.941 | 0.001 | 0.322 | 0.319 | 0.946 | -0.003 | 0.204 | 0.205 | 0.950 |
| | | | β_1 | -0.005 | 0.146 | 0.142 | 0.939 | 0.000 | 0.096 | 0.095 | 0.946 | -0.001 | 0.061 | 0.061 | 0.949 |
| | | | ν_0 | 0.046 | 0.452 | 0.443 | 0.946 | 0.027 | 0.296 | 0.294 | 0.948 | 0.007 | 0.183 | 0.182 | 0.949 |
| | | | ν_1 | 0.002 | 0.107 | 0.106 | 0.946 | 0.002 | 0.068 | 0.068 | 0.949 | 0.000 | 0.042 | 0.042 | 0.951 |
| | | | $\log(\alpha)$ | 0.004 | 0.352 | 0.331 | 0.952 | -0.001 | 0.231 | 0.224 | 0.948 | 0.000 | 0.142 | 0.139 | 0.947 |
| | | 0.9 | β_0 | -0.001 | 0.620 | 0.591 | 0.930 | -0.002 | 0.369 | 0.363 | 0.942 | 0.002 | 0.237 | 0.236 | 0.950 |
| | | | β_1 | 0.004 | 0.177 | 0.169 | 0.932 | 0.002 | 0.112 | 0.111 | 0.943 | 0.001 | 0.072 | 0.072 | 0.948 |
| | | | ν_0 | 0.060 | 0.461 | 0.443 | 0.943 | 0.024 | 0.289 | 0.283 | 0.943 | 0.007 | 0.184 | 0.182 | 0.946 |
| | | | ν_1 | 0.006 | 0.103 | 0.100 | 0.938 | 0.002 | 0.066 | 0.065 | 0.946 | 0.000 | 0.043 | 0.043 | 0.945 |
| | | | $\log(\alpha)$ | 0.010 | 0.362 | 0.334 | 0.946 | 0.003 | 0.234 | 0.224 | 0.949 | 0.002 | 0.140 | 0.139 | 0.949 |
| loglog | logit | 0.1 | β_0 | 0.008 | 0.175 | 0.168 | 0.931 | 0.002 | 0.116 | 0.113 | 0.938 | 0.000 | 0.071 | 0.071 | 0.949 |
| | | | β_1 | 0.001 | 0.039 | 0.037 | 0.935 | 0.000 | 0.026 | 0.025 | 0.937 | 0.000 | 0.016 | 0.016 | 0.948 |
| | | | ν_0 | 0.020 | 0.413 | 0.398 | 0.944 | 0.005 | 0.280 | 0.275 | 0.946 | -0.001 | 0.165 | 0.167 | 0.950 |
| | | | ν_1 | 0.000 | 0.096 | 0.092 | 0.939 | -0.002 | 0.067 | 0.065 | 0.942 | -0.001 | 0.039 | 0.039 | 0.949 |
| | | | $\log(\alpha)$ | 0.153 | 1.175 | 2.515 | 0.964 | 0.035 | 0.349 | 0.324 | 0.961 | 0.014 | 0.178 | 0.174 | 0.956 |
| | | 0.5 | β_0 | -0.002 | 0.130 | 0.128 | 0.944 | -0.003 | 0.090 | 0.090 | 0.951 | 0.001 | 0.061 | 0.061 | 0.946 |
| | | | β_1 | 0.000 | 0.031 | 0.030 | 0.945 | -0.001 | 0.021 | 0.021 | 0.949 | 0.000 | 0.014 | 0.014 | 0.947 |
| | | | ν_0 | 0.007 | 0.386 | 0.376 | 0.944 | 0.003 | 0.264 | 0.261 | 0.947 | 0.006 | 0.175 | 0.175 | 0.950 |
| | | | ν_1 | -0.003 | 0.093 | 0.091 | 0.945 | -0.002 | 0.063 | 0.062 | 0.949 | 0.000 | 0.041 | 0.041 | 0.950 |
| | | | $\log(\alpha)$ | 0.143 | 1.070 | 2.042 | 0.965 | 0.041 | 0.306 | 0.290 | 0.962 | 0.012 | 0.177 | 0.174 | 0.951 |
| | | 0.9 | β_0 | -0.005 | 0.178 | 0.175 | 0.939 | -0.004 | 0.141 | 0.139 | 0.942 | -0.002 | 0.082 | 0.082 | 0.947 |
| | | | β_1 | -0.001 | 0.042 | 0.041 | 0.940 | -0.001 | 0.033 | 0.032 | 0.943 | 0.000 | 0.019 | 0.019 | 0.947 |
| | | | ν_0 | 0.012 | 0.387 | 0.374 | 0.940 | 0.010 | 0.296 | 0.288 | 0.944 | 0.004 | 0.174 | 0.173 | 0.949 |
| | | | ν_1 | -0.002 | 0.094 | 0.091 | 0.940 | 0.000 | 0.071 | 0.069 | 0.944 | 0.000 | 0.041 | 0.041 | 0.949 |
| | | | $\log(\alpha)$ | 0.133 | 0.968 | 1.596 | 0.965 | 0.042 | 0.311 | 0.290 | 0.961 | 0.014 | 0.177 | 0.174 | 0.952 |

Table 3. Estimated bias, standard error of the estimated parameters (SE_1), the mean of the estimated standard errors (SE_2), and 95% coverage probabilities (CP), when G and the link are correctly specified (case G is the cdf of the normal distribution).

| G | Link | q | Parameter | $n = 100$ | | | | $n = 200$ | | | | $n = 500$ | | | |
|--------|-------|-----|----------------|-----------|--------|--------|-------|-----------|--------|--------|-------|-----------|--------|--------|-------|
| | | | | Bias | SE_1 | SE_2 | CP | Bias | SE_1 | SE_2 | CP | Bias | SE_1 | SE_2 | CP |
| normal | logit | 0.1 | β_0 | -0.004 | 0.725 | 0.711 | 0.939 | 0.000 | 0.470 | 0.473 | 0.952 | -0.002 | 0.289 | 0.291 | 0.951 |
| | | | β_1 | -0.005 | 0.202 | 0.198 | 0.939 | -0.002 | 0.133 | 0.134 | 0.951 | -0.002 | 0.081 | 0.082 | 0.949 |
| | | | ν_0 | 0.912 | 2.171 | 0.730 | 0.847 | 0.243 | 1.024 | 0.450 | 0.946 | 0.045 | 0.270 | 0.248 | 0.954 |
| | | | ν_1 | 0.000 | 0.083 | 0.079 | 0.932 | 0.001 | 0.055 | 0.054 | 0.945 | 0.000 | 0.032 | 0.032 | 0.951 |
| | | | $\log(\alpha)$ | -1.763 | 4.569 | 1.650 | 0.867 | -0.462 | 2.167 | 1.019 | 0.960 | -0.082 | 0.612 | 0.565 | 0.956 |
| | | 0.5 | β_0 | -0.006 | 0.464 | 0.452 | 0.942 | -0.005 | 0.330 | 0.324 | 0.945 | 0.002 | 0.196 | 0.194 | 0.946 |
| | | | β_1 | -0.004 | 0.135 | 0.131 | 0.940 | -0.002 | 0.094 | 0.092 | 0.941 | 0.000 | 0.056 | 0.056 | 0.946 |
| | | | ν_0 | 0.944 | 2.251 | 0.703 | 0.841 | 0.215 | 0.966 | 0.450 | 0.949 | 0.040 | 0.281 | 0.250 | 0.952 |
| | | | ν_1 | 0.002 | 0.082 | 0.079 | 0.939 | 0.001 | 0.056 | 0.055 | 0.947 | 0.000 | 0.034 | 0.033 | 0.950 |
| | | | $\log(\alpha)$ | -1.806 | 4.729 | 1.597 | 0.862 | -0.398 | 2.046 | 1.012 | 0.961 | -0.071 | 0.625 | 0.564 | 0.954 |
| | | 0.9 | β_0 | -0.028 | 0.595 | 0.550 | 0.910 | -0.001 | 0.393 | 0.375 | 0.934 | -0.004 | 0.244 | 0.242 | 0.947 |
| | | | β_1 | -0.002 | 0.165 | 0.153 | 0.912 | 0.003 | 0.111 | 0.106 | 0.933 | 0.000 | 0.069 | 0.069 | 0.949 |
| | | | ν_0 | 0.923 | 2.248 | 0.712 | 0.852 | 0.235 | 1.009 | 0.450 | 0.947 | 0.047 | 0.279 | 0.253 | 0.949 |
| | | | ν_1 | 0.006 | 0.088 | 0.084 | 0.937 | 0.001 | 0.057 | 0.055 | 0.941 | 0.001 | 0.035 | 0.035 | 0.949 |
| | | | $\log(\alpha)$ | -1.733 | 4.706 | 1.576 | 0.871 | -0.434 | 2.133 | 1.008 | 0.961 | -0.083 | 0.614 | 0.563 | 0.956 |

Table 3. Cont.

| G | Link | q | Parameter | n = 100 | | | | n = 200 | | | | n = 500 | | | |
|--------|------|---|----------------|---------|-----------------|-----------------|-------|---------|-----------------|-----------------|-------|---------|-----------------|-----------------|-------|
| | | | | Bias | SE ₁ | SE ₂ | CP | Bias | SE ₁ | SE ₂ | CP | Bias | SE ₁ | SE ₂ | CP |
| loglog | 0.1 | | β_0 | 0.005 | 0.156 | 0.152 | 0.935 | 0.005 | 0.115 | 0.114 | 0.942 | 0.001 | 0.070 | 0.069 | 0.946 |
| | | | β_1 | 0.000 | 0.035 | 0.034 | 0.936 | 0.001 | 0.026 | 0.026 | 0.945 | 0.000 | 0.016 | 0.015 | 0.946 |
| | | | ν_0 | 0.085 | 0.834 | 0.530 | 0.963 | 0.024 | 0.371 | 0.351 | 0.951 | 0.006 | 0.209 | 0.209 | 0.952 |
| | | | ν_1 | −0.004 | 0.077 | 0.076 | 0.942 | −0.001 | 0.059 | 0.058 | 0.946 | −0.001 | 0.035 | 0.035 | 0.951 |
| | | | $\log(\alpha)$ | 1.090 | 23.978 | 3.284 | 0.965 | 0.103 | 1.677 | 1.200 | 0.958 | 0.027 | 0.661 | 0.658 | 0.961 |
| | 0.5 | | β_0 | 0.002 | 0.116 | 0.114 | 0.942 | 0.000 | 0.084 | 0.083 | 0.946 | 0.000 | 0.054 | 0.053 | 0.948 |
| | | | β_1 | 0.000 | 0.026 | 0.025 | 0.942 | 0.000 | 0.019 | 0.019 | 0.947 | 0.000 | 0.012 | 0.012 | 0.947 |
| | | | ν_0 | 0.123 | 0.990 | 0.539 | 0.954 | 0.017 | 0.379 | 0.348 | 0.955 | 0.009 | 0.212 | 0.212 | 0.952 |
| | | | ν_1 | −0.004 | 0.082 | 0.079 | 0.939 | −0.002 | 0.059 | 0.059 | 0.946 | −0.001 | 0.036 | 0.036 | 0.950 |
| | | | $\log(\alpha)$ | 0.612 | 16.917 | 3.114 | 0.964 | 0.091 | 1.453 | 1.150 | 0.963 | 0.017 | 0.654 | 0.645 | 0.957 |
| | 0.9 | | β_0 | −0.016 | 0.224 | 0.219 | 0.935 | −0.007 | 0.169 | 0.167 | 0.939 | −0.004 | 0.095 | 0.095 | 0.943 |
| | | | β_1 | −0.002 | 0.051 | 0.050 | 0.937 | −0.001 | 0.040 | 0.039 | 0.940 | −0.001 | 0.022 | 0.022 | 0.946 |
| | | | ν_0 | 0.125 | 0.934 | 0.538 | 0.958 | 0.029 | 0.386 | 0.357 | 0.951 | 0.008 | 0.208 | 0.207 | 0.951 |
| | | | ν_1 | 0.000 | 0.083 | 0.080 | 0.940 | 0.000 | 0.061 | 0.060 | 0.948 | 0.000 | 0.035 | 0.034 | 0.952 |
| | | | $\log(\alpha)$ | 0.428 | 12.877 | 2.696 | 0.964 | 0.088 | 1.443 | 1.192 | 0.957 | 0.034 | 0.657 | 0.647 | 0.961 |

