Stat Methods Appl (2010) 19:379–397
DOI 10.1007/s10260-010-0135-y

Influence diagnostics in the tobit censored response model

Michelli Barros · Manuel Galea · Manuel González · Víctor Leiva

Accepted: 11 March 2010 / Published online: 27 April 2010

Abstract In this article, we develop influence diagnostic tools for the tobit model. Specifically, we discuss global influence methods based on the Cook distance and residuals with envelopes, and total and conformal local influence techniques. In order to analyze the sensitivity of the maximum likelihood estimators of the parameters of the model to small perturbations on the assumptions of the model and/or data, we consider several perturbation schemes, such as case-weight and response perturbations. Finally, we illustrate the developed methodology by means of a real data set.

Keywords Cook distance · Likelihood methods · Local influence · Residual analysis

M. Barros
Unidade Acadêmica de Matemática e Estatística, Universidade Federal de Campina Grande, Paraíba, Brazil
e-mail: michelli.karinne@gmail.com; michelli@dme.ufcg.edu.br

M. Galea
Departamento de Estadística, Pontificia Universidad Católica de Chile, Santiago, Chile
e-mail: mgalea@mat.puc.cl

M. González
Empresa El Mercurio SAP, Santiago, Chile
e-mail: manuelgonzalez@mercurio.cl; mgonzalezlucero@gmail.com

V. Leiva (✉)
Departamento de Estadística, CIMFAV, Universidad de Valparaíso, Avda Gran Bretaña 1111, Playa Ancha, 2360102 Valparaíso, Chile
e-mail: victor.leiva@yahoo.com; victor.leiva@uv.cl

1 Introduction

Classic regression models describe the functional relationship between a dependent quantitative variable usually called response and a set of independent variables usually called predictors. Some more specific regression models describe the relationship between a binary response and a set of one or more predictors. Traditionally, this last relationship has been considered by using logit and probit models. In the probit model, the response is a (unobserved) latent variable, say Y_i^*, for which the regression model

$$Y_i^* = \boldsymbol{x}_i^\top \boldsymbol{\beta} + \varepsilon_i, \quad i = 1, \ldots, n, \tag{1}$$

could be considered, where $\boldsymbol{x}_i = [x_{i1}, \ldots, x_{ip}]^\top$ is an $n \times 1$ vector that contains the values of the predictors, with x_{ij} being the value of the ith case for the jth independent variable, $\boldsymbol{\beta} = [\beta_1, \ldots, \beta_p]^\top$ is a $p \times 1$ vector of unknown regression coefficients to be estimated, and ε_i is the error term of the model. However, since it is not possible to observe the latent response Y_i^* given in Eq. (1), then an indicator variable defined by

$$Y_i = \begin{cases} 1, & Y_i^* > \tau; \\ 0, & Y_i^* \leq \tau; \end{cases}$$

is observed. Thus, instead of using $\mathrm{E}[Y^*|\boldsymbol{x}]$ in the linear form $\boldsymbol{x}^\top \boldsymbol{\beta}$,

$$\mathrm{E}\,[Y|\boldsymbol{x}] = \mathrm{P}\,(Y = 1) = \mathrm{P}\left(Y^* > \tau\right) = \mathrm{P}\left(\boldsymbol{x}^\top \boldsymbol{\beta} + \varepsilon > \tau\right) = 1 - F\left(\tau - \boldsymbol{x}^\top \boldsymbol{\beta}\right)$$

is considered, where $F(\cdot)$ is the cumulative distribution function (cdf) of ε given in Eq. (1). Now, suppose that Y_i^* is observed if $Y_i^* > \tau$ and it is not observed if $Y_i^* \leq \tau$, i.e., Y^* is a censored response to the left. In this case, every $Y_i^* \leq \tau$ is replaced by a known fixed value, say τ_y. In a censored sample, the $\boldsymbol{x}_i$'s are considered even when $Y_i^* \leq \tau$. However, in the case of a truncated sample, when Y_i^* is not observed, the value for the predictor neither is observed so that sample size decreases due to the unobserved part is not considered. (For more details about censored and truncated samples, see Schneider (1986) and Cohen (1991).) Consider a regression model with censored response to the left given by

$$Y_i = \begin{cases} Y_i^* = \boldsymbol{x}_i^\top \boldsymbol{\beta} + \varepsilon_i, & Y_i^* > \tau; \\ \tau_y, & Y_i^* \leq \tau; \end{cases} \tag{2}$$

where $\boldsymbol{\beta}$ and $\boldsymbol{x}_i$ are given in Eq. (1) and ε_i are independent and identically distributed random variables according to the normal model with mean zero and variance σ^2. Thus, Y_i is the response for the ith case, which is observed for values greater than τ and it is censored for values less than or equal to τ.

