UNIVERSIDADE ESTADUAL DE CAMPINAS

UNICAMP

IMECC 50 ANOS

16ª EMR - Escola de Modelos de Regressão

Estimation and diagnostics for partially linear censored regression models based on heavy-tailed distributions

Larissa Avila Matos

Co-Authors: Marcela N. Lemus, Victor H. Lachos and Christian E. Galarza

FAPESP

Março/2019

Contents

Summary

Introduction

Motivation

- In many studies, limited or censored data are collected. This occurs, in several practical situations, for reasons such as limitations of measuring instruments or due to experimental design. So, the responses can be either left, interval or right censored.

- Nowadays, many data problems have structures whose approach goes beyond simple linear regression. The semiparametric regression models are statistical models that allow the mean response of interest to be linearly dependent on some explanatory variables and in other variables it not.

Introduction

- **Goal:** Modeling censored responses in Partially linear regression (PLR) models, from a frequentist perspective.

- **Problem:** The random errors are routinely assumed to follow a normal distribution, however, if the random error distribution its tails are heavier than normal ones and there are presence of outliers, then might lead to biased estimates.

- **Proposal:** We propose a robust partial censored regression (PCR) models based on the scale mixtures of normal (SMN) distributions and we present some influence diagnostic techniques, based on case deletion and local influence approaches.

- **Alternative solutions:** Castro *et al.* (2014) advocated the use of the SMN class of distributions in PCR (SMN-PCR) models and adopted a Bayesian framework to carry out posterior inference.

Preliminaries

Scale mixtures of normal distributions (Andrews & Mallows, 1974)

Stochastic representation

$$Y \stackrel{d}{=} \mu + k(U)^{1/2} Z, \qquad (1)$$

where

- $\mu \in \mathcal{R}$ is location parameter and $\sigma^2 \in (0, \infty)$ is scale parameter;
- Z and U are independent random variables, $Z \sim \mathrm{N}(0, \sigma^2)$;
- $k(\cdot)$ is a positive weight function;
- U is a mixing positive random variable with pdf $h(\cdot|\boldsymbol{\nu})$ and cdf $H(\cdot|\boldsymbol{\nu})$;
- $\boldsymbol{\nu}$ is a scalar or vector parameter indexing the distribution of U;
- Notation: $Y \sim \mathrm{SMN}(\mu, \sigma^2, \boldsymbol{\nu})$.

Preliminaries

From (1), we have that $Y|k(U) = k(u) \sim \mathrm{N}(\mu, k(u)\sigma^2)$. So, the pdf of Y is given by

$$f_{SMN}(y|\mu, \sigma^2, \boldsymbol{\nu}) = (2\pi\sigma^2)^{-1/2} \int_0^\infty k(u)^{-1/2} \exp\left[-k^{-1}(u)\frac{(\mathbf{y}-\mu)^2}{2\sigma^2}\right] dH(u|\boldsymbol{\nu}).$$

Particular cases: For $k(u) = u^{-1}$, we have

- Normal: If $\mathrm{P}(U = 1) = 1 \implies Y \sim \mathrm{N}(\mu, \sigma^2)$.
- Student-t: If $U \sim \mathrm{Gamma}(\nu/2, \nu/2) \implies Y \sim \mathrm{T}(\mu, \sigma^2, \nu)$.
- Slash: If $U \sim \mathrm{Beta}(\nu, 1) \implies Y \sim \mathrm{SL}(\mu, \sigma^2, \nu)$.
- Contaminated Normal: If U is a discrete random variable $\implies$ $Y \sim \mathrm{CN}(\mu, \sigma^2, \boldsymbol{\nu})$.

Preliminaries

The ECME algorithm

Let $\boldsymbol{\theta}$ be the parameter vector and $\mathbf{z}$ be the vector of complete data, i.e., the observed data and the missing/censored data (or the latent variables, depending on the situation). The ECME algorithm consists of the following steps:

- **E-step**: Compute the conditional expectation conditioned to the observed data vector, $Q(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}}^{(k)}) = \mathrm{E}_{\boldsymbol{\theta}^{(k)}}[\ell_c(\boldsymbol{\theta}|\mathbf{z})|\mathbf{Y}_{\mathrm{obs}}, \widehat{\boldsymbol{\theta}}^{(k)}]$, where $\widehat{\boldsymbol{\theta}}^{(k)}$ is the estimate of $\boldsymbol{\theta}$ at the (k)-th iteration.

- **CM-step**: Calculate $\widehat{\boldsymbol{\theta}}^{(k+1)}$ by maximizing $Q(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}}^{(k)})$, with $\boldsymbol{\nu}$ fixed at $\widehat{\boldsymbol{\nu}}^{(k)}$, such that, $\widehat{\boldsymbol{\theta}}^{(k+1)} = \underset{\boldsymbol{\theta}\in\Theta}{\arg\max}\{Q(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}}^{(k)})\}$.

- **CML-step**: Chooses $\widehat{\boldsymbol{\nu}}^{(k+1)}$ to maximize the constrained actual marginal log-likelihood function $\ell(\boldsymbol{\theta})$, with $\boldsymbol{\theta}$ obtained in the $(k+1)$-th iteration $\longrightarrow$ $\widehat{\boldsymbol{\theta}}^{(k+1)}$.

Summary

SMN-PCR model

The model

Let us consider a partially linear model:

$$Y_i = \mathbf{x}_i^\top \boldsymbol{\beta} + f(t_i) + \varepsilon_i, \qquad (2)$$

where

- $\mathbf{x}_i = (x_{1i}, \ldots, x_{pi})^\top$ is a $p \times 1$ vector of explanatory variable values;
- $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_p)^\top$ is a $p \times 1$ vector of regression parameters;
- t_i is a scalar that may represent a value of a continuous variable;
- $f(\cdot)$ is the smoothing function; and
- $\varepsilon_i \overset{\text{iid.}}{\sim} \text{SMN}(0, \sigma^2, \boldsymbol{\nu}), \quad i = 1, \ldots, n.$

SMN-PCR model

Censored Response

Considering the left censored case, we have

$$Y_{\text{obs}_i} = \begin{cases} \kappa_i & \text{if} \quad Y_i \leq \kappa_i; \\ Y_i & \text{if} \quad Y_i > \kappa_i., \end{cases}$$

for some threshold point κ_i, $i = 1, \ldots, n$. Alternatively, the model (2) can be written as:

$$Y_i = \mathbf{x}_i^\top \boldsymbol{\beta} + \mathbf{n}_i^\top \mathbf{f} + \varepsilon_i,$$

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{N}\mathbf{f} + \boldsymbol{\varepsilon},$$

where

- $\mathbf{f} = (f(t_1^0), \ldots, f(t_r^0))^\top$ is a vector of dimension $r \times 1$;
- $\mathbf{n}_i$: is the incidence vector of dimension $r \times 1$, with the s-th element equal $I(t_i = t_s^0)$ for $s = 1, \ldots, r$; and
- $\mathbf{N}$: is the incidence matrix of dimension $n \times r$, with the (i, s)-th element equal $I(t_i = t_s^0)$ for $s = 1, \ldots, r$.

SMN-PCR model

The log-likelihood function

Let $\mathbf{Y}_{\text{obs}} = (y_1, \ldots, y_n)^\top$ be an observed sample of $\mathbf{Y} = (Y_1, \ldots, Y_n)^\top$, such that $\mathbf{Y}_{\text{obs}} = \{\kappa_1, \ldots, \kappa_m, y_{m+1}, \ldots, y_n\}$, where m are censored and $n - m$ uncensored. The log-likelihood function is:

$$\ell(\boldsymbol{\theta}) \quad = \quad \sum_{i=1}^{m} \log\left[F_{SMN}\left(\frac{\kappa_i - \mu_i}{\sigma}\right)\right] + \sum_{i=m+1}^{n} \log\left[f_{SMN}(y_i|\mu_i, \sigma^2, \boldsymbol{\nu})\right], \qquad (3)$$

where $\mu_i = \mathbf{x}_i^\top \boldsymbol{\beta} + \mathbf{n}_i^\top \mathbf{f}$.

The penalized log-likelihood function

$$\ell_p(\boldsymbol{\theta}, \alpha) \quad = \quad \ell(\boldsymbol{\theta}) - \frac{\alpha}{2} J(\mathbf{f}), \qquad (4)$$

where $J(\mathbf{f})$ denotes the penalty function over $\mathbf{f}(\cdot)$ and α is a smoothing parameter.

SMN-PCR model

Parameter estimation via ECME algorithm

Using the stochastic representation given in (1), the hierarchical representation for the model defined in (2) is

$$\begin{aligned} Y_i|U_i &= u_i \sim \mathrm{N}(\mu_i, u_i^{-1}\sigma^2), \qquad (5) \\ U_i &\sim \mathrm{H}(\cdot|\boldsymbol{\nu}). \end{aligned}$$

Consider the augmented dataset $\boldsymbol{z} = \{\kappa_1, \ldots, \kappa_m, y_{m+1}, \ldots, y_n, u_1, \ldots, u_n\}$, the complete-data penalized log-likelihood, is given by:

$$\ell_{cp}(\boldsymbol{\theta}|\boldsymbol{z}) = -\frac{n}{2}\log\sigma^2 + \frac{1}{2}\sum_{i=1}^{n}\log u_i - \frac{1}{2\sigma^2}\sum_{i=1}^{n}u_i(y_i-\mu_i)^2 + \sum_{i=1}^{n}\log h(u_i|\nu) - \frac{\alpha}{2}J(\mathbf{f}) + C,$$

where $h(\cdot|\nu)$ is the density of the mixing variable U and C is a constant independent of the parameter vector $\boldsymbol{\theta} = (\widehat{\boldsymbol{\beta}}^\top, \widehat{\mathbf{f}}^\top \widehat{\sigma^2})^\top$

SMN-PCR model

Parameter estimation via ECME algorithm

Like Ibacache-Pulgar *et al.* (2013) and Ferreira & Paula (2017), we consider the following penalty function:

$$J(\mathbf{f}) = \int_a^b [f''(t)]^2 dt.$$

As in Green & Silverman (1993), we use the natural cubic spline as a solution for the smoothing function $f(\cdot)$, therefore $J(\mathbf{f}) = \mathbf{f}^\top \mathbf{K}\mathbf{f}$, where $\mathbf{K} \in \mathcal{R}^{r\times r}$ is a non-negative definite matrix that depends only on the knot differences.

SMN-PCR model

Q-function

The Q-function, $Q(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}}^{(k)}) = \mathrm{E}_{\boldsymbol{\theta}^{(k)}}[\ell_{cp}(\boldsymbol{\theta}|\boldsymbol{z})|\mathbf{Y}_{j\mathrm{obs}}, \widehat{\boldsymbol{\theta}}^{(k)}]$, can be written as:

$$Q(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}}^{(k)}) \propto -\frac{n}{2}\log\widehat{\sigma^2}^{(k)} - \frac{1}{2\widehat{\sigma^2}^{(k)}}\sum_{i=1}^{n}\left[\xi_{2i}(\widehat{\boldsymbol{\theta}}^{(k)}) - 2\xi_{1i}(\widehat{\boldsymbol{\theta}}^{(k)})\widehat{\mu}_i^{(k)} + \xi_{0i}(\widehat{\boldsymbol{\theta}}^{(k)})\widehat{\mu}_i^{(k)^2}\right] - \frac{\alpha}{2}\widehat{\mathbf{f}}^{(k)^\top}\mathbf{K}\widehat{\mathbf{f}}^{(k)},$$

or in the matrix form, we have

$$Q(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}}^{(k)}) \propto -\frac{n}{2}\log\widehat{\sigma^2}^{(k)} - \frac{1}{2\widehat{\sigma^2}^{(k)}}(\mathbf{1}_n^\top\boldsymbol{\xi}_2^{(k)} - 2\widehat{\boldsymbol{\mu}}^{(k)\top}\boldsymbol{\xi}_1^{(k)} + \widehat{\boldsymbol{\mu}}^{(k)\top}\boldsymbol{\Omega}^{(k)}\widehat{\boldsymbol{\mu}}^{(k)}) - \frac{\alpha}{2}\widehat{\mathbf{f}}^{(k)^\top}\mathbf{K}\widehat{\mathbf{f}}^{(k)},$$

where $\xi_{si}(\boldsymbol{\theta}^{(k)}) = \mathrm{E}_{\boldsymbol{\theta}^{(k)}}[U_i Y_i^s|y_{\mathrm{obs}i}]$ for $s = 0, 1, 2$, $\boldsymbol{\xi}_s^{(k)} = [\xi_{si}(\widehat{\boldsymbol{\theta}}^{(k)}), \ldots, \xi_{sn}(\widehat{\boldsymbol{\theta}}^{(k)})]^\top$ and $\mathbf{1}_n$ a $n \times 1$ vector of ones.