Table 4. Percentage of time where the maximization algorithm converges with the initial value as the vector zero.

| G | Link | q = 0.1 | | | q = 0.5 | | | q = 0.9 | | |
|----------|--------|---------|--------|--------|---------|--------|--------|---------|--------|--------|
| | | 100 | 200 | 500 | 100 | 200 | 500 | 100 | 200 | 500 |
| logistic | logit | 100.00 | 100.00 | 100.00 | 100.00 | 100.00 | 100.00 | 100.00 | 100.00 | 100.00 |
| | loglog | 99.71 | 100.00 | 100.00 | 99.83 | 100.00 | 100.00 | 99.85 | 100.00 | 100.00 |
| normal | logit | 90.77 | 98.40 | 100.00 | 89.43 | 98.65 | 99.99 | 90.38 | 98.59 | 99.99 |
| | loglog | 99.43 | 99.99 | 100.00 | 99.01 | 99.98 | 100.00 | 99.05 | 99.98 | 100.00 |

5. Data Analysis

In this section, we apply the model to a real data set related to the mortality rate of COVID-19 in different countries to illustrate the performance of the $RPGJSB1_q$ and $RPGJSB2_q$ proposed regression models.

5.1. COVID-19 Data Set

The COVID-19 pandemic has had an unprecedented effect throughout the world. Specifically, it has yielded high mortality rates since its emergence in December 2019, generating social, economic, cultural, and political imbalances. Early studies have shown that statistical analysis can be applied to COVID-19 problems to build predictive models that can assess risk factors and mortality rates [21–23]. Furthermore, the overall mortality rate has been about 5%, while the statistics have shown a rate of around 20% for elderly patients [24]. We considered the following information for countries with at least 1000 reported cases of COVID-19 and at least 100 deaths attributed to the disease, totaling 137 countries as of 25 May 2021.

- **mort:** mortality rate (reported deaths/reported cases since the pandemic started). Mean = 0.020, Median = 0.018, standard deviation = 0.013, minimum = 0.003, and maximum = 0.092.
- **surface:** area of the country (in km²).
- **population:** official estimated population of the country.
- **cont:** continent to which the country belongs (categorized as 1: Africa, Asia, or Oceania; 2: the Americas; 3: Europe; with 69, 29, and 39 countries, respectively. This categorization was based on our previous analysis).

The information was taken from the World Health Organization [25]. It is of interest to model the mortality rate in terms of the surface and the continent of each country (previous analyses suggest that the population was not significant for modeling the mortality rate).

Figure 2 shows the plots for $Q(\text{mort})$ for different link functions versus the $\log(\text{surface})$ and $\log(\text{population})$ separated by cont.

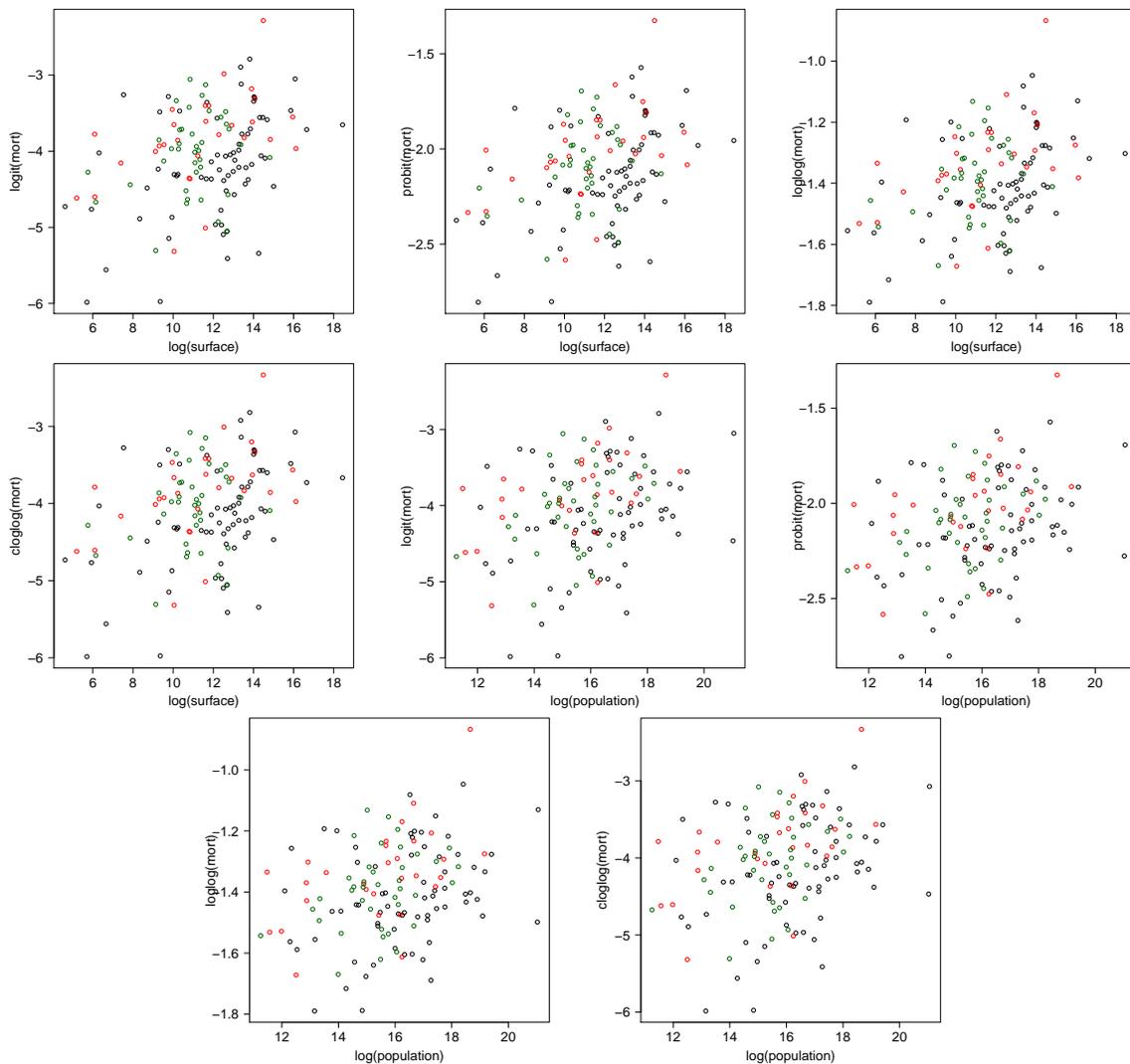


Figure 2. Descriptive plots for $Q(\text{mort})$ versus $\log(\text{surface})$ and versus $\log(\text{population})$ for different link functions: logit, probit, loglog, and cloglog and separated by continent: Africa, Asia, or Oceania (black), the Americas (red), and Europe (green).

5.1.1. Estimation

In view of the above, we modeled the mortality rate using $\text{mort}_i \sim \text{RPGJSB1}_q(\psi_i, \delta_i, \alpha)$, with

$$Q(\psi_i) = \beta_0 + \beta_1 \times \log(\text{population}_i) + \beta_2 \times \log(\text{surface}_i) + \beta_3 \times \text{America}_i + \beta_4 \times \text{Europe}_i$$

and

$$\log(\delta_i) = \nu_0 + \nu_1 \times \text{America}_i + \nu_2 \times \text{Europe}_i$$

or, alternatively, $\text{mort}_i \sim \text{RPGJSB2}_q(\xi_i, \delta_i)$, where $Q(\xi_i) = \beta_0 + \beta_1 \times \log(\text{population}_i) + \beta_2 \times \log(\text{surface}_i) + \beta_3 \times \text{America}_i + \beta_4 \times \text{Europe}_i$ and δ_i was modeled in the same way. In Tables A1 and A2 in Appendix B.1, we present the AIC and BIC for q ranging in the set $\{0.05, 0.10, \dots, 0.90, 0.95\}$ and the RPGJSB1_q and RPGJSB2_q models. Note that the RPGJSB2_q provides a lower AIC than the RPGJSB1_q for all the considered q . Below, we focus on the RPGJSB2_q model, specifically, where G is the cdf of the normal model and the loglog link (which provide the lowest AIC for greater values of q). Tables 5 and A3 in

Appendix B.2 present the estimated parameters for this model for five selected quantiles. We also present the KS, SW, AD, and CVM tests to check the normality of the RQRs. Note that the coefficients related to the $\log(\text{surface})$ and America were significant for modeling the quantile (with a nominal level of 5%) for all the considered q . This can be explained because countries with larger areas may have greater difficulties in providing medical coverage to their inhabitants compared to countries with smaller areas, and some countries in the Americas have been hit hard by the pandemic. The coefficients related to $\log(\text{population})$ and Europe were significant for lower quantiles and not significant for higher quantiles.