The configuration given in Eq. (2) is known as tobit model due to its similarity with the probit model; see Tobin (1958). This model is also known as censored response normal regression model. The tobit model has been used in diverse applications; see Amemiya (1984). This model can also be used for other types of censoring different from those to the left; see Scott (1997).

Another aspect that should be considered is that probit and tobit models for the latent variable (Y^*) are the same, but models for the observed response (Y) are different. In the tobit model, we know the value of Y^* when $Y^* > \tau$, while in the probit model we just know that $Y^* > \tau$, but we do not know its value. Due to this, there is more information in the tobit model than in the probit model. In addition, the estimates of $\boldsymbol{\beta}$ in the tobit model are more efficient than those obtained from the probit model. Moreover, for the censored cases of the probit model, it is not possible to estimate the variance of Y^* so that $\mathrm{Var}[\varepsilon|Y] = 1$. However, in the tobit model, $\mathrm{Var}[\varepsilon|Y]$ can be estimated from the data. For more details, see Scott (1997, p. 199).

A relevant point in regression models is to study how the estimates of the regression coefficients are affected by atypical observations. Diagnostic methods for normal regression models have been widely studied in the statistical literature; see Belsley et al. (1980), Cook and Weisberg (1982), and Chatterjee and Hadi (1988). The most of the works related to this topic have discussed how potentially influential cases affect the results of the model, in particular, on the estimates of the parameters of this model. Cook (1986) proposed a method called local influence in order to detect the effect of small perturbations in the model and/or the data on the parameter estimates. This method has had a great impact in the literature of the topic in the last decades. Several authors have extended the technique of local influence to more general regression models; see, e.g., Escobar and Meeker (1992), Paula (1993), Lawrence (1998), Galea et al. (1997, 2000, 2004), Ortega et al. (2003), Díaz-García et al. (2003), Leiva et al. (2007), Barros et al. (2008), and Paula et al. (2009).

The aims of this article are: (i) to develop diagnostics for the tobit model and (ii) to apply the obtained results to real data. The model, likelihood function, score vector and Hessian matrix are introduced in the following section. Diagnostic methods based on global and local influence for the tobit model are discussed in the third section. An application with real data that allows us to illustrate the developed methodology is presented in the fourth section. Some conclusions are given in the final section.

2 The tobit model

We consider the configuration for the tobit model given in Eq. (2), i.e., we assume that $\varepsilon_i \sim \mathrm{N}(0, \sigma^2)$, for $i = 1, \ldots, n$. A random sample of size n from this model includes m censored data and $n - m$ observed data. The m censored data correspond to the values of Y less than a threshold point τ, so that all these data take the value τ_y. The other $n - m$ data correspond to values of Y greater than τ. Thus, the likelihood function for $\boldsymbol{\theta} = [\boldsymbol{\beta}^\top, \sigma]^\top$ based on a censored random sample of size n, say $\boldsymbol{Y} = [Y_1, \ldots, Y_m, Y_{m+1}, \ldots Y_n]^\top$, from the tobit model is given by

$$\mathrm{L}(\boldsymbol{\theta}) = \prod_{i=1}^{m} \Phi\left(\frac{\tau - \boldsymbol{x}_i^\top \boldsymbol{\beta}}{\sigma}\right) \prod_{i=m+1}^{n} \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left[\frac{y_i - \boldsymbol{x}_i^\top \boldsymbol{\beta}}{\sigma}\right]^2\right), \quad (3)$$

where $\Phi(\cdot)$ is the standard normal cdf.