SMN-PCR model

E-step

1. Given $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}^{(k)}$, compute $\xi_{s_{\boldsymbol{l}}}(\widehat{\boldsymbol{\theta}}^{(k)})$ or $\boldsymbol{\xi}_s$ in matrix form, for $s = 0, 1, 2$:

   $*$ For a *censored observation* i, $Y_{\text{obs}_i} = \kappa_i$ iff $Y_i \leq \kappa_i$, we have

$$\xi_{s_{\boldsymbol{l}}}(\boldsymbol{\theta}^{(k)}) = \mathrm{E}_{\boldsymbol{\theta}^{(k)}}[U_i Y_i^s | Y_i \leq \kappa_i],$$

   From the works of Genç (2013) and Garay *et al.* (2017, Proposition 1) for $a < b$, $r \geq 1$, $A = (a, b)$ and $s = 0, 1, 2$, the $E[U^r Y^s | Y \in A]$ is given by

$$E\left[U^r | Y \in A\right] = \mathrm{B} * \left[E_{\Phi}(r, b) - E_{\Phi}(r, a)\right], \tag{6}$$

$$E\left[U^r Y | Y \in A\right] = \mathrm{B} * \left[E_{\phi}(r - 0.5, a) - E_{\phi}(r - 0.5, b)\right], \tag{7}$$

$$E\left[U^r Y^2 | Y \in A\right] = \mathrm{B} * \left[E_{\Phi}(r - 1, b) - E_{\Phi}(r - 1, a) + aE_{\phi}(r - 0.5,) - bE_{\phi}(r - 0.5, b)\right], \tag{8}$$

   with $\mathrm{B} = \left(F_{SMN}(b) - F_{SMN}(a)\right)^{-1}$.

Table: $E_\phi\,(r,h)$ and $E_\Phi\,(r,h)$ for some members of the SMN family of distributions.

Distribution	$E_\phi\,(r,h)$	$E_\Phi\,(r,h)$
Student-t	$\frac{\Gamma\left(\frac{\nu+2r}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)\sqrt{2\pi}}\left(\frac{\nu}{2}\right)^{\nu/2}\left(\frac{h^2+\nu}{2}\right)^{-\frac{(\nu+2r)}{2}}$	$\frac{\Gamma\left(\frac{\nu+2r}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)}\left(\frac{\nu}{2}\right)^{-r}F_{PVII}(h\|\nu+2r,\nu)$
Slash	$\frac{\nu}{\sqrt{2\pi}}\left(\frac{h^2}{2}\right)^{-(\nu+1)}\Gamma\left(\nu+r,0.5h^2\right)$	$\left(\frac{\nu}{\nu+1}\right)F_{SL}(h\|\nu+r)$
Contaminated normal	$\varphi\gamma^r\phi(h\sqrt{\gamma})+(1-\varphi)\phi(h)$	$\gamma^r F_{CN}(h\|\varphi,\gamma)+(1-\gamma^r)\Phi(h)$

SMN-PCR model

∗ For an *uncensored observation* i,

$$\xi_{s_i}(\boldsymbol{\theta}^{(k)}) = y_i^s \mathrm{E}_{\boldsymbol{\theta}^{(k)}}[U_i|y_i].$$

Thus, by Osorio *et al.* (2007):

Table: $\mathrm{E}_{\boldsymbol{\theta}^{(k)}}[U_i|Y_i]$ for some members of the SMN family of distributions.

Distribution		
Student-t	Slash	Contaminated normal
$\frac{(\nu+1)}{\nu + d^2(\boldsymbol{\theta}^{(k)}, y_i)}$	$\frac{\Gamma\left(\nu + 1.5, d^2(\boldsymbol{\theta}^{(k)}, y_i)/2\right)}{\Gamma\left(\nu + 0.5, d^2(\boldsymbol{\theta}^{(k)}, y_i)/2\right)}$	$\frac{1-\varphi+\varphi\gamma^{1.5}e^{0.5(1-\gamma)d^2(\boldsymbol{\theta}^{(k)},y_i)}}{1-\varphi+\varphi\gamma^{0.5}e^{0.5(1-\gamma)d^2(\boldsymbol{\theta}^{(k)},y_i)}}$

SMN-PCR model

CM-step

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

$$
\begin{aligned}
\widehat{\boldsymbol{\beta}}^{(k+1)} &= \left[\sum_{i=1}^{n} \xi_{0_{\boldsymbol{I}}}(\widehat{\boldsymbol{\theta}}^{(k)})\mathbf{x}_i\mathbf{x}_i^\top\right]^{-1} \sum_{i=1}^{n} \mathbf{x}_i \left[\xi_{1_{\boldsymbol{I}}}(\widehat{\boldsymbol{\theta}}^{(k)}) - \xi_{0_{\boldsymbol{I}}}(\widehat{\boldsymbol{\theta}}^{(k)})\mathbf{n}_i^\top \widehat{\mathbf{f}}^{(k)}\right] \\
&= \left(\mathbf{X}^\top \boldsymbol{\Omega}^{(k)} \mathbf{X}\right)^{-1} \mathbf{X}^\top \left(\boldsymbol{\xi}_1^{(k)} - \boldsymbol{\Omega}^{(k)} \mathbf{N}^\top \widehat{\mathbf{f}}^{(k)}\right), \\
\widehat{\mathbf{f}}^{(k+1)} &= \left[\sum_{i=1}^{n} \xi_{0_{\boldsymbol{I}}}(\widehat{\boldsymbol{\theta}}^{(k)})\mathbf{n}_i\mathbf{n}_i^\top + \widehat{\alpha}^{(k)}\widehat{\sigma^2}^{(k)}\mathbf{K}\right]^{-1} \sum_{i=1}^{n} \mathbf{n}_i \left[\xi_{1_{\boldsymbol{I}}}(\widehat{\boldsymbol{\theta}}^{(k)}) - \xi_{0_{\boldsymbol{I}}}(\widehat{\boldsymbol{\theta}}^{(k)})\mathbf{x}_i^\top \widehat{\boldsymbol{\beta}}^{(k+1)}\right] \\
&= \left(\mathbf{N}^\top \boldsymbol{\Omega}^{(k)} \mathbf{N} + \widehat{\alpha}^{(k)}\widehat{\sigma^2}^{(k)}\mathbf{K}\right)^{-1} \mathbf{N}^\top \left(\boldsymbol{\xi}_1^{(k)} - \boldsymbol{\Omega}^{(k)} \mathbf{X} \widehat{\boldsymbol{\beta}}^{(k+1)}\right), \\
\widehat{\sigma^2}^{(k+1)} &= \frac{1}{n} \sum_{i=1}^{n} \left[\xi_{2_{\boldsymbol{I}}}(\widehat{\boldsymbol{\theta}}^{(k)}) - 2\xi_{1_{\boldsymbol{I}}}(\widehat{\boldsymbol{\theta}}^{(k)})\widehat{\mu}_i^{(k+1)} + \xi_{0_{\boldsymbol{I}}}(\widehat{\boldsymbol{\theta}}^{(k)})\widehat{\mu}_i{}^{(k+1)^2}\right] \\
&= \frac{1}{n}\left(\mathbf{1}_n^\top \boldsymbol{\xi}_2^{(k)} - 2\widehat{\boldsymbol{\mu}}^{(k+1)\top} \boldsymbol{\xi}_1^{(k)} + \widehat{\boldsymbol{\mu}}^{(k+1)\top} \boldsymbol{\Omega}^{(k)} \widehat{\boldsymbol{\mu}}^{(k+1)}\right),
\end{aligned}
$$

where $\boldsymbol{\Omega}$ is a diagonal matrix with elements $\xi_{0_{\boldsymbol{I}}}(\widehat{\boldsymbol{\theta}}^{(k)})$ of dimension $n \times n$, $\boldsymbol{\xi}_1^{(k)}$ and $\boldsymbol{\xi}_2^{(k)}$ are vectors of dimension $n \times 1$.

3. Update $\boldsymbol{\nu}^{(k)}$ by maximizing the actual marginal log-likelihood function, obtaining

$$\boldsymbol{\nu}^{(k+1)} = \underset{\boldsymbol{\nu}}{\text{argmax}} \left\{ \sum_{i=1}^{m} \log \left[F_{SMN}\left(\frac{\kappa_i - \widehat{\mu}_i^{(k+1)}}{\widehat{\sigma}^{(k+1)}} \right) \right] + \sum_{i=m+1}^{n} \log \left[f_{SMN}(y_i | \widehat{\mu}_i^{(k+1)}, \widehat{\sigma^2}^{(k+1)}, \boldsymbol{\nu}) \right] \right\}. \quad (9)$$

The algorithm iterates between the *E*- and *CML-steps* until reaching convergence, i.e., until some distance involving two successive evaluations of the actual log-likelihood, like $||\ell(\boldsymbol{\theta}^{(k+1)})/\ell(\boldsymbol{\theta}^{(k)}) - 1||$ is small enough.

Remark. Starting values:

- Calculate $\widehat{\boldsymbol{\beta}}^{(0)}$ and $\widehat{\sigma^2}^{(0)}$ as the solution of the least squares regression model of $\mathbf{Y}$ on $\mathbf{X}$, considering the censoring values as observed.

- $\widehat{\mathbf{f}}^{(0)} = (\mathbf{N}^\top \mathbf{N} + \alpha \widehat{\sigma^2}^{(0)} \mathbf{K})^{-1} \mathbf{N}^\top (\mathbf{Y}_j - \mathbf{X}\widehat{\boldsymbol{\beta}}^{(0)})$.

SMN-PCR model

Model selection and estimation of α

Following Ferreira & Paula (2017), the AIC for PLR models is defined by:

$$\mathrm{AIC}(\alpha) = -2\ell_{cp}(\widehat{\boldsymbol{\theta}}, \alpha) + 2[p + q + \mathrm{df}(\alpha)],$$

The degrees of freedom (df) can be approximated by:

$$\mathrm{df}(\alpha) = tr\{\mathbf{I}_r + \alpha Ł\},$$

where $Ł = \widehat{\sigma^2}\mathbf{B}^{-1/2}\mathbf{K}\mathbf{B}^{-1/2}$, with $\mathbf{B} = \mathbf{N}^\top\mathbf{N}$.

SMN-PCR model

Standard error approximation

The approximate variance-covariance matrix of $\boldsymbol{\theta}$ is derived from the inverse of the observed information matrix (Mark, Bacchetti,& Jewell1994). In effect,

$$\widehat{Var}_{approx}(\widehat{\boldsymbol{\theta}}) = I^{-1}_{\boldsymbol{\theta}\boldsymbol{\theta}}(\boldsymbol{\theta}|y)|_{\widehat{\boldsymbol{\theta}}} \longrightarrow \mathrm{I}_{\boldsymbol{\theta}\boldsymbol{\theta}}(\boldsymbol{\theta}|y) = -\sum_{i=1}^{n} \frac{\partial^2 \ell_{cp_i}(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}^\top},$$

$$\ell_{cp}(\boldsymbol{\theta}) = \sum_{i=1}^{m} \log\left[\Psi_i(\boldsymbol{\theta})\right] + \sum_{i=m+1}^{n} \left\{-\frac{1}{2}\log 2\pi - \frac{1}{2}\log\sigma^2 + \log[\psi_i(\boldsymbol{\theta})]\right\} - \frac{\alpha}{2}\mathbf{f}^\top \mathbf{K}\mathbf{f},$$

with

$$\Psi_i(\boldsymbol{\theta}) = \int_0^\infty \Phi\left[k^{-1/2}(u_i)\Delta_i\right] dH(u_i|\boldsymbol{\nu})$$

and

$$\psi_i(\boldsymbol{\theta}) = \int_0^\infty k^{-1/2}(u_i)\exp\left[\frac{-k^{-1}(u_i)d_i}{2}\right] dH(u_i|\boldsymbol{\nu}),$$

where $d_i = \frac{(y_i-\mu_i)^2}{\sigma^2}$ and $\Delta_i = \sqrt{d_i}$.