Figure 3 presents the point estimation and the 90%, 95%, and 99% confidence intervals for the parameters in terms of quantile q . From Figure 3, the intercept for the quantile increased as q increased, whereas the coefficients related to the quantile of America and Europe decreased when q was increased. Furthermore, the coefficients related to the quantile for $\log(\text{population})$ and $\log(\text{surface})$ and the coefficients related to the scale of America and Europe remained similar for all q . Figure 4 presents the estimated 0.05, 0.25, 0.50, 0.75, and 0.95 quantiles for the mortality rate for different values of $\log(\text{surface})$.

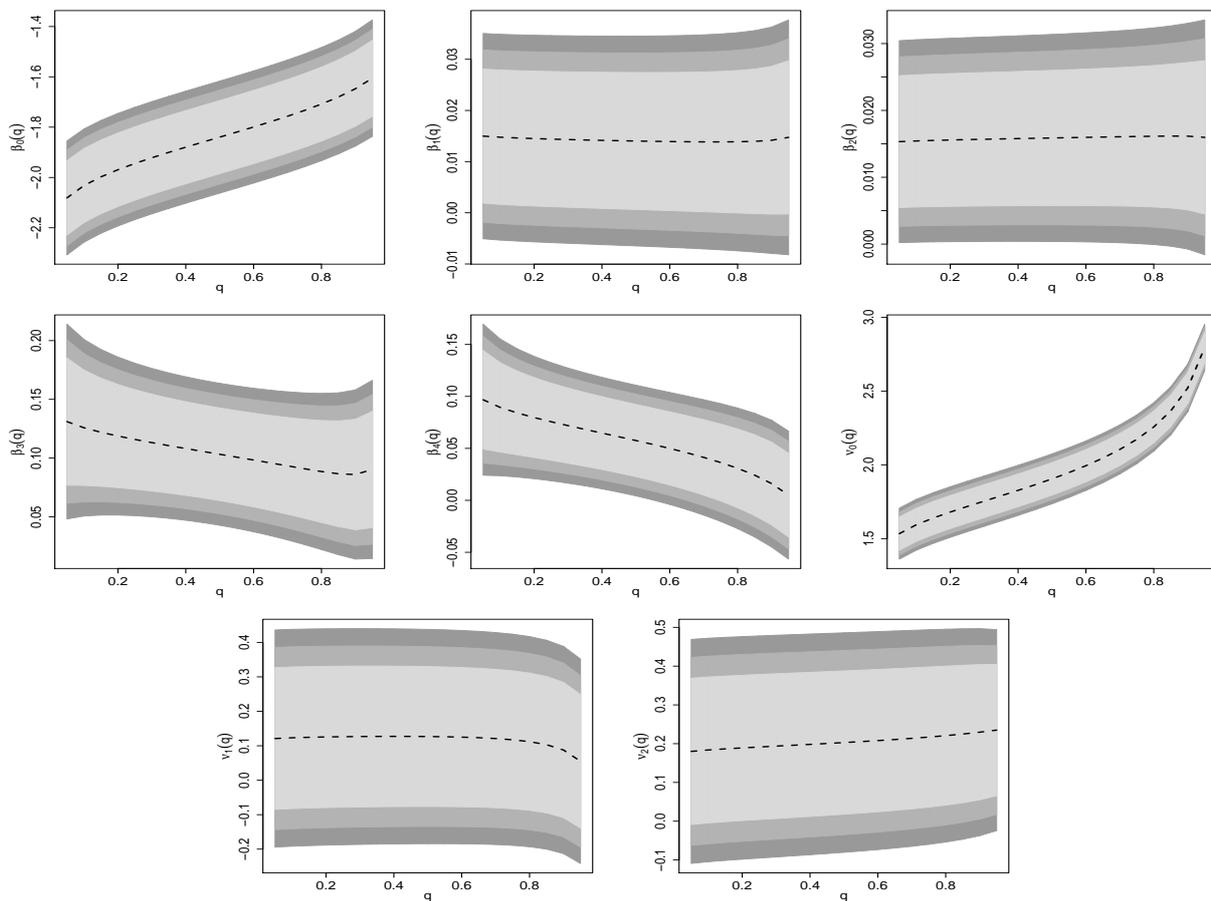


Figure 3. Point estimation and 90%, 95%, and 99% confidence intervals for parameters estimated in the $RPGJSB2_q$ model for different quantiles (log link and G the cdf of the normal model).

Table 5. The estimated parameters and standard errors (s.e.) for different quantiles in the $RPGJSB2_{q=0.5}$ model of the COVID-19 data set with G the cdf of the normal model and loglog link. The p -values for the traditional normality test for the RQRs are also presented.

| q | Parameter | Estimated | s.e. | z-Value | p -Value | p -Values for Quantile Residuals | | | |
|-----|-----------|-----------|--------|---------|------------|------------------------------------|-------|-------|-------|
| | | | | | | KS | SW | AD | CVM |
| 0.5 | β_0 | -1.8396 | 0.1136 | -16.19 | <0.0001 | 0.626 | 0.249 | 0.133 | 0.094 |
| | β_1 | 0.0140 | 0.0105 | 1.34 | 0.0899 | | | | |
| | β_2 | 0.0159 | 0.0079 | 2.01 | 0.0223 | | | | |
| | β_3 | 0.1032 | 0.0308 | 3.35 | 0.0004 | | | | |
| | β_4 | 0.0575 | 0.0272 | 2.12 | 0.0171 | | | | |
| | ν_0 | 1.9051 | 0.0862 | 22.11 | <0.0001 | | | | |
| | ν_1 | 0.1268 | 0.1587 | 0.80 | 0.2121 | | | | |
| | ν_2 | 0.2028 | 0.1445 | 1.40 | 0.0802 | | | | |

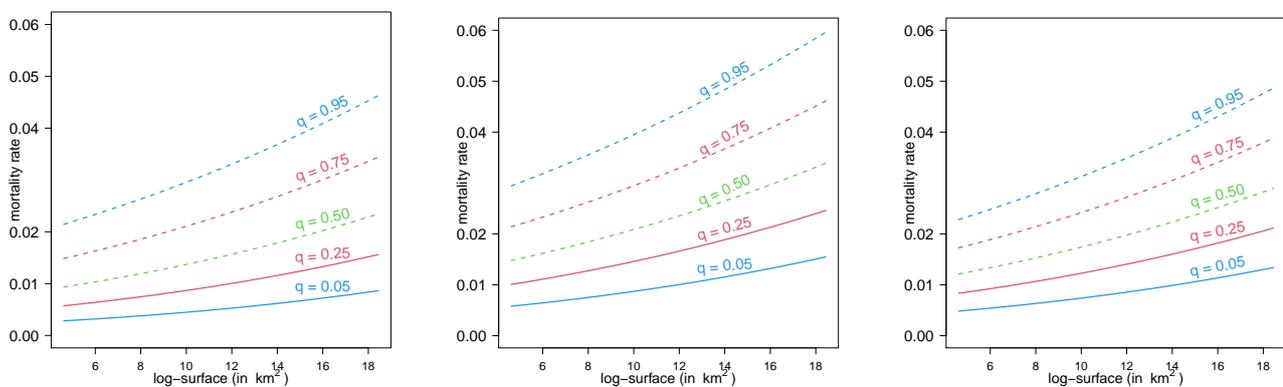


Figure 4. The estimated $100 \times q$ th quantile in the $RPGJSB2_q$ model for $\log(\text{population}) = 16$ (around 9 million people) varying the $\log(\text{surface})$ for countries in Africa, Asia, or Oceania (left panel), the Americas (center panel), and Europe (right panel) considering the cloglog link and G the cdf from the logistic model.

5.1.2. Local Influence Analysis

We performed a local influence analysis for the selected model under the three perturbation schemes discussed in Section 3.3. Figure 5 shows the analysis for the $RPGJSB2$ model with $q = 0.5$ using the loglog link and G the cdf of the normal model in the COVID-19 data set. In Figures A1–A4 in Appendix B.3, the same analysis for the other selected quantiles is presented. Note that, considering all the cases, the observation 100 appears in at least one case, namely Mexico (the Americas). Mexico reported the highest mortality rate (9.2%, with 221,647 accumulated deaths and 2,396,604 accumulated cases, respectively). Evidently there is a problem in managing the pandemic situation in this country as compared to the rest of the world. Table 6 presents the relative change for the parameters (RC), for their estimated standard errors (RCSE) and the respective p -values, for the estimation without Mexico. We highlight that the significance of the parameters related to the quantile was maintained for all the cases (except $\beta_1(q)$), suggesting that the inferences drawn from the model are robust to estimate the different quantiles in this problem.

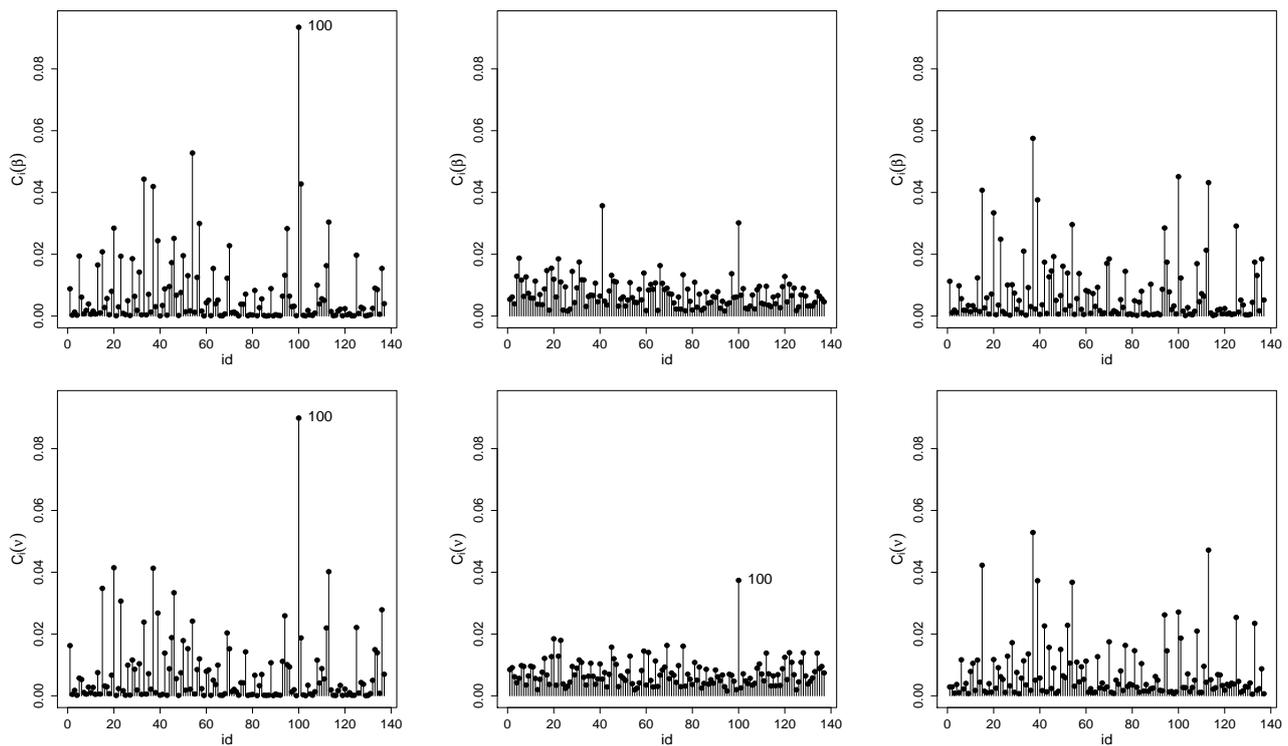


Figure 5. The index plots of C_i for $\hat{\beta}$ (upper) and \hat{v} (lower) under the weight perturbation (left), response perturbation (center), and covariate perturbation (right) schemes for the RPGJSB2 $_{q=0.5}$ model (cloglog link and G the cdf from the logistic model) of the COVID-19 data set.

Table 6. The relative changes (in %) in the ML estimates (RC) and their corresponding standard errors (RCSE) for the indicated parameter and respective p -values for the COVID-19 data set, when observation 100 is dropped.

| | | q | | | | |
|------------|--------------|---------|---------|---------|---------|---------|
| Parameter | | 0.10 | 0.25 | 0.50 | 0.75 | 0.90 |
| RC | $\beta_0(q)$ | 239.95 | 243.30 | 253.85 | 267.69 | 279.66 |
| RCSE | | 18.10 | 18.71 | 19.62 | 20.63 | 21.45 |
| p -value | | <0.0001 | <0.0001 | <0.0001 | <0.0001 | <0.0001 |
| RC | $\beta_1(q)$ | 298.10 | 295.33 | 296.05 | 292.45 | 276.05 |
| RCSE | | 224.15 | 228.55 | 233.01 | 233.17 | 223.76 |
| p -value | | 0.1751 | 0.1879 | 0.1979 | 0.2067 | 0.2151 |
| RC | $\beta_2(q)$ | 313.98 | 310.46 | 305.38 | 300.83 | 300.56 |
| RCSE | | 161.89 | 158.80 | 155.11 | 150.94 | 147.41 |
| p -value | | 0.0508 | 0.0491 | 0.0479 | 0.0469 | 0.046 |
| RC | $\beta_3(q)$ | 398.39 | 402.09 | 431.38 | 472.57 | 475.72 |
| RCSE | | 104.18 | 103.37 | 108.53 | 113.21 | 109.20 |
| p -value | | 0.0002 | 0.0001 | 0.0001 | 0.0001 | 0.0002 |
| RC | $\beta_4(q)$ | 406.56 | 446.90 | 569.98 | 895.37 | 2014.00 |
| RCSE | | 139.08 | 150.33 | 184.91 | 270.23 | 566.60 |
| p -value | | 0.0041 | 0.0037 | 0.0039 | 0.0044 | 0.0054 |
| RC | $v_0(q)$ | 96.05 | 92.62 | 90.77 | 89.93 | 89.73 |
| RCSE | | 0.02 | 0.01 | 0.06 | 0.13 | 0.21 |
| p -value | | 0.4717 | 0.1467 | 0.0439 | 0.0120 | 0.0029 |
| RC | $v_1(q)$ | 112.14 | 113.02 | 116.81 | 138.60 | 224.79 |
| RCSE | | 0.41 | 0.90 | 1.91 | 3.98 | 8.62 |
| p -value | | 0.1048 | 0.0950 | 0.0879 | 0.0820 | 0.0767 |
| RC | $v_2(q)$ | 15.21 | 11.87 | 6.52 | 0.32 | 4.46 |
| RCSE | | 0.18 | 0.17 | 0.92 | 2.27 | 4.21 |
| p -value | | 0.1493 | 0.1440 | 0.1399 | 0.1362 | 0.1329 |

6. Conclusions

In this paper, we proposed two classes of parametric quantile regression models for studying the association between a bounded response and covariates by inferring the conditional quantile of the response. The proposed quantile regression model was based on power Johnson SB distribution [11] using a new parameterization of this distribution that was indexed by quantile, dispersion, and shape parameters (RPGJSB1_q(ψ, δ, α)) or quantile and dispersion parameters (RPGJSB2_q(ψ, δ)). The first proposed quantile model had an extra-parameter $\alpha > 0$ associated with the “tailedness”; the second proposed quantile model had lower computing costs. The ML inference was implemented to estimate the model parameters, which were satisfactory considering the simulation studies, where parameters were recovered for different sample sizes. Furthermore, we developed model diagnostic tools for each proposed quantile regression model.

There are several extensions of the new models not considered in this work that can be addressed in future research; in particular, an extension of the methods developed in this paper would be to consider in Equation (2) a much more general family of distributions; that is, to consider models for zero-inflated and one-inflated data set. Directions related to random effects in the model also can be addressed in future works.

Author Contributions: Conceptualization, D.I.G. and M.B.; methodology, D.I.G., M.B., and Y.M.G.; software, D.I.G. and Y.M.G.; validation, C.C.-C. and O.V.; data curation, C.C.-C.; writing—original draft preparation, D.I.G. and M.B.; writing—review and editing, Y.M.G., C.C.-C., and O.V.; funding acquisition, O.V. All authors have read and agreed to the published version of the manuscript.

Funding: Caamaño-Carrillo’s research was funded by FONDECYT (Chile) grant No. 11220066 and by Proyecto Regular DIUBB 2120538 IF/R de la Universidad del Bío-Bío.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The COVID-19 data set was taken from the World Health Organization [25].

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

| | |
|---------------------|---|
| AD | Anderson–Darling (test) |
| AIC | Akaike information criterion |
| BIC | Bayesian information criterion |
| cdf | cumulative distribution function |
| CP | coverage probabilities |
| CVM | Cramér–Von-Mises (test) |
| GJS | generalized Johnson S_B distribution |
| KS | Kolmogorov–Smirnov (test) |
| LD | likelihood displacement |
| ML | maximum likelihood |
| pdf | probability distribution function |
| PJSB | power Johnson S_B distribution |
| PGJSB | power generalized Johnson S_B distribution |
| PGJSB1 _q | model 1 with reparametrization power generalized Johnson S_B distribution |
| PGJSB2 _q | model 2 with reparametrization power generalized Johnson S_B distribution |
| RQRs | randomized quantile residuals |
| SW | Shapiro–Wilks (test) |

Appendix A. Details of Local Influence

In these Appendices, we provide details of the matrices involved in the local influence for the RPGJSB1_q and RPGJSB2_q models. Furthermore, we present more results of the COVID-19 mortality rate discussed in Section 5.

In Appendix A, we present details of the matrices related to the different perturbation schemes discussed in Section 3.3.

Appendix A.1. Perturbation of the Response

The *i*th element of matrices **D**₁₀ and **D**₁₁ for model RPGJSB1_q are given by

$$\begin{aligned} \ddot{d}_{\psi,w} &= \delta_i^2 y_i \dot{Q}_{\psi_i}(\psi_i) \dot{Q}_{w_i}(0) \left(\frac{(\alpha - 1)[\dot{G}_{\psi}(\tau(\delta, \psi)_{1i})^2 - G(\tau(\delta, \psi)_{1i})\ddot{G}_{\psi,w_i}(\tau(\delta, \psi)_{1i})]}{G(\tau(\delta, \psi)_{1i})^2} + \frac{\dot{g}_{\psi}(\tau(\delta, \psi)_{1i})^2}{g(\tau(\delta, \psi)_{1i})^2} \right. \\ &\quad \left. - \frac{\ddot{g}_{\psi,w_i}(\tau(\delta, \psi)_{1i})}{g(\tau(\delta, \psi)_{1i})} \right) \\ \ddot{d}_{\delta,w} &= y_i \dot{Q}_{\delta_i}(0) \left(\frac{\dot{g}_{\delta_i}(\tau(\delta, \psi)_{1i})}{g(\tau(\delta, \psi)_{1i})} - \frac{\tau(\delta, \psi)_{1i} \dot{g}_{\delta_i}(\tau(\delta, \psi)_{1i})^2}{g(\tau(\delta, \psi)_{1i})^2} + \frac{(\alpha - 1)\dot{G}_{\delta_i}(\tau(\delta, \psi)_{1i})}{G(\tau(\delta, \psi)_{1i})} \right. \\ &\quad \left. - \frac{(\alpha - 1)\tau(\delta, \psi)_{1i} \dot{G}_{\delta_i}(\tau(\delta, \psi)_{1i})^2}{G(\tau(\delta, \psi)_{1i})^2} + \frac{\tau(\delta, \psi)_{1i} \ddot{g}_{\delta_i,w_i}(\tau(\delta, \psi)_{1i})}{g(\tau(\delta, \psi)_{1i})} + \frac{(\alpha - 1)\tau(\delta, \psi)_{1i} \ddot{G}_{\delta_i,w_i}(\tau(\delta, \psi)_{1i})}{G(\tau(\delta, \psi)_{1i})} \right), \end{aligned}$$

respectively, with $\tau(\delta, \psi)_{1i} = \delta_i(Q(0) - Q(\psi_i))$. On the other hand, the *i*th element of matrices **D**₁₂, **D**₁₃, and **D**₁₄ for model RPGJSB2_q are given by