Thus, the matrix of second derivatives $\mathsf{I}_{\boldsymbol{\theta}\boldsymbol{\theta}}(\boldsymbol{\theta}|y)$ can be represented as:

$$\mathsf{I}_{\boldsymbol{\theta}\boldsymbol{\theta}}(\boldsymbol{\theta}|y) = -\sum_{i=1}^{n} \frac{\partial^2 \ell_{cp_i}(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}^\top} = \mathsf{I}^1(\boldsymbol{\theta}) + \mathsf{I}^2(\boldsymbol{\theta}) + \mathsf{I}^3(\boldsymbol{\theta}),$$

where

$$\begin{aligned}
\mathsf{I}^1(\boldsymbol{\theta}) &= -\sum_{i=1}^{m} \left\{ \frac{\partial^2}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}^\top} \log[\Psi_i(\boldsymbol{\theta})] \right\} = \sum_{i=1}^{m} \left[\frac{1}{\Psi_i^2(\boldsymbol{\theta})} \frac{\partial\Psi_i(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}} \frac{\partial\Psi_i(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}^\top} - \frac{1}{\Psi_i(\boldsymbol{\theta})} \frac{\partial^2\Psi_i(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}^\top} \right], \\
\mathsf{I}^2(\boldsymbol{\theta}) &= -\sum_{i=m+1}^{n} \left\{ \frac{\partial^2}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}^\top} \left[-\frac{1}{2}\log(2\pi) - \frac{1}{2}\log(\sigma^2) - \frac{\alpha}{2(n-m)} \mathbf{f}^\top \mathbf{K}\mathbf{f} \right] \right\}, \\
\mathsf{I}^3(\boldsymbol{\theta}) &= -\sum_{i=m+1}^{n} \left\{ \frac{\partial^2}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}^\top} \log[\psi_i(\boldsymbol{\theta})] \right\} = \sum_{i=1}^{m} \left[\frac{1}{\psi_i^2(\boldsymbol{\theta})} \frac{\partial\psi_i(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}} \frac{\partial\psi_i(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}^\top} - \frac{1}{\psi_i(\boldsymbol{\theta})} \frac{\partial^2\psi_i(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}^\top} \right].
\end{aligned}$$

Then,

$$\frac{\partial \psi_i(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = -\frac{1}{2}\mathbb{I}_i^{\phi}(3/2)\frac{\partial d_i}{\partial \boldsymbol{\theta}}, \qquad \frac{\partial^2 \psi_i(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}\partial \boldsymbol{\theta}^\top} = \frac{1}{4}\mathbb{I}_i^{\phi}(5/2)\frac{\partial d_i}{\partial \boldsymbol{\theta}}\frac{\partial d_i}{\partial \boldsymbol{\theta}^\top} - \frac{1}{2}\mathbb{I}_i^{\phi}(3/2)\frac{\partial^2 d_i}{\partial \boldsymbol{\theta}\partial \boldsymbol{\theta}^\top},$$

$$\frac{\partial \Psi_i(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \mathbb{I}_i^{\Phi}(1/2)\frac{\partial \Delta_i}{\partial \boldsymbol{\theta}}, \qquad \frac{\partial^2 \Psi_i(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}\partial \boldsymbol{\theta}^\top} = -\frac{1}{2}\mathbb{I}_i^{\Phi}(3/2)\frac{\partial \Delta_i^2}{\partial \boldsymbol{\theta}}\frac{\partial \Delta_i}{\partial \boldsymbol{\theta}} + \mathbb{I}_i^{\Phi}(1/2)\frac{\partial^2 \Delta_i}{\partial \boldsymbol{\theta}\partial \boldsymbol{\theta}^\top}.$$

Using the same notation as in Lachos *et al.* (2011), we have that:

$$\mathbb{I}_i^{\phi}(\omega) = \int_0^{\infty} k^{-\omega}(u_i)\exp\left[\frac{-k^{-1}(u_i)d_i}{2}\right] dH(u_i|\boldsymbol{\nu}) \tag{10}$$

SMN-PCR model

Standard error approximation

Since, $\mathbb{I}_i^{\Phi}(\omega) = \frac{1}{\sqrt{2\pi}}\mathbb{I}_i^{\phi}(\omega)$, for each distribution considered, the integral defined in (10) can be written as:

- Student-t distribution

$$\mathbb{I}_i^{\phi}(\omega) = \frac{\nu^{\nu/2}2^{\omega}\Gamma(\omega + \frac{\nu}{2})}{\Gamma\left(\frac{\nu}{2}\right)(\nu + d_i)^{\omega+\nu/2}};$$

- Slash distribution

$$\mathbb{I}_i^{\phi}(\omega) = \nu\int_0^1 u_i^{\omega+\nu-1}\exp\left(-\frac{u_i}{2}d_i\right)du_i;$$

- Contaminated normal distribution

$$\mathbb{I}_i^{\phi}(\omega) = \sqrt{2\pi}\left[\varphi\gamma^{\omega-1/2}\phi(\sqrt{d_i};0,1/\gamma) + (1-\varphi)\phi(\sqrt{d_i};0,1)\right].$$

Diagnostic analysis

Case deletion

Zhu & Lee (2001) proposed the following one-step pseudo approximation:

$$\widehat{\boldsymbol{\theta}}^{*}_{[-i]} = \widehat{\boldsymbol{\theta}} + \{-\ddot{Q}(\widehat{\boldsymbol{\theta}}|\widehat{\boldsymbol{\theta}})\}^{-1}\dot{Q}_{[-i]}(\widehat{\boldsymbol{\theta}}|\widehat{\boldsymbol{\theta}}), \qquad (11)$$

where

$$\ddot{Q}(\widehat{\boldsymbol{\theta}}|\widehat{\boldsymbol{\theta}}) = \frac{\partial^2 Q(\theta|\widehat{\theta})}{\partial\theta\partial\theta^\top}\Big|_{\theta=\widehat{\theta}} \quad \text{and} \quad \dot{Q}_{[-i]}(\widehat{\boldsymbol{\theta}}|\widehat{\boldsymbol{\theta}}) = \frac{\partial Q_{[-i]}(\theta|\widehat{\theta})}{\partial\theta}\Big|_{\theta=\widehat{\theta}},$$

with $\dot{Q}_{[-i]}(\widehat{\boldsymbol{\theta}}|\widehat{\boldsymbol{\theta}}) = (\dot{Q}_{[-i]_{\boldsymbol{\beta}}}(\widehat{\boldsymbol{\theta}}|\widehat{\boldsymbol{\theta}}),\ \dot{Q}_{[-i]_{\mathbf{f}}}(\widehat{\boldsymbol{\theta}}|\widehat{\boldsymbol{\theta}}),\ \dot{Q}_{[-i]_{\sigma^2}}(\widehat{\boldsymbol{\theta}}|\widehat{\boldsymbol{\theta}}))^\top$ has its elements as:

$$\begin{aligned}
\dot{Q}_{[-i]_{\boldsymbol{\beta}}}(\widehat{\boldsymbol{\theta}}|\widehat{\boldsymbol{\theta}}) &= \frac{1}{\widehat{\sigma^2}}\sum_{i\neq j}\left[\xi_{1_j}(\widehat{\boldsymbol{\theta}})\mathbf{x}_j - \xi_{0_j}(\widehat{\boldsymbol{\theta}})\mathbf{x}_j\widehat{\mu}_j\right], \\
\dot{Q}_{[-i]_{\mathbf{f}}}(\widehat{\boldsymbol{\theta}}|\widehat{\boldsymbol{\theta}}) &= \frac{1}{\widehat{\sigma^2}}\sum_{i\neq j}\left[\xi_{1_j}(\widehat{\boldsymbol{\theta}})\mathbf{n}_j - \xi_{0_j}(\widehat{\boldsymbol{\theta}})\mathbf{n}_j\widehat{\mu}_j\right] - \frac{\widehat{\alpha}}{n}\mathbf{K}\widehat{\mathbf{f}} \quad \text{and} \\
\dot{Q}_{[-i]_{\sigma^2}}(\widehat{\boldsymbol{\theta}}|\widehat{\boldsymbol{\theta}}) &= -\frac{1}{2\widehat{\sigma^2}}\sum_{i\neq j}\left\{1 - \frac{1}{\widehat{\sigma^2}}\left[\xi_{2_j}(\widehat{\boldsymbol{\theta}}) - 2\xi_{1_j}(\widehat{\boldsymbol{\theta}})\widehat{\mu}_j + \xi_{0_j}(\widehat{\boldsymbol{\theta}})\widehat{\mu}_j^2\right]\right\}.
\end{aligned}$$

Diagnostic analysis

Case deletion

To measure the distance between $\widehat{\boldsymbol{\theta}}_{[-i]}$ and $\widehat{\boldsymbol{\theta}}$, we compute the *generalized Cook's distance* as:

$$GD_i = \left(\widehat{\boldsymbol{\theta}}_{[-i]} - \widehat{\boldsymbol{\theta}}\right)^{\top} \left\{-\ddot{Q}(\widehat{\boldsymbol{\theta}}|\widehat{\boldsymbol{\theta}})\right\} \left(\widehat{\boldsymbol{\theta}}_{[-i]} - \widehat{\boldsymbol{\theta}}\right), \qquad i = 1, \ldots, n \tag{12}$$

and by substituting Equation (12) into (11), we obtain the approximation of the *generalized Cook's distance*

$$GD_i^1 = \dot{Q}_{[-i]}(\widehat{\boldsymbol{\theta}}|\widehat{\boldsymbol{\theta}})^{\top} \left\{-\ddot{Q}(\widehat{\boldsymbol{\theta}}|\widehat{\boldsymbol{\theta}})\right\}^{-1} \dot{Q}_{[-i]}(\widehat{\boldsymbol{\theta}}|\widehat{\boldsymbol{\theta}}) \qquad i = 1, \ldots, n.$$

Diagnostic analysis

The Hessian matrix, $\ddot{Q}(\widehat{\boldsymbol{\theta}}|\widehat{\boldsymbol{\theta}})$

After some rearrangement of terms and evaluation of the derivatives at $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}$, we obtain the Hessian matrix $\ddot{Q}(\widehat{\boldsymbol{\theta}}|\widehat{\boldsymbol{\theta}})$ with elements given by:

$$\ddot{Q}_{\boldsymbol{\beta}}(\widehat{\boldsymbol{\theta}}|\widehat{\boldsymbol{\theta}}) = \frac{\partial^2 Q(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}})}{\partial\boldsymbol{\beta}\partial\boldsymbol{\beta}^\top}\Big|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}} = -\frac{1}{\widehat{\sigma^2}}\sum_{i=1}^{n}\xi_{0i}(\widehat{\boldsymbol{\theta}})\mathbf{x}_i\mathbf{x}_i^\top,$$

$$\ddot{Q}_{\mathbf{f}}(\widehat{\boldsymbol{\theta}}|\widehat{\boldsymbol{\theta}}) = \frac{\partial^2 Q(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}})}{\partial\mathbf{f}\partial\mathbf{f}^\top}\Big|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}} = -\frac{1}{\widehat{\sigma^2}}\sum_{i=1}^{n}\xi_{0i}(\widehat{\boldsymbol{\theta}})\mathbf{n}_i\mathbf{n}_i^\top - \widehat{\alpha}\mathbf{K},$$

$$\ddot{Q}_{\sigma^2}(\widehat{\boldsymbol{\theta}}|\widehat{\boldsymbol{\theta}}) = \frac{\partial^2 Q(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}})}{\partial\sigma^2\partial\sigma^2}\Big|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}} = -\frac{n}{2\widehat{\sigma^2}} + \frac{1}{(\widehat{\sigma^2})^3}\sum_{i=1}^{n}\left[\xi_{2i}(\widehat{\boldsymbol{\theta}}) - 2\xi_{1i}(\widehat{\boldsymbol{\theta}})\widehat{\mu}_i + \xi_{0i}(\widehat{\boldsymbol{\theta}})\widehat{\mu}_i^2\right],$$

$$\ddot{Q}_{\boldsymbol{\beta}\mathbf{f}}(\widehat{\boldsymbol{\theta}}|\widehat{\boldsymbol{\theta}}) = \frac{\partial^2 Q(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}})}{\partial\boldsymbol{\beta}\partial\mathbf{f}^\top}\Big|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}} = -\frac{1}{\widehat{\sigma^2}}\sum_{i=1}^{n}\xi_{0i}(\widehat{\boldsymbol{\theta}})\mathbf{x}_i\mathbf{n}_i^\top,$$

$$\ddot{Q}_{\boldsymbol{\beta}\sigma^2}(\widehat{\boldsymbol{\theta}}|\widehat{\boldsymbol{\theta}}) = \frac{\partial^2 Q(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}})}{\partial\boldsymbol{\beta}\partial\sigma^2}\Big|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}} = -\frac{1}{(\widehat{\sigma^2})^2}\sum_{i=1}^{n}\left[\xi_{1i}(\widehat{\boldsymbol{\theta}})\mathbf{x}_i - \xi_{0i}(\widehat{\boldsymbol{\theta}})\mathbf{x}_i\widehat{\mu}_i\right],$$

$$\ddot{Q}_{\mathbf{f}\sigma^2}(\widehat{\boldsymbol{\theta}}|\widehat{\boldsymbol{\theta}}) = \frac{\partial^2 Q(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}})}{\partial\mathbf{f}\partial\sigma^2}\Big|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}} = -\frac{1}{(\widehat{\sigma^2})^2}\sum_{i=1}^{n}\left[\xi_{1i}(\widehat{\boldsymbol{\theta}})\mathbf{n}_i - \xi_{0i}(\widehat{\boldsymbol{\theta}})\mathbf{n}_i\widehat{\mu}_i\right].$$