$$\begin{aligned} \ddot{d}_{\xi,w} &= \delta_i^2 y_i \dot{Q}_{\xi_i}(\xi_i) \dot{Q}_{w_i}(0) \left(\frac{(\alpha - 1)[\dot{G}_{\xi}(\tau(\xi, \delta, \alpha)_{2i})^2 - G(\tau(\xi, \delta, \alpha)_{2i})\ddot{G}_{\xi,w_i}(\tau(\xi, \delta, \alpha)_{2i})]}{G(\tau(\xi, \delta, \alpha)_{2i})^2} + \frac{\dot{g}_{\xi}(\tau(\xi, \delta, \alpha)_{2i})^2}{g(\tau(\xi, \delta, \alpha)_{2i})^2} \right. \\ &\quad \left. - \frac{\ddot{g}_{\xi,w_i}(\tau(\xi, \delta, \alpha)_{2i})}{g(\tau(\xi, \delta, \alpha)_{2i})} \right) \\ \ddot{d}_{\delta,w} &= y_i \dot{Q}_{\delta_i}(0) \left(\frac{\dot{g}_{\delta_i}(\tau(\xi, \delta, \alpha)_{2i})}{g(\tau(\xi, \delta, \alpha)_{2i})} - \frac{\tau(\xi, \delta)_{2i} \dot{g}_{\delta_i}(\tau(\xi, \delta, \alpha)_{2i})^2}{g(\tau(\xi, \delta, \alpha)_{2i})^2} + \frac{(\alpha - 1)\dot{G}_{\delta_i}(\tau(\xi, \delta, \alpha)_{2i})}{G(\tau(\xi, \delta, \alpha)_{2i})} \right. \\ &\quad \left. - \frac{(\alpha - 1)\tau(\xi, \delta)_{2i} \dot{G}_{\delta_i}(\tau(\xi, \delta, \alpha)_{2i})^2}{G(\tau(\xi, \delta, \alpha)_{2i})^2} + \frac{\tau(\xi, \delta)_{2i} \ddot{g}_{\delta_i,w_i}(\tau(\xi, \delta, \alpha)_{2i})}{g(\tau(\xi, \delta, \alpha)_{2i})} + \frac{(\alpha - 1)\tau(\xi, \delta)_{2i} \ddot{G}_{\delta_i,w_i}(\tau(\xi, \delta, \alpha)_{2i})}{G(\tau(\xi, \delta, \alpha)_{2i})} \right) \\ \ddot{d}_{\alpha,w} &= \delta_i y_i \dot{Q}_{\alpha}(0) \left(\frac{(1 - \alpha)\dot{G}_{\alpha}(\tau(\xi, \delta, \alpha)_{2i})^2 \dot{x}_{\alpha}}{G(\tau(\xi, \delta, \alpha)_{2i})^2} + \frac{\dot{x}_{\alpha}[g(\tau(\xi, \delta, \alpha)_{2i}) - \dot{g}_{\alpha}(\tau(\xi, \delta, \alpha)_{2i})^2 + \ddot{g}_{\alpha,w_i}(\tau(\xi, \delta, \alpha)_{2i})]}{g(\tau(\xi, \delta, \alpha)_{2i})^2} \right. \\ &\quad \left. + \frac{\dot{G}_{\alpha}(\tau(\xi, \delta, \alpha)_{2i}) + (\alpha - 1)\dot{x}_{\alpha} \ddot{G}_{\alpha,w_i}(\tau(\xi, \delta, \alpha)_{2i})}{G(\tau(\xi, \delta, \alpha)_{2i})} \right), \end{aligned}$$

respectively, with $\tau(\xi, \delta, \alpha)_{2i} = \delta_i(Q(0) - Q(\xi_i)) + x_q^*(\alpha)$, $\tau(\xi, \delta)_{2i} = \delta_i(Q(0) - Q(\xi_i))$, and $\dot{x}_{\alpha} = \partial x_q^*(\alpha) / \partial \alpha$.

Appendix A.2. Perturbation of the Predictor

The *i*th element of matrices **D**₁₅ and **D**₁₆ for model RPGJSB1_q are given by

$$\begin{aligned} \ddot{d}_{\beta,w} &= \left(\frac{(1 - \alpha)\dot{G}_{\beta_i}(Q(y_i))}{G(Q(y_i))} - \frac{\dot{g}_{\beta_i}(Q(y_i))}{g(Q(y_i))} \right) \\ \ddot{d}_{v,w} &= \left(1 + \frac{Q(y_i)\dot{g}_{v_i}(Q(y_i))}{g(Q(y_i))} + \frac{(\alpha - 1)Q(y_i)\dot{G}_{v_i}(Q(y_i))}{G(Q(y_i))} \right), \end{aligned}$$

respectively. Similarly, the i th element of matrices D_{17} , D_{18} , and D_{19} for model RPGJSB 2_q are given by

$$\begin{aligned} \ddot{d}_{\beta,w} &= \left(\frac{(1-\alpha)\dot{G}_{\beta_i}(Q(y_i) + x_q^*(\alpha))}{G(Q(y_i) + x_q^*(\alpha))} - \frac{\dot{g}_{\beta_i}(Q(y_i) + x_q^*(\alpha))}{g(Q(y_i) + x_q^*(\alpha))} \right) \\ \ddot{d}_{v,w} &= \left(1 + \frac{Q(y_i)\dot{g}_{v_i}(Q(y_i) + x_q^*(\alpha))}{g(Q(y_i) + x_q^*(\alpha))} + \frac{(\alpha-1)Q(y_i)\dot{G}_{v_i}(Q(y_i) + x_q^*(\alpha))}{G(Q(y_i) + x_q^*(\alpha))} \right) \\ \ddot{d}_{\alpha,w} &= (Q(y_i)v_{iz_{ij}} - \beta_ix_{ij}) \left(\frac{(\alpha-1)\dot{G}_{\alpha}(Q(y_i) + x_q^*(\alpha))^2\dot{x}_a}{G(Q(y_i) + x_q^*(\alpha))^2} + \frac{\dot{G}_{\alpha}(Q(y_i) + x_q^*(\alpha)) + (\alpha-1)\dot{x}_a\ddot{G}_{\alpha,w_i}(Q(y_i) + x_q^*(\alpha))}{G(Q(y_i) + x_q^*(\alpha))} \right) \\ &+ \frac{\dot{x}_a[g(Q(y_i) + x_q^*(\alpha))\ddot{g}_{\alpha,w_i}(Q(y_i) + x_q^*(\alpha)) - \dot{g}_{\alpha}(Q(y_i) + x_q^*(\alpha))^2]}{g(Q(y_i) + x_q^*(\alpha))^2}, \end{aligned}$$

respectively.

Appendix B. COVID-19 Data Set

In Appendix B, we present additional information for the COVID-19 data set analyzed by the RPGJSB 1_q and RPGJSB 2_q regression models.

Appendix B.1. AIC and BIC Criteria

Table A1. AIC and BIC criteria for the RPGJSB 1_q model of the COVID-19 data set considering 3 options for G (normal, logistic, and Cauchy) and the 4 discussed link functions.

| Criteria | q | Normal | | | | Logistic | | | | Cauchy | | | |
|----------|--------|--------|--------|--------|---------|----------|--------|--------|---------|--------|--------|--------|---------|
| | | Logit | Probit | Loglog | Cloglog | Logit | Probit | Loglog | Cloglog | Logit | Probit | Loglog | Cloglog |
| AIC | 0.05 | -871.9 | -871.7 | -871.5 | -871.9 | -871.4 | -871.3 | -871.3 | -871.3 | -840.3 | -841.5 | -842.4 | -840.1 |
| | 0.10 | -871.8 | -871.6 | -871.5 | -871.8 | -871.2 | -871.2 | -871.2 | -871.2 | -839.9 | -841.0 | -842.0 | -839.7 |
| | 0.15 | -871.7 | -871.6 | -871.5 | -871.8 | -871.2 | -871.2 | -871.1 | -871.2 | -839.7 | -840.8 | -841.8 | -839.5 |
| | 0.20 | -871.7 | -871.6 | -871.4 | -871.7 | -871.1 | -871.1 | -871.1 | -871.1 | -839.6 | -840.7 | -841.6 | -839.4 |
| | 0.25 | -871.6 | -871.5 | -871.4 | -871.7 | -871.1 | -871.1 | -871.1 | -871.1 | -839.5 | -840.6 | -841.6 | -839.3 |
| | 0.30 | -871.6 | -871.5 | -871.4 | -871.6 | -871.1 | -871.1 | -871.1 | -871.0 | -839.4 | -840.6 | -841.5 | -839.2 |
| | 0.35 | -871.5 | -871.5 | -871.4 | -871.6 | -871.0 | -871.0 | -871.0 | -871.0 | -839.4 | -840.5 | -841.5 | -839.2 |
| | 0.40 | -871.5 | -871.5 | -871.4 | -871.6 | -871.0 | -871.0 | -871.0 | -871.0 | -839.4 | -840.5 | -841.4 | -839.1 |
| | 0.45 | -871.5 | -871.4 | -871.3 | -871.5 | -871.0 | -871.0 | -871.0 | -870.9 | -839.3 | -840.5 | -841.4 | -839.1 |
| | 0.50 | -871.5 | -871.4 | -871.3 | -871.5 | -870.9 | -871.0 | -871.0 | -870.9 | -839.3 | -840.4 | -841.4 | -839.1 |
| | 0.55 | -871.4 | -871.4 | -871.3 | -871.5 | -870.9 | -870.9 | -870.9 | -870.9 | -839.3 | -840.4 | -841.4 | -839.1 |
| | 0.60 | -871.4 | -871.4 | -871.3 | -871.5 | -870.9 | -870.9 | -870.9 | -870.9 | -839.3 | -840.4 | -841.3 | -839.0 |
| | 0.65 | -871.4 | -871.4 | -871.3 | -871.4 | -870.9 | -870.9 | -870.9 | -870.8 | -839.2 | -840.4 | -841.3 | -839.0 |
| | 0.70 | -871.3 | -871.3 | -871.2 | -871.4 | -870.8 | -870.9 | -870.9 | -870.8 | -839.2 | -840.3 | -841.3 | -839.0 |
| | 0.75 | -871.3 | -871.3 | -871.2 | -871.4 | -870.8 | -870.8 | -870.9 | -870.8 | -839.2 | -840.3 | -841.2 | -838.9 |
| | 0.80 | -871.3 | -871.3 | -871.2 | -871.3 | -870.8 | -870.8 | -870.8 | -870.8 | -839.1 | -840.2 | -841.2 | -838.9 |
| | 0.85 | -871.2 | -871.3 | -871.2 | -871.3 | -870.7 | -870.8 | -870.8 | -870.7 | -839.1 | -840.2 | -841.1 | -838.8 |
| 0.90 | -871.2 | -871.2 | -871.2 | -871.3 | -870.7 | -870.7 | -870.7 | -870.7 | -839.0 | -840.1 | -841.1 | -838.8 | |
| 0.95 | -871.2 | -871.2 | -871.1 | -871.2 | -870.6 | -870.7 | -870.7 | -870.6 | -839.9 | -840.6 | -840.9 | -839.7 | |
| BIC | 0.05 | -839.7 | -839.6 | -839.4 | -839.8 | -839.2 | -839.2 | -839.1 | -839.2 | -808.2 | -809.4 | -810.3 | -808.0 |
| | 0.10 | -839.7 | -839.5 | -839.4 | -839.7 | -839.1 | -839.1 | -839.1 | -839.1 | -807.8 | -808.9 | -809.8 | -807.6 |
| | 0.15 | -839.6 | -839.5 | -839.3 | -839.7 | -839.1 | -839.0 | -839.0 | -839.0 | -807.6 | -808.7 | -809.6 | -807.4 |
| | 0.20 | -839.5 | -839.4 | -839.3 | -839.6 | -839.0 | -839.0 | -839.0 | -839.0 | -807.5 | -808.6 | -809.5 | -807.2 |
| | 0.25 | -839.5 | -839.4 | -839.3 | -839.6 | -839.0 | -839.0 | -839.0 | -838.9 | -807.4 | -808.5 | -809.5 | -807.2 |
| | 0.30 | -839.5 | -839.4 | -839.3 | -839.5 | -838.9 | -838.9 | -838.9 | -838.9 | -807.3 | -808.5 | -809.4 | -807.1 |
| | 0.35 | -839.4 | -839.4 | -839.3 | -839.5 | -838.9 | -838.9 | -838.9 | -838.9 | -807.3 | -808.4 | -809.4 | -807.1 |
| | 0.40 | -839.4 | -839.3 | -839.2 | -839.5 | -838.9 | -838.9 | -838.9 | -838.9 | -807.2 | -808.4 | -809.3 | -807.0 |
| | 0.45 | -839.4 | -839.3 | -839.2 | -839.4 | -838.8 | -838.9 | -838.9 | -838.8 | -807.2 | -808.4 | -809.3 | -807.0 |
| | 0.50 | -839.3 | -839.3 | -839.2 | -839.4 | -838.8 | -838.8 | -838.8 | -838.8 | -807.2 | -808.3 | -809.3 | -807.0 |
| | 0.55 | -839.3 | -839.3 | -839.2 | -839.4 | -838.8 | -838.8 | -838.8 | -838.8 | -807.2 | -808.3 | -809.2 | -806.9 |
| | 0.60 | -839.3 | -839.3 | -839.2 | -839.3 | -838.8 | -838.8 | -838.8 | -838.7 | -807.1 | -808.3 | -809.2 | -806.9 |
| | 0.65 | -839.3 | -839.2 | -839.2 | -839.3 | -838.7 | -838.8 | -838.8 | -838.8 | -807.1 | -808.2 | -809.2 | -806.9 |
| | 0.70 | -839.2 | -839.2 | -839.1 | -839.3 | -838.7 | -838.7 | -838.8 | -838.7 | -807.1 | -808.2 | -809.2 | -806.8 |
| | 0.75 | -839.2 | -839.2 | -839.1 | -839.2 | -838.7 | -838.7 | -838.7 | -838.7 | -807.0 | -808.2 | -809.1 | -806.8 |
| | 0.80 | -839.2 | -839.2 | -839.1 | -839.2 | -838.7 | -838.7 | -838.7 | -838.6 | -807.0 | -808.1 | -809.1 | -806.8 |
| | 0.85 | -839.1 | -839.1 | -839.1 | -839.2 | -838.6 | -838.6 | -838.7 | -838.6 | -806.9 | -808.1 | -809.0 | -806.7 |
| 0.90 | -839.1 | -839.1 | -839.0 | -839.1 | -838.6 | -838.6 | -838.6 | -838.6 | -806.9 | -808.0 | -808.9 | -806.7 | |
| 0.95 | -839.0 | -839.1 | -839.0 | -839.1 | -838.5 | -838.5 | -838.6 | -838.5 | -807.7 | -808.5 | -808.8 | -807.5 | |

Table A2. AIC and BIC criteria for the RPGJSB_{2q} model of the COVID-19 data set considering 3 options for G (normal, logistic, and Cauchy) and the 4 discussed link functions.

| Criteria | q | Normal | | | | Logistic | | | | Cauchy | | | |
|----------|--------|---------------|---------------|---------------|---------------|----------|--------|--------|---------|--------|--------|--------|---------|
| | | Logit | Probit | Loglog | Cloglog | Logit | Probit | Loglog | Cloglog | Logit | Probit | Loglog | Cloglog |
| AIC | 0.05 | -869.0 | -872.3 | -873.4 | -868.7 | -857.4 | -863.8 | -867.6 | -856.9 | -758.5 | -770.5 | -779.3 | -757.6 |
| | 0.10 | -869.8 | -872.7 | -873.5 | -869.6 | -860.0 | -865.7 | -869.1 | -859.6 | -779.2 | -789.0 | -796.4 | -778.4 |
| | 0.15 | -870.4 | -872.9 | -873.5 | -870.1 | -862.1 | -867.2 | -870.1 | -861.7 | -795.1 | -803.3 | -809.4 | -794.5 |
| | 0.20 | -870.8 | -873.1 | -873.4 | -870.6 | -864.0 | -868.5 | -871.0 | -863.6 | -808.5 | -815.2 | -820.1 | -808.0 |
| | 0.25 | -871.2 | -873.2 | -873.4 | -871.0 | -865.7 | -869.7 | -871.8 | -865.3 | -819.6 | -825.0 | -828.9 | -819.1 |
| | 0.30 | -871.6 | -873.3 | -873.3 | -871.4 | -867.2 | -870.7 | -872.3 | -866.9 | -828.4 | -832.6 | -835.7 | -828.0 |
| | 0.35 | -871.9 | -873.4 | -873.1 | -871.7 | -868.7 | -871.6 | -872.8 | -868.4 | -835.0 | -838.1 | -840.4 | -834.6 |
| | 0.40 | -872.2 | -873.4 | -873.0 | -872.0 | -870.0 | -872.3 | -873.0 | -869.8 | -839.3 | -841.4 | -843.0 | -839.0 |
| | 0.45 | -872.4 | -873.4 | -872.8 | -872.3 | -871.2 | -872.7 | -872.9 | -871.0 | -841.2 | -842.5 | -843.3 | -841.0 |
| | 0.50 | -872.6 | -873.4 | -872.5 | -872.5 | -872.1 | -872.9 | -872.6 | -872.0 | -840.6 | -841.0 | -841.1 | -840.4 |
| | 0.55 | -872.9 | -873.3 | -872.2 | -872.8 | -872.7 | -872.8 | -871.8 | -872.7 | -837.1 | -836.6 | -836.0 | -837.0 |
| | 0.60 | -873.0 | -873.2 | -871.8 | -873.0 | -872.9 | -872.0 | -870.4 | -872.9 | -830.1 | -828.7 | -827.5 | -830.1 |
| | 0.65 | -873.2 | -873.0 | -871.3 | -873.2 | -872.4 | -870.5 | -868.1 | -872.4 | -818.9 | -816.7 | -814.8 | -818.9 |
| | 0.70 | -873.3 | -872.7 | -870.8 | -873.3 | -870.8 | -867.9 | -864.6 | -871.0 | -802.3 | -799.2 | -796.6 | -802.4 |
| | 0.75 | -873.3 | -872.3 | -870.0 | -873.4 | -867.7 | -863.5 | -859.3 | -868.0 | -778.5 | -774.5 | -771.2 | -778.6 |
| | 0.80 | -873.2 | -871.7 | -868.9 | -873.3 | -862.2 | -856.5 | -851.1 | -862.6 | -744.2 | -739.3 | -735.4 | -744.3 |
| | 0.85 | -873.0 | -870.8 | -867.5 | -873.1 | -852.8 | -845.3 | -838.5 | -853.3 | -693.1 | -687.5 | -682.9 | -693.4 |
| | 0.90 | -872.3 | -869.3 | -865.2 | -872.5 | -836.9 | -827.1 | -818.5 | -837.6 | -611.5 | -605.2 | -600.1 | -611.8 |
| 0.95 | -870.7 | -866.3 | -861.1 | -871.1 | -810.5 | -797.4 | -786.0 | -811.6 | -453.7 | -446.8 | -431.4 | -454.1 | |
| BIC | 0.05 | -839.8 | -843.1 | -844.2 | -839.5 | -828.2 | -834.6 | -838.4 | -827.7 | -729.3 | -741.3 | -750.1 | -728.4 |
| | 0.10 | -840.6 | -843.5 | -844.3 | -840.4 | -830.8 | -836.5 | -839.9 | -830.4 | -750.0 | -759.8 | -767.2 | -749.2 |
| | 0.15 | -841.2 | -843.7 | -844.3 | -840.9 | -832.9 | -838.0 | -840.9 | -832.5 | -765.9 | -774.1 | -780.2 | -765.3 |
| | 0.20 | -841.6 | -843.9 | -844.2 | -841.4 | -834.8 | -839.3 | -841.8 | -834.4 | -779.3 | -786.0 | -790.9 | -778.8 |
| | 0.25 | -842.0 | -844.0 | -844.2 | -841.8 | -836.5 | -840.5 | -842.6 | -836.1 | -790.4 | -795.8 | -799.7 | -789.9 |
| | 0.30 | -842.4 | -844.1 | -844.1 | -842.2 | -838.0 | -841.5 | -843.1 | -837.7 | -799.2 | -803.4 | -806.5 | -798.8 |
| | 0.35 | -842.7 | -844.2 | -843.9 | -842.5 | -839.5 | -842.4 | -843.6 | -839.2 | -805.8 | -808.9 | -811.2 | -805.4 |
| | 0.40 | -843.0 | -844.2 | -843.8 | -842.8 | -840.8 | -843.1 | -843.8 | -840.6 | -810.1 | -812.2 | -813.8 | -809.8 |
| | 0.45 | -843.2 | -844.2 | -843.6 | -843.1 | -842.0 | -843.5 | -843.7 | -841.8 | -812.0 | -813.3 | -814.1 | -811.8 |
| | 0.50 | -843.4 | -844.2 | -843.3 | -843.3 | -842.9 | -843.7 | -843.4 | -842.8 | -811.4 | -811.8 | -811.9 | -811.2 |
| | 0.55 | -843.7 | -844.1 | -843.0 | -843.6 | -843.5 | -843.6 | -842.6 | -843.5 | -807.9 | -807.4 | -806.8 | -807.8 |
| | 0.60 | -843.8 | -844.0 | -842.6 | -843.8 | -843.7 | -842.8 | -841.2 | -843.7 | -800.9 | -799.5 | -798.3 | -800.9 |
| | 0.65 | -844.0 | -843.8 | -842.1 | -844.0 | -843.2 | -841.3 | -838.9 | -843.2 | -789.7 | -787.5 | -785.6 | -789.7 |
| | 0.70 | -844.1 | -843.5 | -841.6 | -844.1 | -841.6 | -838.7 | -835.4 | -841.8 | -773.1 | -770.0 | -767.4 | -773.2 |
| | 0.75 | -844.1 | -843.1 | -840.8 | -844.2 | -838.5 | -834.3 | -830.1 | -838.8 | -749.3 | -745.3 | -742.0 | -749.4 |
| | 0.80 | -844.0 | -842.5 | -839.7 | -844.1 | -833.0 | -827.3 | -821.9 | -833.4 | -715.0 | -710.1 | -706.2 | -715.1 |
| | 0.85 | -843.8 | -841.6 | -838.3 | -843.9 | -823.6 | -816.1 | -809.3 | -824.1 | -663.9 | -658.3 | -653.7 | -664.2 |
| | 0.90 | -843.1 | -840.1 | -836.0 | -843.3 | -807.7 | -797.9 | -789.3 | -808.4 | -582.3 | -576.0 | -570.9 | -582.6 |
| 0.95 | -841.5 | -837.1 | -831.9 | -841.9 | -781.3 | -768.2 | -756.8 | -782.4 | -424.5 | -417.6 | -402.2 | -424.9 | |