Diagnostic analysis

Local influence

We will follow the approach proposed by Zhu & Lee (2001). Consider $\boldsymbol{\omega} = (\omega_1, \ldots, \omega_n)^\top$ restricted to $\boldsymbol{\Omega} \in \mathcal{R}^n$ and we assume that a $\boldsymbol{\omega}_0 \in \boldsymbol{\Omega}$ exists such that $\ell_{cp}(\boldsymbol{\theta}, \boldsymbol{\omega}|\boldsymbol{z}) = \ell_{cp}(\boldsymbol{\theta}|\boldsymbol{z})$ for all $\boldsymbol{\theta}$. Let $\widehat{\boldsymbol{\theta}}(\boldsymbol{\omega}) = (\widehat{\boldsymbol{\beta}}(\boldsymbol{\omega})^\top, \widehat{\mathbf{f}}(\boldsymbol{\omega})^\top, \widehat{\sigma^2}(\boldsymbol{\omega}))^\top$ denote the maximum of the function $Q(\boldsymbol{\theta}, \boldsymbol{\omega}|\widehat{\boldsymbol{\theta}}) = \mathrm{E}[\ell_{cp}(\boldsymbol{\theta}, \boldsymbol{\omega}|\boldsymbol{z})|\boldsymbol{Y}_{obs}, \widehat{\boldsymbol{\theta}}]$. Then, the influence graph is defined as $\boldsymbol{\alpha}(\boldsymbol{\omega}) = (\boldsymbol{\omega}^\top, f_Q(\boldsymbol{\omega}))^\top$, where

$$f_Q(\boldsymbol{\omega}) = 2\Big[Q(\widehat{\boldsymbol{\theta}}|\widehat{\boldsymbol{\theta}}) - Q(\widehat{\boldsymbol{\theta}}(\boldsymbol{\omega})|\widehat{\boldsymbol{\theta}})\Big],$$

is the Q-displacement function.

Diagnostic analysis

Local influence

To approximate the Q-displacement function, the normal curvature $C_{f_Q,\boldsymbol{h}}(\boldsymbol{\theta})$ of $\boldsymbol{\alpha}(\boldsymbol{\omega})$ at $\boldsymbol{\omega}_0$ in the direction of a unit vector $\boldsymbol{h}$ is used to summarize the local behavior of $f_Q(\boldsymbol{\omega})$:

$$C_{f_Q,\boldsymbol{h}}(\boldsymbol{\theta}) = -2\boldsymbol{h}^\top \ddot{Q}_{\boldsymbol{\omega}_0}\boldsymbol{h} \;=\; 2\boldsymbol{h}^\top \Delta_{\boldsymbol{\theta},\boldsymbol{\omega}_0}^\top \{\ddot{Q}(\widehat{\boldsymbol{\theta}}|\widehat{\boldsymbol{\theta}})\}^{-1}\Delta_{\boldsymbol{\theta},\boldsymbol{\omega}_0}\boldsymbol{h},$$

leading to

$$-\ddot{Q}_{\boldsymbol{\omega}_0} = \Delta_{\boldsymbol{\theta},\boldsymbol{\omega}_0}^\top \{\ddot{Q}(\widehat{\boldsymbol{\theta}}|\widehat{\boldsymbol{\theta}})\}^{-1}\Delta_{\boldsymbol{\theta},\boldsymbol{\omega}_0},$$

where $\Delta_{\boldsymbol{\theta},\boldsymbol{\omega}_0} = \partial^2 Q(\boldsymbol{\theta},\boldsymbol{\omega}|\widehat{\boldsymbol{\theta}})/\partial\boldsymbol{\theta}\partial\boldsymbol{\omega}^\top = (\Delta_{\boldsymbol{\beta},\boldsymbol{\omega}_0}^\top, \Delta_{\mathbf{f},\boldsymbol{\omega}_0}^\top, \Delta_{\sigma^2,\boldsymbol{\omega}_0}^\top)^\top$ is the matrix of dimension $(p+r+1)\times n$ evaluated at $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}$.

Diagnostic analysis

Local influence

We use the conformal normal curvature (Poon & Poon, 1999), given by:

$$B_{f_Q,\boldsymbol{h}}(\boldsymbol{\theta}) = \frac{C_{f_Q,\boldsymbol{h}}(\boldsymbol{\theta})}{\mathrm{tr}\left[-2\ddot{Q}_{\boldsymbol{\omega}_0}\right]} \quad \Rightarrow \quad B_{f_Q,\boldsymbol{h}_l}(\boldsymbol{\theta}) = \frac{\Delta_{l_{\boldsymbol{\theta},\boldsymbol{\omega}_0}}^{\top}\{\ddot{Q}(\widehat{\boldsymbol{\theta}}|\widehat{\boldsymbol{\theta}})\}^{-1}\Delta_{l_{\boldsymbol{\theta},\boldsymbol{\omega}_0}}}{\mathrm{tr}\left[\Delta_{\boldsymbol{\theta},\boldsymbol{\omega}_0}^{\top}\{\ddot{Q}(\widehat{\boldsymbol{\theta}}|\widehat{\boldsymbol{\theta}})\}^{-1}\Delta_{\boldsymbol{\theta},\boldsymbol{\omega}_0}\right]}, \quad (13)$$

$l = 1, \ldots, n$, and $0 \leq B_{f_Q,\boldsymbol{h}_l}(\boldsymbol{\theta}) \leq 1$.

Benchmark value: Lee & Xu (2004) proposed to use

$$M(0)_l > \overline{M(0)} + c^* SM(0),$$

with c^* being a selected constant greater than 2.

(a) Case-weight perturbation: The so-called perturbed Q-function, considering an arbitrary attribution of weights given by

$$Q(\boldsymbol{\theta}, \boldsymbol{\omega}|\widehat{\boldsymbol{\theta}}) = \sum_{i=1}^{n} \omega_i \mathrm{E}[\ell_{cp_i}(\boldsymbol{\theta}|\boldsymbol{z})|\boldsymbol{Y}_{obs}, \widehat{\boldsymbol{\theta}}] = \sum_{i=1}^{n} \omega_i Q_i(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}}) - \frac{\widehat{\alpha}}{2}\widehat{\mathbf{f}}^{\top}\mathbf{K}\widehat{\mathbf{f}}.$$

In this perturbation scheme, the matrix $\Delta_{\boldsymbol{\theta},\boldsymbol{\omega}_0}$ has elements given by:

$$\Delta_{\boldsymbol{\beta},\boldsymbol{\omega}_0} = \frac{\partial^2 Q(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}})}{\partial\boldsymbol{\beta}\partial\omega_i}\Big|_{\boldsymbol{\omega}=\boldsymbol{\omega}_0} = \frac{\mathbf{x}_i}{\widehat{\sigma^2}}\left[\xi_{1_i}(\widehat{\boldsymbol{\theta}}) - \xi_{0_i}(\widehat{\boldsymbol{\theta}})\widehat{\mu}_i\right],$$

$$\Delta_{\mathbf{f},\boldsymbol{\omega}_0} = \frac{\partial^2 Q(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}})}{\partial\mathbf{f}\partial\omega_i}\Big|_{\boldsymbol{\omega}=\boldsymbol{\omega}_0} = \frac{\mathbf{n}_i}{\widehat{\sigma^2}}\left[\xi_{1_i}(\widehat{\boldsymbol{\theta}}) - \xi_{0_i}(\widehat{\boldsymbol{\theta}})\widehat{\mu}_i\right] \quad \text{and}$$

$$\Delta_{\sigma^2,\boldsymbol{\omega}_0} = \frac{\partial^2 Q(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}})}{\partial\sigma^2\partial\omega_i}\Big|_{\boldsymbol{\omega}=\boldsymbol{\omega}_0} = -\frac{1}{2\widehat{\sigma^2}} + \frac{1}{2(\widehat{\sigma^2})^2}\left[\xi_{2_i}(\widehat{\boldsymbol{\theta}}) - 2\xi_{1_i}(\widehat{\boldsymbol{\theta}})\widehat{\mu}_i + \xi_{0_i}(\widehat{\boldsymbol{\theta}})\widehat{\mu}_i^2\right].$$

(b) Scale perturbation: We assume that $Y_i \sim \mathrm{SMN}(\mu_i, \sigma^2(\omega_i), \boldsymbol{\nu})$, with $\sigma^2(\omega_i) = \omega_i^{-1}\sigma^2$, $\omega_i > 0$ for $i = 1, \ldots, n$. The perturbed Q-function is

$$Q(\boldsymbol{\theta}, \boldsymbol{\omega}|\widehat{\boldsymbol{\theta}}) = \sum_{i=1}^{n}\left\{-\frac{1}{2}\log\left(\frac{\widehat{\sigma^2}}{\omega_i}\right) - \frac{\omega_i}{2\widehat{\sigma^2}}\left[\xi_{2i}(\widehat{\boldsymbol{\theta}}) - 2\xi_{1i}(\widehat{\boldsymbol{\theta}})\widehat{\mu}_i + \xi_{0i}(\widehat{\boldsymbol{\theta}})\widehat{\mu}_i^2\right]\right\} - \frac{\widehat{\alpha}}{2}\widehat{\mathbf{f}}^{\top}\mathbf{K}\widehat{\mathbf{f}}.$$

So, the matrix $\Delta_{\boldsymbol{\theta},\boldsymbol{\omega}_0}$, has elements given by:

$$\Delta_{\boldsymbol{\beta},\boldsymbol{\omega}_0} = \frac{\partial^2 Q(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}})}{\partial\boldsymbol{\beta}\partial\omega_i}\Big|_{\boldsymbol{\omega}=\boldsymbol{\omega}_0} = \frac{\mathbf{x}_i}{\widehat{\sigma^2}}\left[\xi_{1i}(\widehat{\boldsymbol{\theta}}) - \xi_{0i}(\widehat{\boldsymbol{\theta}})\widehat{\mu}_i\right],$$

$$\Delta_{\mathbf{f},\boldsymbol{\omega}_0} = \frac{\partial^2 Q(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}})}{\partial\mathbf{f}\partial\omega_i}\Big|_{\boldsymbol{\omega}=\boldsymbol{\omega}_0} = \frac{\mathbf{n}_i}{\widehat{\sigma^2}}\left[\xi_{1i}(\widehat{\boldsymbol{\theta}}) - \xi_{0i}(\widehat{\boldsymbol{\theta}})\widehat{\mu}_i\right] \quad \text{and}$$

$$\Delta_{\sigma^2,\boldsymbol{\omega}_0} = \frac{\partial^2 Q(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}})}{\partial\sigma^2\partial\omega_i}\Big|_{\boldsymbol{\omega}=\boldsymbol{\omega}_0} = \frac{1}{2(\widehat{\sigma^2})^2}\left[\xi_{2i}(\widehat{\boldsymbol{\theta}}) - 2\xi_{1i}(\widehat{\boldsymbol{\theta}})\widehat{\mu}_i + \xi_{0i}(\widehat{\boldsymbol{\theta}})\widehat{\mu}_i^2\right].$$

(c) Explanatory variable perturbation: The r-th explanatory variable of the design matrix is perturbed as $\mathbf{x}_{i\omega}^{\top} = \mathbf{x}_i^{\top} + \omega_i \mathcal{S}_r \mathbf{e}_r^{\top}$ for $r = 1, \ldots, p$, so

$$Q(\boldsymbol{\theta}, \boldsymbol{\omega}|\widehat{\boldsymbol{\theta}}) = -\frac{n}{2}\log\widehat{\sigma^2} - \frac{1}{2\widehat{\sigma^2}}\sum_{i=1}^{n}\left[\xi_{2i}(\widehat{\boldsymbol{\theta}}) - 2\xi_{1i}(\widehat{\boldsymbol{\theta}})\widehat{\mu}_i^{*} + \xi_{0i}(\widehat{\boldsymbol{\theta}})\widehat{\mu}_i^{2*}\right] - \frac{\widehat{\alpha}}{2}\widehat{\mathbf{f}}^{\top}\mathbf{K}\widehat{\mathbf{f}},$$

with $\widehat{\mu}_i^{*} = \mathbf{x}_{i\omega}^{\top}\widehat{\boldsymbol{\beta}} + \mathbf{n}_i^{\top}\widehat{\mathbf{f}}$. Thus,

$$\Delta_{\boldsymbol{\beta},\boldsymbol{\omega}_0} = \frac{\partial^2 Q(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}})}{\partial\boldsymbol{\beta}\partial\omega_i}\Big|_{\boldsymbol{\omega}=\boldsymbol{\omega}_0} = \frac{\mathcal{S}_r}{\widehat{\sigma^2}}\left\{\xi_{1i}(\widehat{\boldsymbol{\theta}})\mathbf{e}_r - \xi_{0i}(\widehat{\boldsymbol{\theta}})\left[\widehat{\mu}_i\mathbf{e}_r + \mathbf{e}_r^{\top}\widehat{\boldsymbol{\beta}}\mathbf{x}_i\right]\right\},$$

$$\Delta_{\mathbf{f},\boldsymbol{\omega}_0} = \frac{\partial^2 Q(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}})}{\partial\mathbf{f}\partial\omega_i}\Big|_{\boldsymbol{\omega}=\boldsymbol{\omega}_0} = -\frac{\mathcal{S}_r}{\widehat{\sigma^2}}\xi_{0i}(\widehat{\boldsymbol{\theta}})\mathbf{e}_r^{\top}\widehat{\boldsymbol{\beta}}\ \mathbf{n}_i \quad \text{and}$$

$$\Delta_{\sigma^2,\boldsymbol{\omega}_0} = \frac{\partial^2 Q(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}})}{\partial\sigma^2\partial\omega_i}\Big|_{\boldsymbol{\omega}=\boldsymbol{\omega}_0} = \frac{\mathcal{S}_r}{(\widehat{\sigma^2})^2}\left[\xi_{0i}(\widehat{\boldsymbol{\theta}})\widehat{\mu}_i\mathbf{e}_r^{\top}\widehat{\boldsymbol{\beta}} - \xi_{1i}(\widehat{\boldsymbol{\theta}})\mathbf{e}_r^{\top}\widehat{\boldsymbol{\beta}}\right].$$

(d) Response variable perturbation: To perturb the response variable values, we replace Y_{obs_i} by $Y_{\text{obs}_i}(\omega_i) = Y_{\text{obs}_i} + \mathcal{S}_y\omega_i$ for $i = 1, \ldots, n$, we have

$$Y_{\text{obs}_i}(\omega_i) = \begin{cases} \kappa_i(\omega_i) & \text{if} \quad Y_i \leq \kappa_i; \\ Y_i(\omega_i) & \text{if} \quad Y_i > \kappa_i. \end{cases}$$

Therefore, $Y_i(\omega_i) = Y_i - \mathcal{S}_y\omega_i$ (Matos *et al.*, 2013) and

$$\begin{aligned} Q(\boldsymbol{\theta}, \boldsymbol{\omega}|\widehat{\boldsymbol{\theta}}) &= -\frac{n}{2}\log(\widehat{\sigma^2}) - \frac{1}{2\widehat{\sigma^2}}\sum_{i=1}^{n}\Big[\xi_{2\boldsymbol{i}}(\widehat{\boldsymbol{\theta}}) - 2\xi_{1\boldsymbol{i}}(\widehat{\boldsymbol{\theta}})\mathcal{S}_y\omega_i + \xi_{0\boldsymbol{i}}(\widehat{\boldsymbol{\theta}})\mathcal{S}_y^2\omega_i^2 \\ &- 2\xi_{1\boldsymbol{i}}(\widehat{\boldsymbol{\theta}})\widehat{\mu}_i + 2\xi_{0\boldsymbol{i}}(\widehat{\boldsymbol{\theta}})\mathcal{S}_y\omega_i\widehat{\mu}_i + \xi_{0\boldsymbol{i}}(\widehat{\boldsymbol{\theta}})\widehat{\mu}_i^2\Big] - \frac{\widehat{\alpha}}{2}\widehat{\mathbf{f}}^{\top}\mathbf{K}\widehat{\mathbf{f}}. \end{aligned}$$

Diagnostic analysis

Local influence - Perturbation schemes

The matrix $\Delta_{\boldsymbol{\theta},\boldsymbol{\omega}_0}$ has elements given by:

$$\Delta_{\boldsymbol{\beta},\boldsymbol{\omega}_0} = \frac{\partial^2 Q(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}})}{\partial\boldsymbol{\beta}\partial\omega_i}\Big|_{\boldsymbol{\omega}=\boldsymbol{\omega}_0} = -\frac{\mathcal{S}_y}{\widehat{\sigma^2}}\ \xi_{0\boldsymbol{I}}(\widehat{\boldsymbol{\theta}})\mathbf{x}_i,$$

$$\Delta_{\mathbf{f},\boldsymbol{\omega}_0} = \frac{\partial^2 Q(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}})}{\partial\mathbf{f}\partial\omega_i}\Big|_{\boldsymbol{\omega}=\boldsymbol{\omega}_0} = -\frac{\mathcal{S}_y}{\widehat{\sigma^2}}\ \xi_{0\boldsymbol{I}}(\widehat{\boldsymbol{\theta}})\mathbf{n}_i \quad \text{and}$$

$$\Delta_{\sigma^2,\boldsymbol{\omega}_0} = \frac{\partial^2 Q(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}})}{\partial\sigma^2\partial\omega_i}\Big|_{\boldsymbol{\omega}=\boldsymbol{\omega}_0} = \frac{\mathcal{S}_y}{(\widehat{\sigma^2})^2}\ \left[\xi_{0\boldsymbol{I}}(\widehat{\boldsymbol{\theta}})\widehat{\mu}_i - \xi_{1\boldsymbol{I}}(\widehat{\boldsymbol{\theta}})\right].$$

Summary

Simulation study

The first simulation reports the results of a Monte Carlo (MC) experiment designed to evaluate the performance of the proposed model to analyze the behavior of MPL estimates, the approximated standard errors, as well as its asymptotic properties. We consider the following PLR model:

$$Y_i = \mathbf{x}_i^\top \boldsymbol{\beta} + f(t_i) + \varepsilon_i, \qquad i = 1, \ldots, n, \quad (14)$$

- We generated left-censored samples from the model given in (14) with
  - Censoring levels: $0\%, 10\%, 20\%$ and 30%.
  - Sample sizes: $n = 200, 300, 400$ and 600.
  - $\boldsymbol{\beta} = (2, 4)^\top$, $\sigma^2 = 2$ and $x_{1i} \sim U(0, 1)$, $x_{2i} \sim U(1, 2)$.
  - $f(t) = 10\,\mathrm{sin}(2\pi t)$, with $t \in (0, 1.5)$.
- $\varepsilon_i \overset{\text{iid.}}{\sim} \mathrm{SMN}(0, \sigma^2, \boldsymbol{\nu})$.

Simulation study

- For each combination of censoring level and sample size, we generated 500 samples from the SMN-PCR model, in four different situations: N-PCR, T-PCR ($\nu = 4$), SL-PCR ($\nu = 2$) and CN-PCR ($\boldsymbol{\nu}^\top = (0.1, 0.1)$).

- We computed for each parameter $\boldsymbol{\theta}_k$, $k = 1, 2, 3$, the measures:

$$\overline{\widehat{\theta}}_k = \frac{1}{500}\sum_{j=1}^{500}\widehat{\theta}_k^{(j)}, \qquad \text{MC-SD} = \sqrt{\sum_{j=1}^{500}(\widehat{\theta}_k^{(j)} - \overline{\widehat{\theta}}_k)^2/499} \quad \text{and}$$

$$\text{OM-SD} = \frac{1}{500}\sum_{j=1}^{500}\text{SE}(\widehat{\theta}_k^{(j)}),$$

where MC-SD is the MC standard deviation and OM-SD is the average values of the approximate standard errors.

Simulation study - Parameter recovery

Table: Simulated data. Mean value, MC-SD and OM-SD, considering left censoring and $n = 400$.

		Model/Measure											
		N-PCR			**T-PCR**			**SL-PCR**			**CN-PCR**		
Parameter	C.L.	$\bar{\widehat{\theta}}_k$	MC-SD	OM-SD	$\bar{\widehat{\theta}}_k$	MC-SD	OM-SD	$\bar{\widehat{\theta}}_k$	MC-SD	OM-SD	$\bar{\widehat{\theta}}_k$	MC-SD	OM-SD
β_1	0%	1.998	0.247	0.238	2.039	0.298	0.284	2.011	0.305	0.315	2.033	0.275	0.268
	10%	1.997	0.253	0.245	2.041	0.298	0.291	2.008	0.309	0.323	2.026	0.280	0.275
	20%	2.004	0.268	0.261	2.066	0.327	0.311	2.022	0.336	0.344	2.050	0.298	0.294
	30%	2.012	0.292	0.280	2.072	0.365	0.333	2.053	0.361	0.368	2.065	0.318	0.315
β_2	0%	4.087	0.279	0.249	4.127	0.316	0.296	4.115	0.355	0.328	4.065	0.296	0.280
	10%	4.080	0.289	0.260	4.133	0.319	0.309	4.104	0.370	0.341	4.081	0.304	0.292
	20%	4.113	0.305	0.277	4.184	0.338	0.329	4.117	0.397	0.363	4.117	0.323	0.312
	30%	4.123	0.300	0.293	4.229	0.314	0.347	4.186	0.390	0.384	4.152	0.314	0.330
σ^2	0%	1.842	0.138	0.132	1.832	0.165	0.174	1.855	0.163	0.156	1.829	0.164	0.160
	10%	1.839	0.147	0.138	1.827	0.173	0.182	1.851	0.170	0.164	1.818	0.166	0.167
	20%	1.836	0.157	0.147	1.823	0.184	0.192	1.849	0.178	0.173	1.824	0.186	0.178
	30%	1.839	0.167	0.157	1.837	0.197	0.207	1.848	0.195	0.185	1.830	0.194	0.191

Simulation study - Parameter recovery

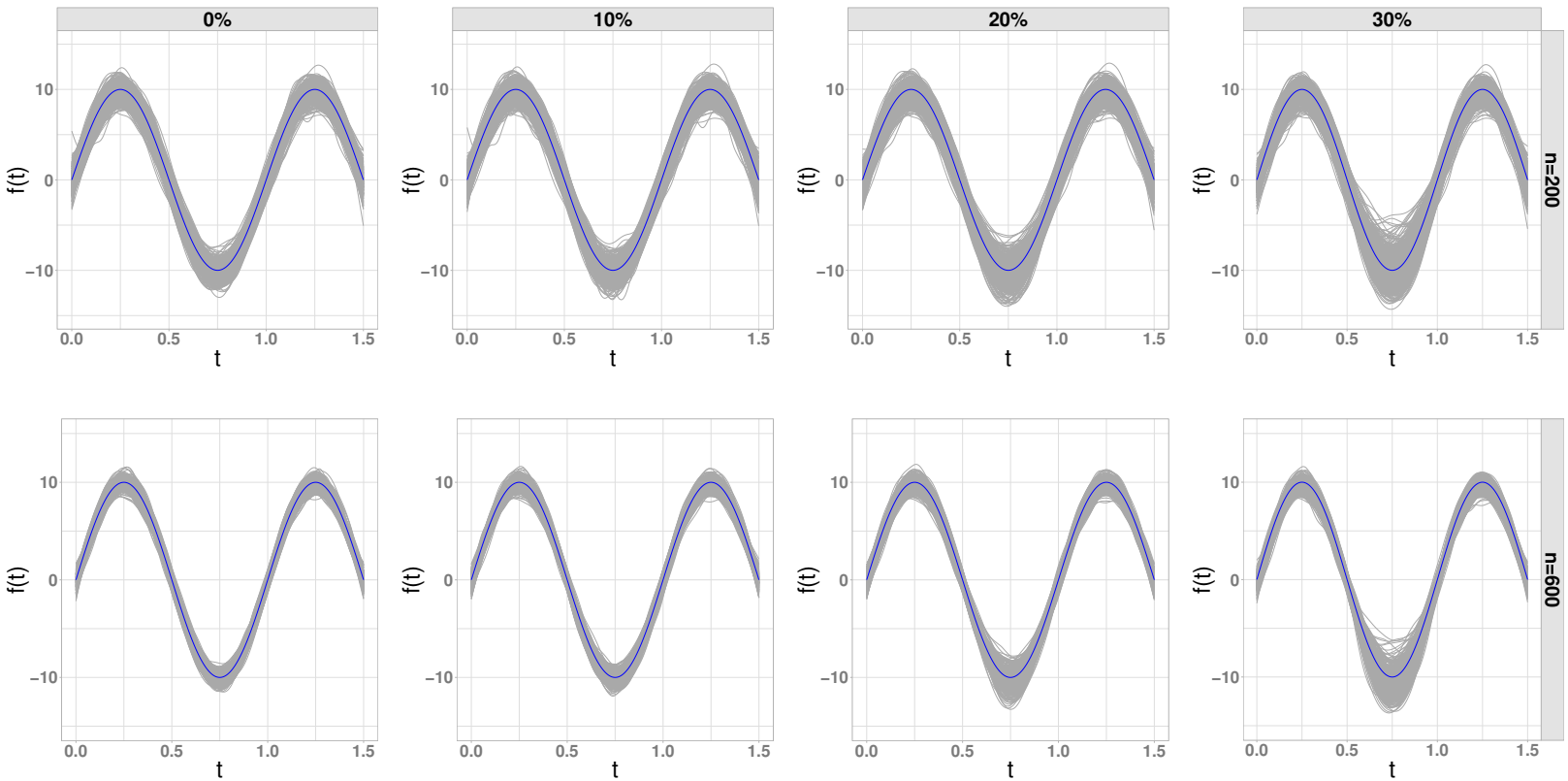

Figure: Simulated data. Behavior of the nonparametric component from the T-PCR model.