Appendix B.2. Estimated Parameters

Table A3. The estimated parameters for different quantiles in the RPGJSB_{2q} model of the COVID-19 data set, with the normal distribution for G and the loglog link. Furthermore, we present the p-values for the traditional normality test for the randomized quantile residuals.

| q | Parameter | Estimated | s.e. | t-Value | p-Value | p-Values for Quantile Residuals | | | |
|------|-----------|-----------|--------|---------|---------|---------------------------------|-------|-------|-------|
| | | | | | | KS | SW | AD | CVM |
| 0.1 | β_0 | -2.0810 | 0.1154 | -18.03 | <0.0001 | 0.441 | 0.260 | 0.128 | 0.099 |
| | β_1 | 0.0150 | 0.0102 | 1.46 | 0.0716 | | | | |
| | β_2 | 0.0153 | 0.0077 | 1.99 | 0.0234 | | | | |
| | β_3 | 0.1310 | 0.0423 | 3.10 | 0.0010 | | | | |
| | β_4 | 0.0967 | 0.0371 | 2.61 | 0.0045 | | | | |
| | ν_0 | 1.5341 | 0.0881 | 17.41 | <0.0001 | | | | |
| 0.25 | β_0 | -1.9443 | 0.1140 | -17.06 | <0.0001 | 0.573 | 0.324 | 0.164 | 0.119 |
| | β_1 | 0.0144 | 0.0103 | 1.40 | 0.0814 | | | | |
| | β_2 | 0.0156 | 0.0078 | 2.01 | 0.0223 | | | | |
| | β_3 | 0.1158 | 0.0331 | 3.50 | 0.0002 | | | | |
| | β_4 | 0.0756 | 0.0290 | 2.61 | 0.0046 | | | | |
| | ν_0 | 1.7178 | 0.0873 | 19.69 | <0.0001 | | | | |
| 0.5 | ν_1 | 0.1263 | 0.1600 | 0.79 | 0.2149 | | | | |
| | ν_2 | 0.1914 | 0.1462 | 1.31 | 0.0953 | | | | |

Table A3. Cont.

| q | Parameter | Estimated | s.e. | t -Value | p -Value | p -Values for Quantile Residuals | | | |
|---------|-----------|-----------|--------|------------|------------|------------------------------------|-------|-------|-------|
| | | | | | | KS | SW | AD | CVM |
| 0.75 | β_0 | -1.7334 | 0.1144 | -15.16 | <0.0001 | 0.396 | 0.070 | 0.048 | 0.035 |
| | β_1 | 0.0139 | 0.0108 | 1.29 | 0.0989 | | | | |
| | β_2 | 0.0161 | 0.0082 | 1.97 | 0.0244 | | | | |
| | β_3 | 0.0909 | 0.0330 | 2.76 | 0.0029 | | | | |
| | β_4 | 0.0364 | 0.0289 | 1.26 | 0.1041 | | | | |
| | ν_0 | 2.1731 | 0.0843 | 25.78 | <0.0001 | | | | |
| | ν_1 | 0.1174 | 0.1564 | 0.75 | 0.2264 | | | | |
| ν_2 | 0.2170 | 0.1412 | 1.54 | 0.0622 | | | | | |
| 0.9 | β_0 | -1.6473 | 0.1163 | -14.16 | <0.0001 | 0.169 | 0.005 | 0.006 | 0.005 |
| | β_1 | 0.0142 | 0.0113 | 1.26 | 0.1044 | | | | |
| | β_2 | 0.0161 | 0.0086 | 1.87 | 0.0304 | | | | |
| | β_3 | 0.0860 | 0.0367 | 2.34 | 0.0096 | | | | |
| | β_4 | 0.0162 | 0.0311 | 0.52 | 0.3018 | | | | |
| | ν_0 | 2.5206 | 0.0818 | 30.81 | <0.0001 | | | | |
| | ν_1 | 0.0877 | 0.1534 | 0.57 | 0.2837 | | | | |
| ν_2 | 0.2296 | 0.1363 | 1.68 | 0.0460 | | | | | |

Appendix B.3. Additional Information for Local Influence

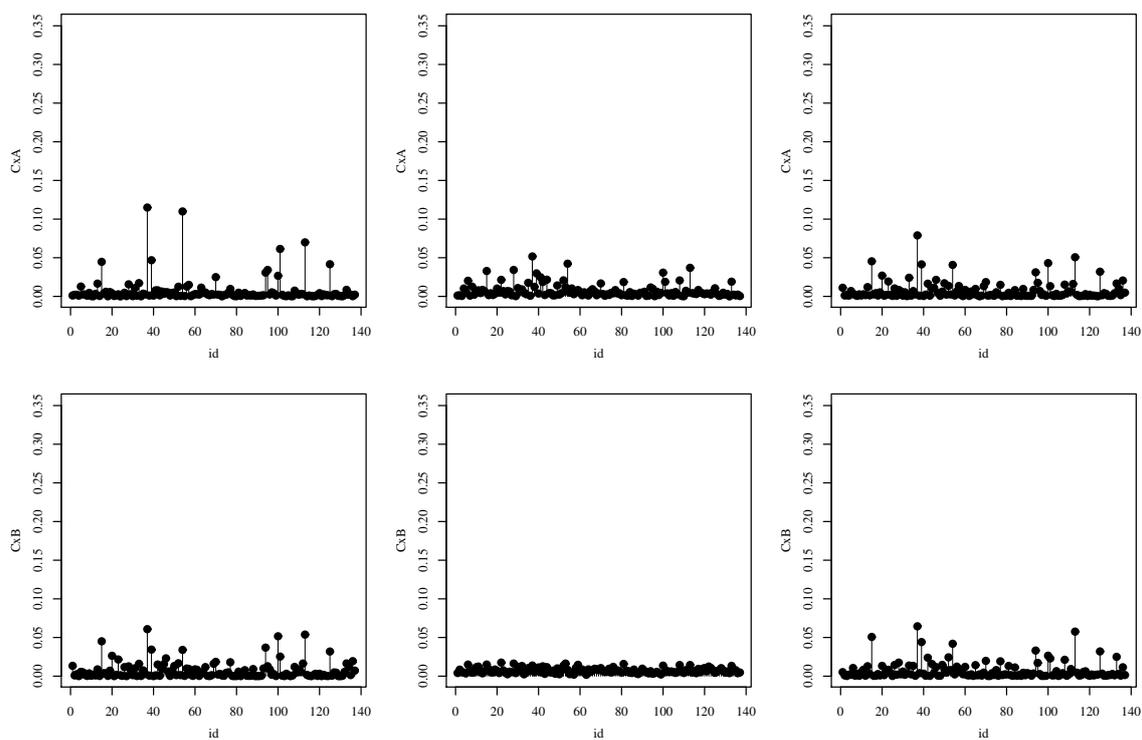


Figure A1. The index plots of C_i for $\hat{\beta}$ (upper) and $\hat{\nu}$ (lower) under the weight perturbation (left), response perturbation (center), and covariate perturbation (right) schemes for the RPGJSB $_q$ model with $q = 0.1$ (link loglog and G the cdf from the normal model) of the COVID-19 data set.

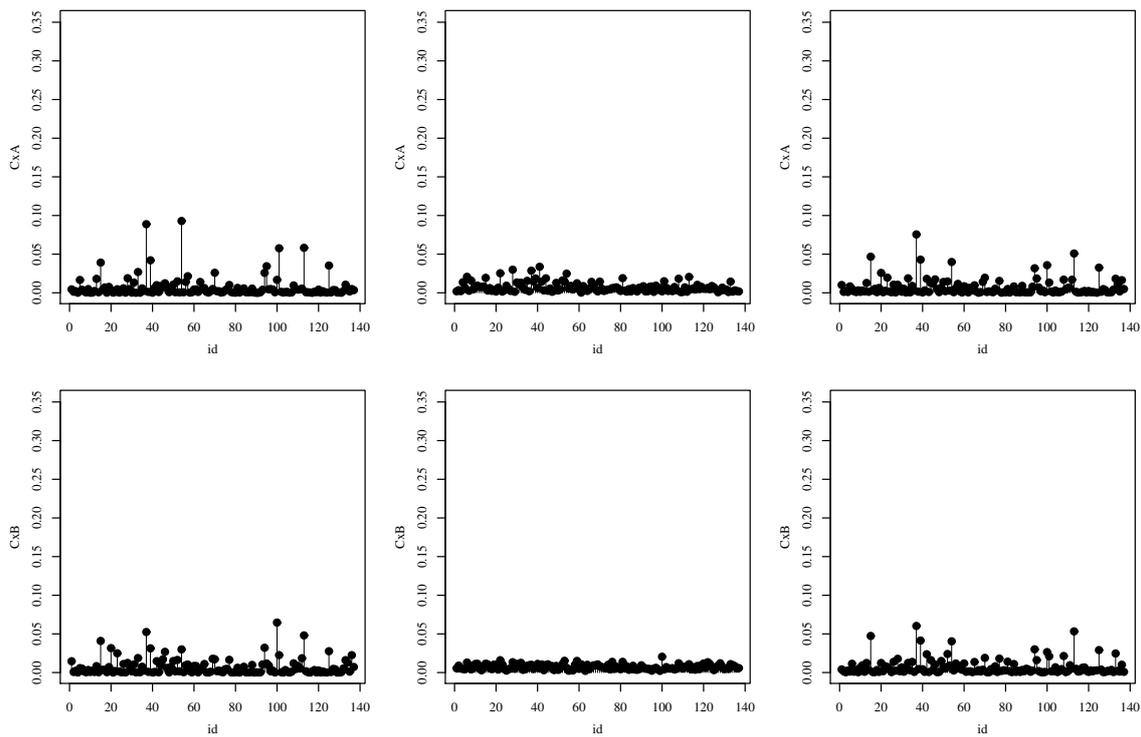


Figure A2. The index plots of C_i for $\hat{\beta}$ (upper) and \hat{v} (lower) under the weight perturbation (left), response perturbation (center), and covariate perturbation (right) schemes for the RPGJSB $_q$ model with $q = 0.25$ (link loglog and G the cdf from the normal model) of the COVID-19 data set.