Simulation study - Asymptotic properties

- $\text{Bias}(\theta_k) = \frac{1}{500}\sum_{i=1}^{500}\left|\widehat{\theta}_k^{(j)} - \theta_k\right|$

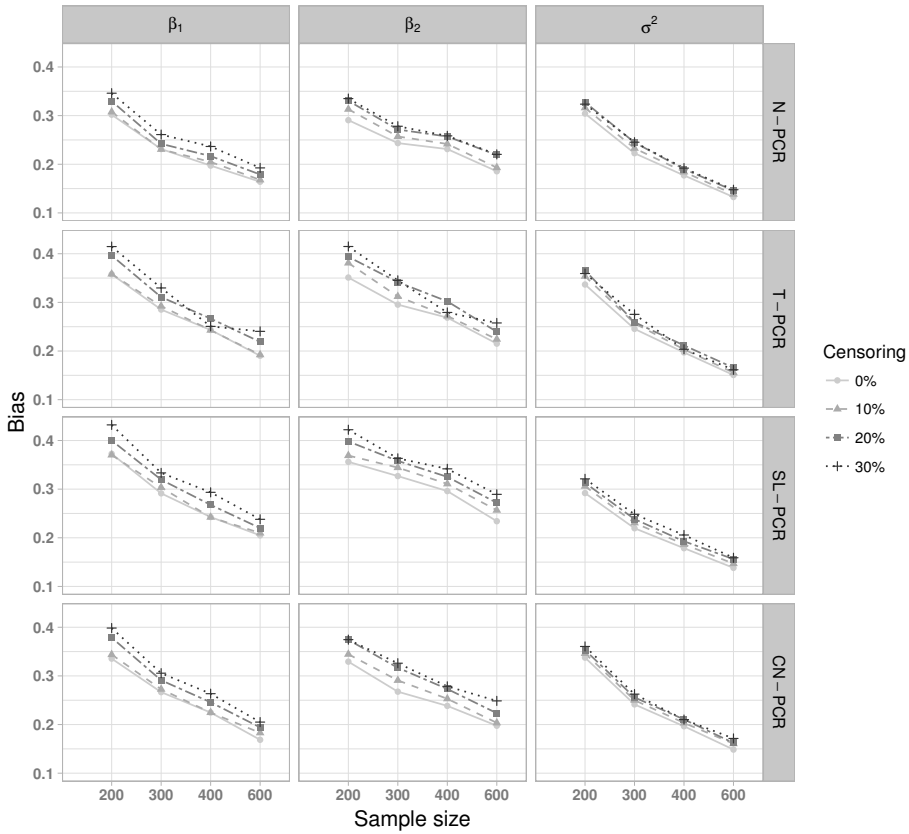

Figure: Simulated data. MC mean of bias.

- $\text{MSE}(\theta_k) = \frac{1}{500}\sum_{j=1}^{500}(\widehat{\theta}_k^{(j)} - \theta_k)^2$

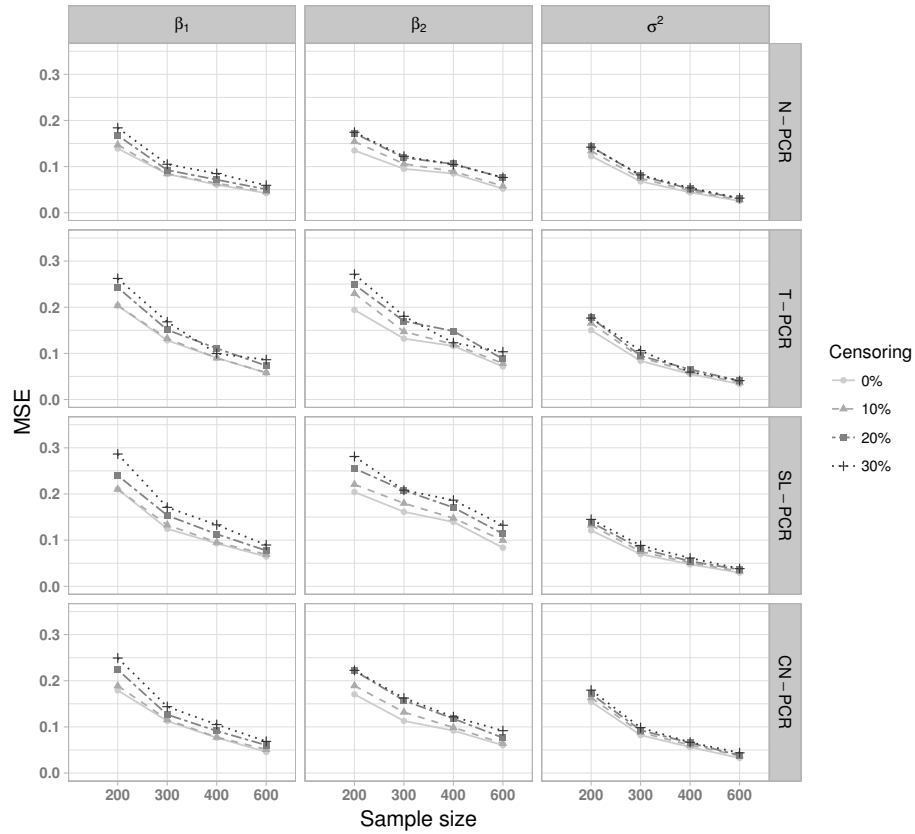

Figure: Simulated data. MC mean of MSE.

Simulation study - Robustness of the MPL estimates

We generate 100 MC samples of size $n = 200$ under N-PCR model. In this case, for the observation #66, we perturbed as

$$y_{66}(\eta) = y_{66} + \eta,$$

for $\eta \in \{2, 4, 6, 8, 10, 12, 14, 16\}$. We define the relative change as:

$$RC(\widehat{\theta}_i) = \left| \frac{\widehat{\theta}_i(\eta) - \widehat{\theta}_i}{\widehat{\theta}_i} \right|.$$

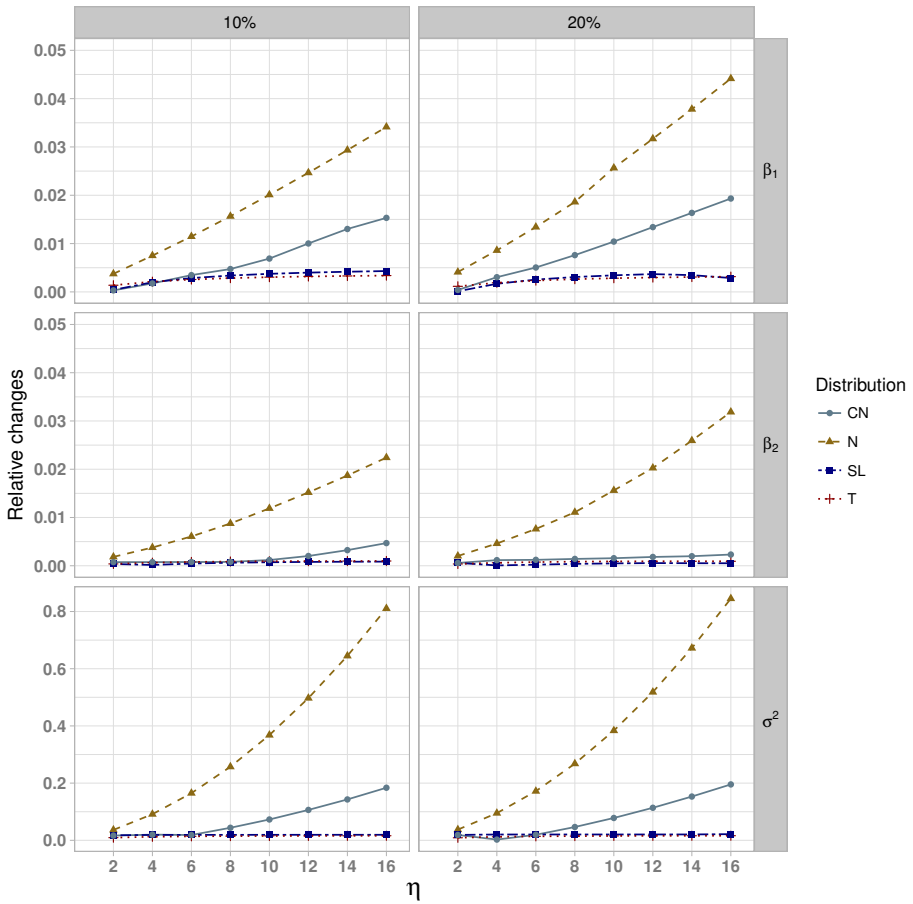

Figure: Simulated data. Relative changes.

Simulation study - Diagnostic measures

We generate 200 MC samples of size $n = 200$ under N-PCR model, 10% of censoring. For each sample, we contaminated the observation # 82 as follows:

- Replace β by 2β to generate the response of the observation #82 $\longrightarrow y_{82}$.
- β by 4β.
- β by 8β.

Table: Simulated data. Success percentages for different perturbation schemes in the N-PCR model and preference percentages under the T, SL and CN models, for different contamination schemes.

	Normal				heavy-tailed		
Contamination	Case-weight	Scale	Explanatory	Response	T	SL	CN
2β	71.5	71.5	72.5	70.5	71.5	70.5	71.5
4β	95.5	95.5	94.5	89.0	94.0	94.0	94.0
8β	97.5	97.5	96.5	96.5	96.5	96.5	95.5

Simulation study - Diagnostic measures

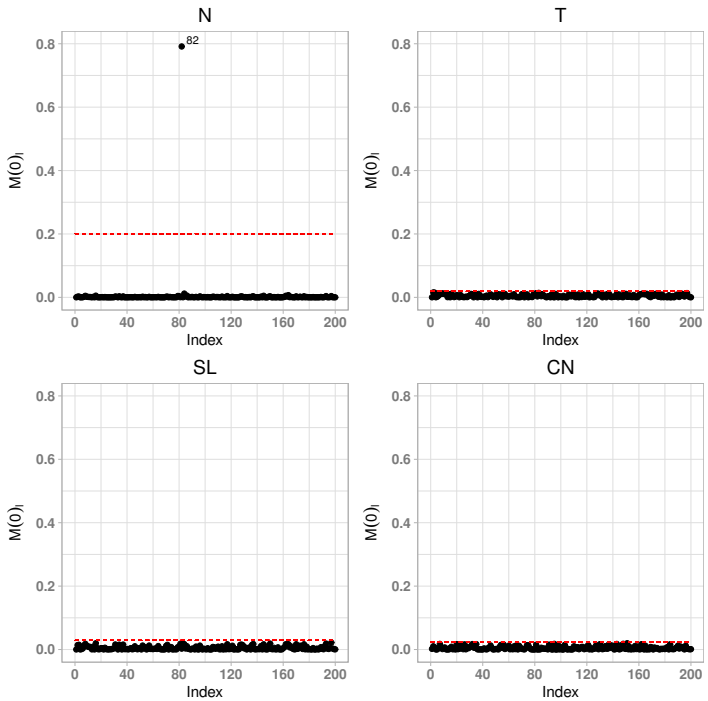

Figure: Simulated data. Index plot for 8β and $c^* = 3.5$. Case-weight perturbation.

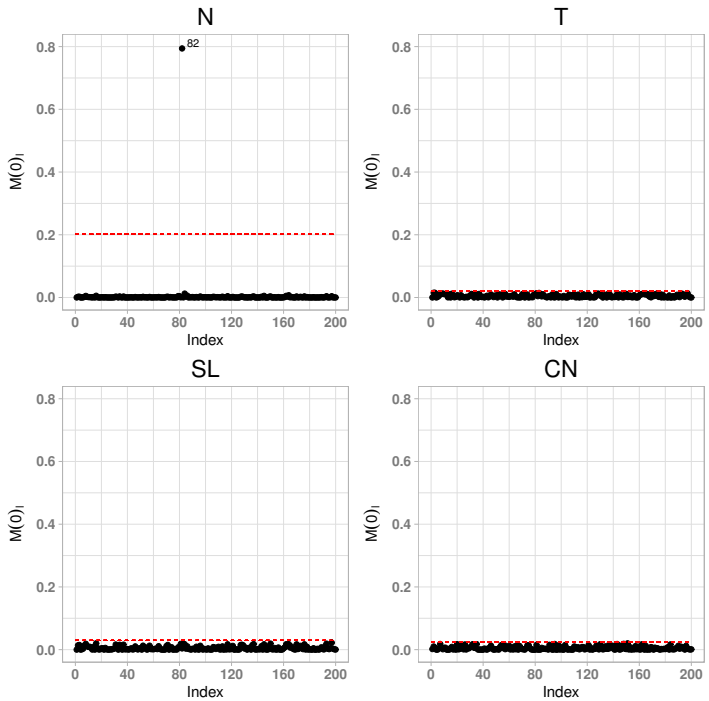

Figure: Simulated data. Index plot for 8β and $c^* = 3.5$. Scale perturbation.

Simulation study - Diagnostic measures

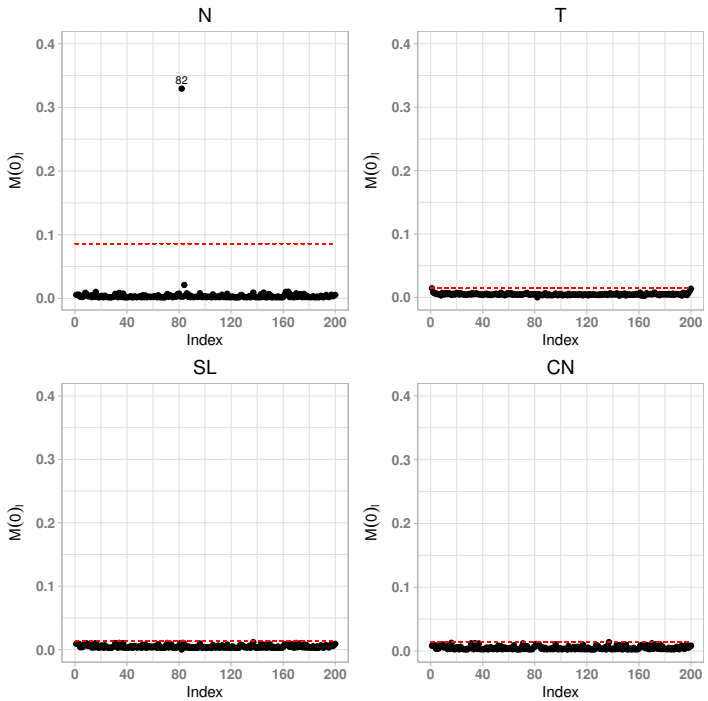

Figure: Simulated data. Index plot for 8β and $c^* = 3.5$. Explanatory variable perturbation.

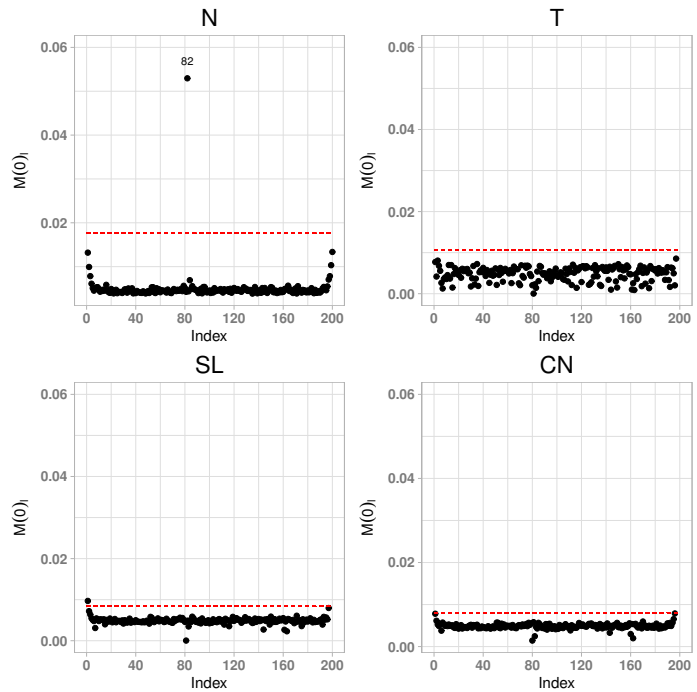

Figure: Simulated data. Index plot for 8β and $c^* = 3.5$. Response variable perturbation.

Wage rate data

We illustrate the performance of the proposed method by analyzing the wage rate dataset described in Mroz (1987).

- 753 married white women between the ages of 30 and 60, with 428 working at some time during the year 1975.
- The variables involved in the study were:
  - Y_i: The average hourly earnings (wage rates).
  - x_{1i}: Years of schooling.
  - x_{2i}: The wife's age.
  - x_{3i}: Husband's hours worked.
  - x_{4i}: Husband's wage in dollars.
  - x_{5i}: Tax rate faced by the wife.
  - x_{6i}: Number of children < 6 years.
  - x_{7i}: Number of children > 6 and < 19 years.
  - t_i: Number of years worked (Experience).

Wage rate data

- For those who did not work in 1975, the wage rate is zero, so the variable can be classified as censored-uncensored.

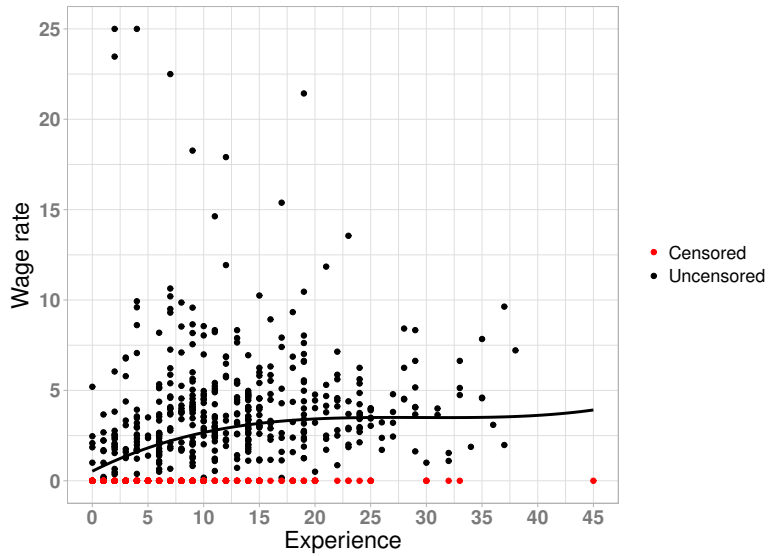

Figure: PSID-1975 dataset. Wage rates vs. experience.

The model:

$$Y_i = \beta_1 x_{1_i} + \beta_2 x_{2_i} + \beta_3 x_{3_i} + \beta_4 x_{4_i} + \beta_5 x_{5_i} + \beta_6 x_{6_i} + \beta_7 x_{7_i} + f(t_i) + \varepsilon_i,$$

$$\varepsilon_i \overset{\text{iid.}}{\sim} \text{SMN}(0, \sigma^2, \boldsymbol{\nu}), \quad i = 1, \ldots, 753.$$

Wage rate data: Analyses of the fitted models

Table: PSID-1975 dataset. Parameter estimates and standard errors (SE) for various fits of the SMN-PCR models.

Parameter	Model							
	N-PCR		**T-PCR**		**SL-PCR**		**CN-PCR**	
	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE
β_1	0.7688	(0.0881)	0.6672	(0.0783)	0.6620	(0.0757)	0.6682	(0.0730)
β_2	-0.0634	(0.0269)	-0.0743	(0.0204)	-0.0745	(0.0201)	-0.0746	(0.0207)
β_3	-0.0008	(0.0004)	-0.0004	(0.0003)	-0.0004	(0.0003)	-0.0005	(0.0003)
β_4	-0.1801	(0.0742)	-0.1321	(0.0642)	-0.1403	(0.0622)	-0.1527	(0.0605)
β_5	-8.6192	(3.8726)	-5.7045	(3.3936)	-6.0435	(3.2619)	-6.3471	(3.1808)
β_6	-1.8132	(0.4056)	-1.7666	(0.3136)	-1.7829	(0.3103)	-1.7874	(0.3254)
β_7	0.3257	(0.1456)	0.1986	(0.1136)	0.2038	(0.1098)	0.2145	(0.1126)
σ^2	15.885	(1.2320)	5.4806	(0.6375)	3.4223	(0.3788)	7.2270	(0.7089)
ν	-	-	2.8655	-	1.1248	-	-	-
φ	-	-	-	-	-	-	0.1	-
γ	-	-	-	-	-	-	0.1	-
α	0.0007	-	3.74e-07	-	3.74e-07	-	3.74e-07	-
$\ell(\widehat{\boldsymbol{\theta}})$	-1378.2	-	-1305.8	-	-1303.3	-	-1303.8	-
AIC	2852.2	-	2709.6	-	2704.5	-	2707.6	-

Wage rate data: Diagnostics analysis

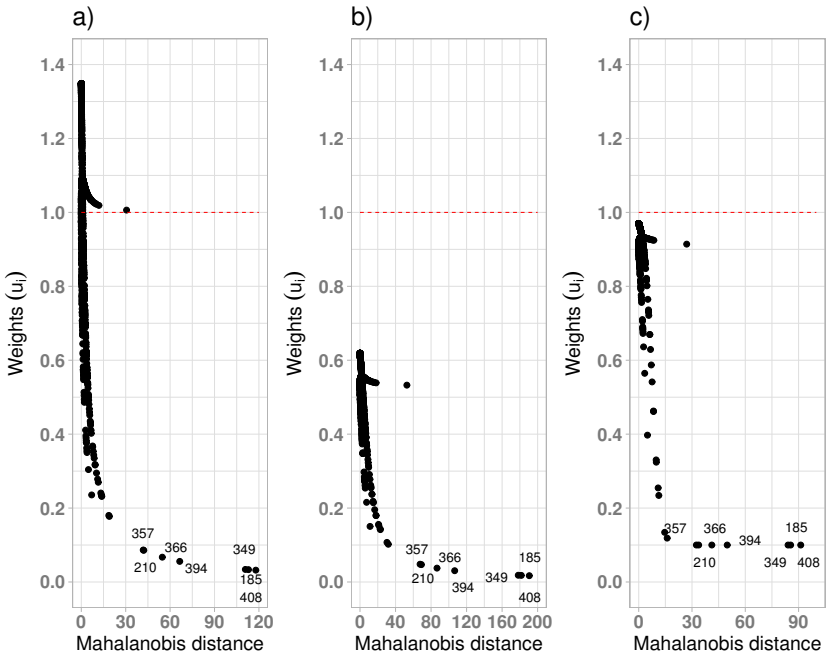

Figure: PSID-1975 dataset. Estimated weights u_i vs Mahalanobis distance d_i^2 for: a) T-PCR, b) SL-PCR and c) CN-PCR models.

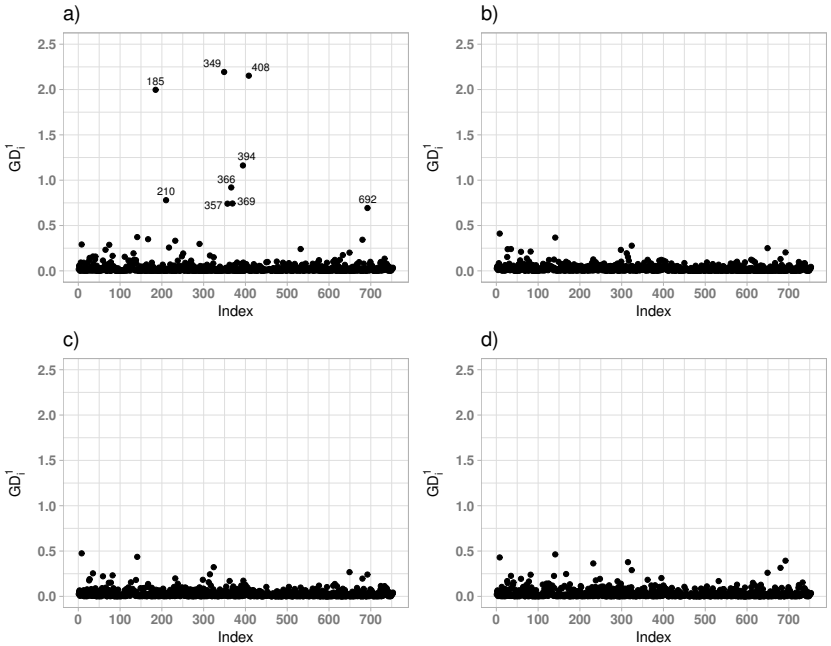

Figure: PSID-1975 dataset. Generalized Cook's distance GD_i^1 for: a) N-PCR, b) T-PCR, c) SL-PCR and d) CN-PCR models.