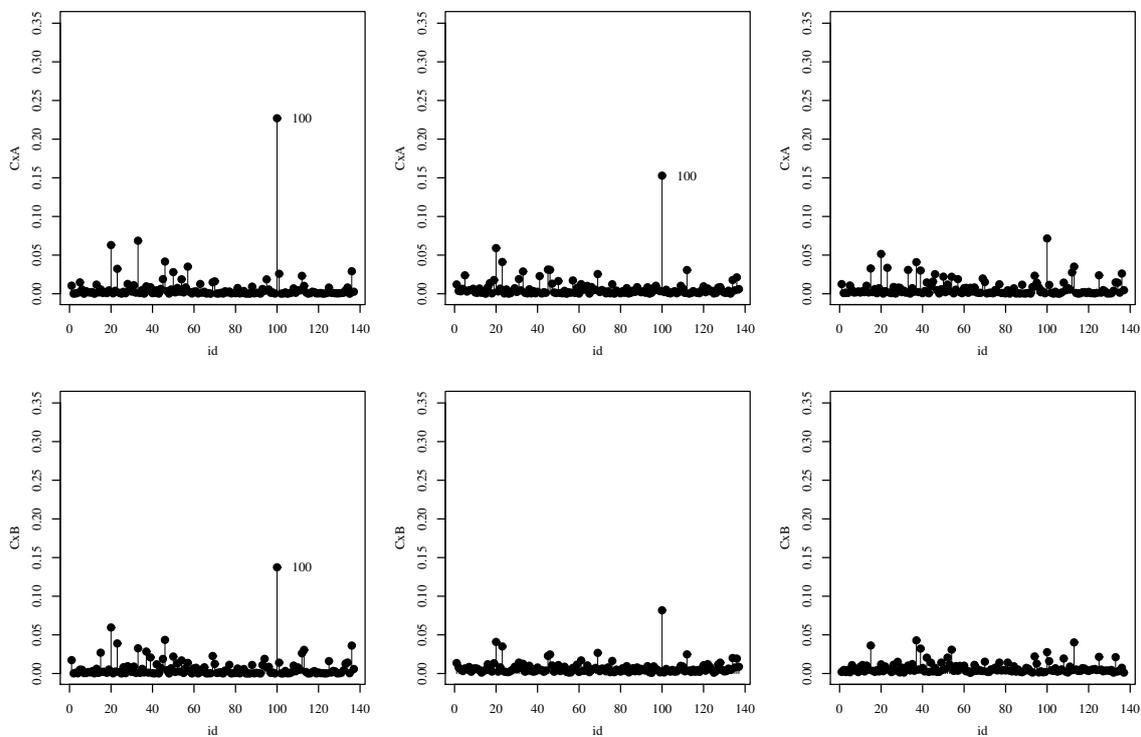


Figure A3. The index plots of C_i for $\hat{\beta}$ (upper) and \hat{v} (lower) under the weight perturbation (left), response perturbation (center), and covariate perturbation (right) schemes for the RPGJSB $_q$ model with $q = 0.75$ (link loglog and G the cdf from the normal model) of the COVID-19 data set.

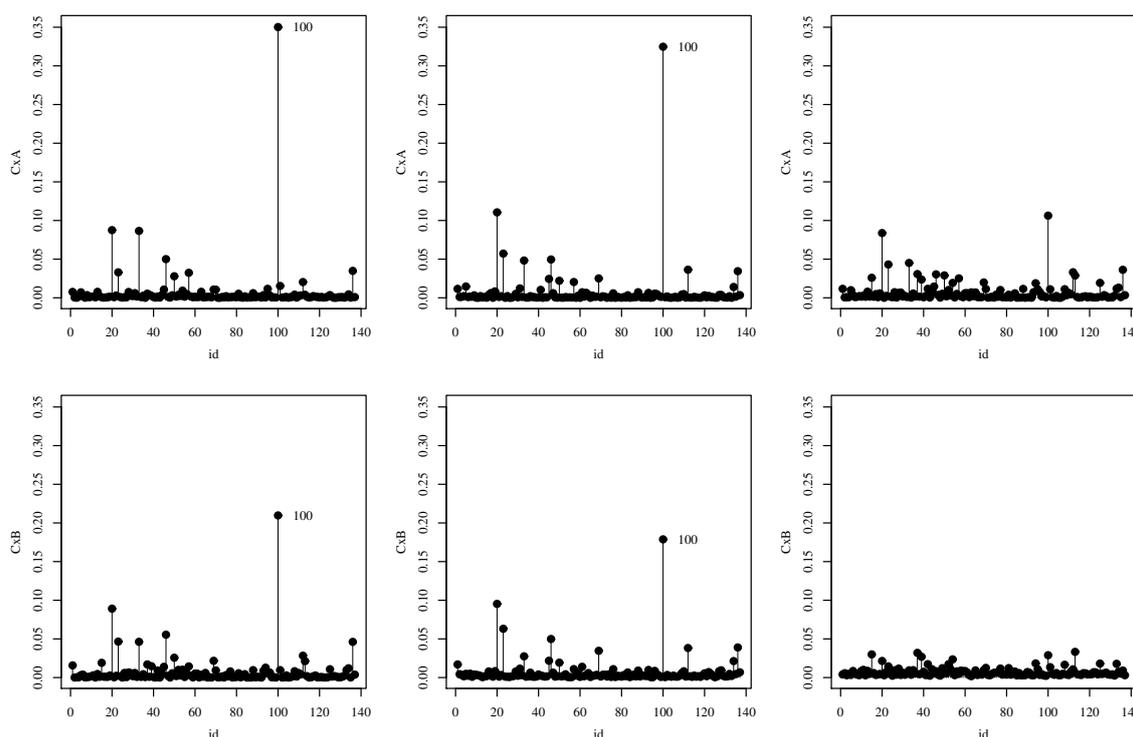


Figure A4. The index plots of C_i for $\hat{\beta}$ (upper) and \hat{v} (lower) under the weight perturbation (left), response perturbation (center), and covariate perturbation (right) schemes for the RPGJSB $_q$ model with $q = 0.9$ (link loglog and G the cdf from the normal model) of the COVID-19 data set.

References

- Ferrari, S.L.; Cribari-Neto, F. Beta regression for modelling rates and proportions. *J. Appl. Stat.* **2004**, *31*, 799–815. [CrossRef]
- Ospina, R.; Ferrari, S.L.P. Inflated beta distributions. *Stat. Pap.* **2008**, *51*, 111–126. [CrossRef]
- Bayes, C.L.; Bazán, J.L.; García, C. A new robust regression model for proportions. *Bayesian Anal.* **2012**, *7*, 841–866. [CrossRef]
- Migliorati, S.; Di Brisco, A.M.; Ongaro, A. A New Regression Model for Bounded Responses. *Bayesian Anal.* **2018**, *13*, 845–872. [CrossRef]
- Koenker, R.; Bassett, G. Regression quantiles. *Econometrica* **1978**, *46*, 33–50. [CrossRef]
- Lemonte, A.J.; Moreno-Arenas, G. On a heavy-tailed parametric quantile regression model for limited range response variables. *Comput. Stat.* **2020**, *35*, 379–398. [CrossRef]
- Su, S. Flexible parametric quantile regression model. *Stat. Comput.* **2015**, *25*, 635–650. [CrossRef]
- Mazucheli, J.; Menezes, A.F.B.; Fernandes, L.B.; Oliveira, R.P.; Ghitany, M.E. The unit-Weibull distribution as an alternative to the Kumaraswamy distribution for the modelling of quantiles conditional on covariates. *J. Appl. Stat.* **2020**, *47*, 954–974. [CrossRef]
- Lemonte, A.J.; Bazán, J.L. New class of Johnson S_B distributions and its associated regression model for rates and proportions. *Biom. J.* **2016**, *58*, 727–746. [CrossRef] [PubMed]
- Johnson, N.L. Systems of frequency curves generated by the methods of translation. *Biometrika* **1949**, *36*, 149–176. [CrossRef]
- Cancho, V.G.; Bazán, J.L.; Dey, D.K. A new class of regression model for a bounded response with application in the study of the incidence rate of colorectal cancer. *Stat. Methods Med. Res.* **2020**, *29*, 2015–2033. [CrossRef] [PubMed]
- Bayes, C.L.; Bazán, J.L.; De Castro, M. A quantile parametric mixed regression model for bounded response variables. *Stat. Interface* **2017**, *10*, 483–493. [CrossRef]
- Durrans, S.R. Distributions of fractional order statistics in hydrology. *Water Resour. Res.* **1992**, *28*, 1649–1655. [CrossRef]
- Lehmann, E.L. The power of rank tests. *Ann. Math. Stat.* **1953**, *24*, 23–43. [CrossRef]
- Fletcher, R. *Practical Methods of Optimization*, 2nd ed.; John Wiley & Sons.: New York, NY, USA, 1987.
- Cox, D.; Hinkley, D. *Theoretical Statistics*; Chapman and Hall: London, UK, 1974.
- Dunn, P.K.; Smyth, G.K. Randomized quantile residuals. *J. Comput. Graph. Stat.* **1996**, *5*, 236–244.
- Yap, B.W.; Sim, C.H. Comparisons of various normality tests. *J. Stat. Comput. Simul.* **2011**, *81*, 2141–2155. [CrossRef]
- Cook, R.D. Assessment of Local Influence. *J. R. Soc. B Stat. Methodol.* **1986**, *48*, 133–155. [CrossRef]
- R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2021. Available online: <https://www.R-project.org/> (accessed on 10 January 2022).

21. Du, R.H.; Liang, L.R.; Yang, C.Q.; Wang, W.; Cao, T.Z.; Li, M.; Guo, G.Y.; Du, J.; Zheng, C.L.; Zhu, Q.; et al. Predictors of mortality for patients with COVID-19 pneumonia caused by SARS-CoV-2: A prospective cohort study. *Eur. Respir. J.* **2020**, *55*, 2000524. [[CrossRef](#)]
22. Ji, J.S.; Liu, Y.; Liu, R.; Zha, Y.; Chang, X.; Zhang, L.; Zhang, Y.; Zeng, J.; Dong, T.; Xu, X.; et al. Survival analysis of hospital length of stay of novel coronavirus (COVID-19) pneumonia patients in Sichuan, China. *medRxiv* **2020**, doi: [[CrossRef](#)]
23. Li, X.; Xu, S.; Yu, M.; Wang, K.; Tao, Y.; Zhou, Y.; Shi, J.; Zhou, M.; Wu, B.; Yang, Z. Risk factors for severity and mortality in adult COVID-19 inpatients in Wuhan. *J. Allergy Clin. Immunol.* **2020**, *146*, 110–118. [[CrossRef](#)]
24. Livingston, E.; Bucher, K. Coronavirus disease 2019 (COVID-19) in Italy. *J. Am. Med Assoc.* **2020**, *323*, 1335. [[CrossRef](#)] [[PubMed](#)]
25. WHO. *Coronavirus Disease (COVID-19) Dashboard*; World Health Organization: Geneva, Switzerland, 2021. Available online: <https://covid19.who.int/> (accessed on 25 May 2021).