Wage rate data: Diagnostics analysis

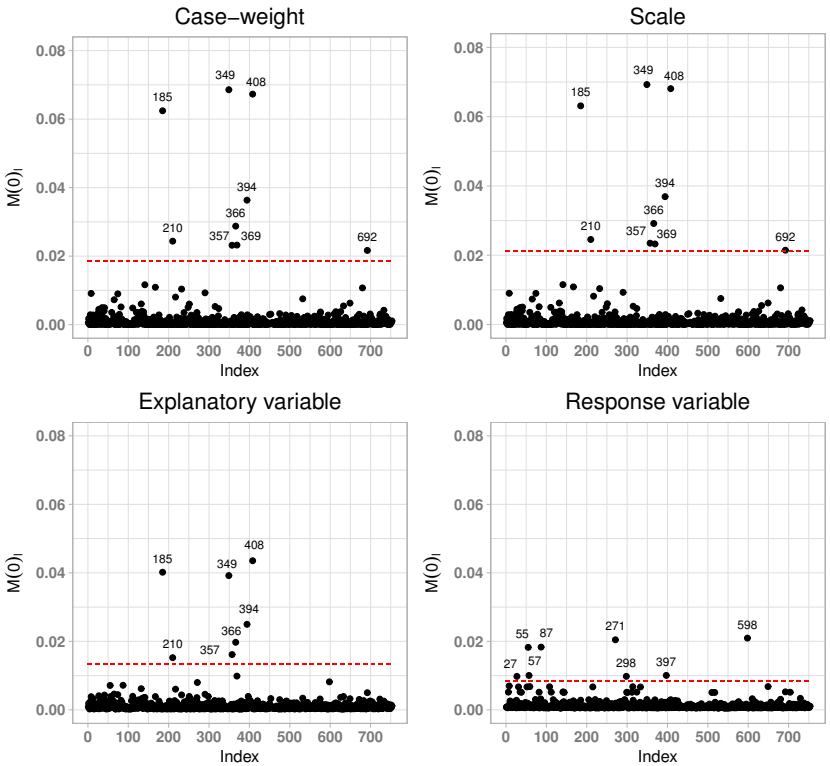

Figure: PSID-1975 dataset. Index plots for assessing local influence in N-PCR model, $c^* = 4$.

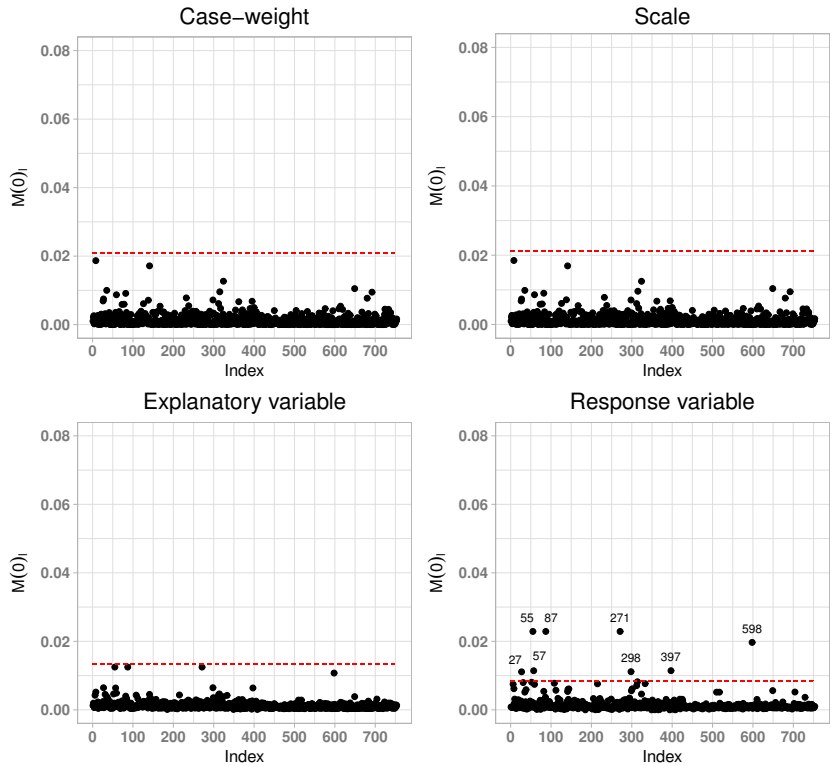

Figure: PSID-1975 dataset. Index plots for assessing local influence SL-PCR model, $c^* = 4$.

Wage rate data: Relative change in the MPL estimates

We define the relative change (RC) as

$$RC_j(\widehat{\boldsymbol{\theta}}) = \left| \frac{\widehat{\boldsymbol{\theta}} - \widehat{\boldsymbol{\theta}}_{[-\mathcal{I}_j]}}{\widehat{\boldsymbol{\theta}}} \right| \times 100\%.$$

Table: PSID-1975 dataset. Relative change (%) of maximum penalized likelihood estimates of $\widehat{\boldsymbol{\beta}}$ and $\widehat{\sigma^2}$ in N-PCR and SL-PCR models.

		Parameter							
Model	Dropped	$RC_{\widehat{\beta}_1}$	$RC_{\widehat{\beta}_2}$	$RC_{\widehat{\beta}_3}$	$RC_{\widehat{\beta}_4}$	$RC_{\widehat{\beta}_5}$	$RC_{\widehat{\beta}_6}$	$RC_{\widehat{\beta}_7}$	$RC_{\widehat{\sigma^2}}$
N(SL)	185	1.908(0.015)	7.551(0.182)	11.70(7.375)	7.446(1.812)	2.109(0.923)	6.693(0.280)	4.575(0.537)	8.718(4.376)
	210	0.855(0.138)	2.244(0.179)	3.378(6.497)	0.163(0.733)	0.216(0.421)	3.675(0.564)	6.568(1.806)	3.235(2.526)
	349	1.218(0.047)	20.28(1.004)	6.792(7.762)	2.789(1.114)	7.210(0.277)	0.485(0.026)	12.48(0.163)	8.291(4.353)
	357	2.260(0.396)	9.032(1.195)	4.944(8.373)	4.104(1.625)	0.002(0.503)	3.061(0.458)	6.400(2.160)	3.677(2.235)
	366	0.030(0.137)	14.11(1.293)	5.085(8.811)	11.98(0.662)	4.974(0.130)	0.348(0.062)	4.735(1.466)	4.481(2.854)
	369	0.425(0.262)	8.584(2.375)	5.543(9.181)	13.18(5.238)	2.333(1.036)	3.420(1.063)	3.465(2.470)	1.614(0.883)
	394	3.525(0.355)	8.433(0.678)	3.675(8.331)	2.474(0.735)	0.257(0.566)	2.077(0.309)	2.860(1.166)	5.467(3.046)
	408	0.335(0.073)	3.442(0.257)	2.979(8.894)	1.401(1.216)	5.536(0.412)	7.793(0.468)	13.77(0.205)	9.524(4.542)
	692	0.709(0.358)	0.713(0.213)	1.587(5.980)	1.687(1.120)	0.390(0.103)	0.231(0.044)	0.921(0.534)	0.576(0.355)
	all	11.24(1.445)	1.558(8.358)	19.32(42.01)	6.239(25.55)	14.59(12.68)	10.10(5.076)	10.33(24.97)	46.95(10.29)

Summary

Conclusions

- From a frequentist perspective, we develop the SMN-PCR model considering the SMN distributions and proposed a maximum penalized likelihood implementation of the partially linear regression (PCR) model with censored response, generalizing the papers of Garay *et al.* (2017) and Massuia *et al.* (2015), where the stochastic representation of the model allows a simple implementation of an EM-type algorithm (ECME algorithm).

- We also propose influence diagnostic tools for detecting influential observations in the context of PCR models with heavy-tailed distribution errors.

- The codes implemented in the work, were organized in the package `PartCensReg` and give computational support for estimation procedure and diagnostic analysis. The package is available in the CRAN repository.

Future research

- A natural extension would be to incorporate skewness and heavy tailedness simultaneously using scale mixtures of skew-normal (SMSN) distributions, as proposed in Lachos *et al.* (2010).

- Other extensions include considering semiparametric mixed effects models with censored data, following the same lines of ideas proposed by Matos *et al.* (2013) and Matos *et al.* (2015).

- Extend the analysis of local influence to a subset of the parameters of interest (local influence analysis on sub-vectors), following the work of Zhu *et al.* (2003), Ibacache & Paula (2011), Chen *et al.* (2012) and Relvas & Paula (2016).

Summary

Bibliography I

Andrews, D. F. & Mallows, C. L. (1974). Scale mixtures of normal distributions. *Journal of the Royal Statistical Society, Series B*, **36**, 99–102.

Castro, L, M., Lachos, V, H., Ferreira, G, P. & Arellano-Valle R, B. (2014). Partially linear censored regression models using heavy-tailed distributions: A bayesian approach. *Statistical Methodology*, **18**, 345–352.

Chen, X.-d., Tang, N.-s. & Wang, X.-r. (2012). Local influence analysis for semiparametric reproductive dispersion nonlinear models. *Acta Mathematicae Applicatae Sinica, English Series*, **28**(1), 75–90.

Ferreira, C. S. & Paula, G. A. (2017). Estimation and diagnostic for skew-normal partially linear models. *Journal of Applied Statistics*, **44**(16), 3033–3053.

Garay, A. M., Lachos, V. H., Bolfarine, H. & Cabral, C. R. (2017). Linear censored regression models with scale mixtures of normal distributions. *Statistical Papers*, **58**, 247–278.

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

Green, P. J. & Silverman, B. W. (1993). *Nonparametric regression and generalized linear models: a roughness penalty approach*. CRC Press.

Ibacache, P. G. & Paula, G. A. (2011). Local influence for student-t partially linear models. *Computational Statistics & Data Analysis*, **55**(3), 1462–1478.

Ibacache-Pulgar, G., Paula, G. & Cysneiros, F. (2013). Semiparametric additive models under symmetric distributions. *Test*, **22**, 103–121.

Lachos, V. H., Ghosh, P. & Arellano-Valle, R. B. (2010). Likelihood based inference for skew-normal independent linear mixed models. *Statistica Sinica*, **20**. SS-08-045.

Lachos, V. H., Angolini, T. & Abanto-Valle, C. A. (2011). On estimation and local influence analysis for measurement errors models under heavy-tailed distributions. *Statistical Papers*, **52**(3), 567–590.

Lee, S. Y. & Xu, L. (2004). Influence analysis of nonlinear mixed-effects models. *Computational Statistics and Data Analysis*, **45**, 321–341.

Massuia, B. M., Barbosa, C. C., Matos, A. L. & Lachos, V. H. (2015). Influence diagnostics for student-t censored linear regression models. *Statistics-A Journal of Theoretical and Applied Statistics*, **49**(5), 1074–1094.

Matos, L. A., Lachos, V. H., Balakrishnan, N. & Labra, F. V. (2013). Influence diagnostics in linear and nonlinear mixed-effects models with censored data. *Computational Statistics & Data Analysis*, **57**(1), 450–464.

Bibliography II

Matos, L. A., Bandyopadhyay, D., Castro, L. M. & Lachos, V. H. (2015). Influence assessment in censored mixed-effects models using the multivariate student's-t distribution. *Journal of multivariate analysis*, **141**, 104–117.

Mroz, T. A. (1987). The sensitivity of an empirical model of married women's hours of work to economic and statistical assumptions. *Econometrica: Journal of the Econometric Society*, pages 765–799.

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

Poon, W.-Y. & Poon, Y. S. (1999). Conformal normal curvature and assessment of local influence. *Journal of the Royal Statistical Society: Series B (Statistical Methodology*, **61**(1), 51–61.

Relvas, C. E. M. & Paula, G. A. (2016). Partially linear models with first-order autoregressive symmetric errors. *Statistical Papers*, **57**(3), 795–825.

Zhu, H. & Lee, S. (2001). Local influence for incomplete-data models. *Journal of the Royal Statistical Society, Series B*, **63**, 111–126.

Zhu, Z.-Y., He, X. & Fung, W.-K. (2003). Local influence analysis for penalized gaussian likelihood estimators in partially linear models. *Scandinavian Journal of Statistics*, **30**(4), 767–780.