Remark 1 We note that second expression of the right side of Eq. (3) corresponds to a uncensored normal linear regression model, while the first one corresponds to the probit model given in Eq. (1) when the distribution of the error term is $F(\cdot) = \Phi(\cdot)$.

Remark 2 In the sequel, for simplicity, we consider that the observations indexed by $i = 1, \ldots, m$ in the tobit model given in Eq. (2) correspond to the censored cases, while the indexes $i = m + 1, \ldots, n$ correspond to the uncensored cases.

The log-likelihood function obtained from Eq. (2) is $\ell(\boldsymbol{\theta}) = \sum_{i=1}^{n} \ell_i(\boldsymbol{\theta})$, where

$$\ell_i(\boldsymbol{\theta}) = \begin{cases} \log\left(\Phi(\delta_i)\right), & i = 1, \ldots, m; \\ -\frac{1}{2}\log(2\pi) - \log(\sigma) - \frac{1}{2}\delta_i^2, & i = m + 1, \ldots, n; \end{cases} \tag{4}$$

with

$$\delta_i = \begin{cases} \dfrac{\tau - \boldsymbol{x}_i^\top \boldsymbol{\beta}}{\sigma}, & i = 1, \ldots, m; \\ \dfrac{y_i - \boldsymbol{x}_i^\top \boldsymbol{\beta}}{\sigma}, & i = m + 1, \ldots, n. \end{cases}$$

To estimate the parameters of the model given in Eq. (2), it is necessary to maximize the log-likelihood function given in Eq. (4) requiring the score vector given by

$$\dot{\ell} = \frac{\partial \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \sum_{i=1}^{n} \dot{\ell}_i, \quad \text{where} \quad \dot{\ell}_i = \left[\dot{\ell}_{i\beta}^\top, \dot{\ell}_{i\sigma}\right]^\top, \tag{5}$$

with

$$\dot{\ell}_{i\beta} = \begin{cases} -\frac{1}{\sigma} W(\delta_i)\, \boldsymbol{x}_i, & i=1, \ldots, m; \\ \frac{1}{\sigma}\, \delta_i\, \boldsymbol{x}_i, & i=m + 1, \ldots, n; \end{cases} \qquad \dot{\ell}_{i\sigma} = \begin{cases} -\frac{1}{\sigma} W(\delta_i)\, \delta_i, & i=1, \ldots, m; \\ -\frac{1}{\sigma} + \frac{1}{\sigma}\delta_i^2, & i=m + 1, \ldots, n; \end{cases}$$

$W(\delta_i) = \phi_i(\delta_i)/\Phi_i(\delta_i)$, δ_i given in Eq. (4), and $\phi(\cdot)$ denoting the standard normal probability density function. The maximum likelihood (ML) estimator of $\boldsymbol{\theta}$ is obtained from the likelihood equations given by $\dot{\ell} = \mathbf{0}$. However, these equations produce a nonlinear system so that a numerical iterative procedure is necessary in this case.

The observed information matrix for the tobit model is obtained as $\mathcal{J}(\boldsymbol{\theta}) = -\ddot{\ell}$, where $\ddot{\ell}$ denotes the Hessian matrix, which is given by

$$\ddot{\ell} = \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^\top} = \sum_{i=1}^{n} \ddot{\ell}_i, \quad \text{where} \quad \ddot{\ell}_i = \begin{bmatrix} \ddot{\ell}_{i\beta\beta} & \ddot{\ell}_{i\beta\sigma} \\ \ddot{\ell}_{i\sigma\beta} & \ddot{\ell}_{i\sigma\sigma} \end{bmatrix}, \tag{6}$$

with

$$\ddot{\ell}_{i\beta\beta} = \begin{cases} \frac{1}{\sigma^2} W'(\delta_i)\, \boldsymbol{x}_i \boldsymbol{x}_i^\top, & i = 1, \ldots, m; \\ -\frac{1}{\sigma^2} \boldsymbol{x}_i \boldsymbol{x}_i^\top, & i = m+1, \ldots, n; \end{cases}$$

$$\ddot{\ell}_{i\beta\sigma} = \begin{cases} \frac{1}{\sigma^2} \left[W(\delta_i) + \delta_i W'(\delta_i)\right] \boldsymbol{x}_i, & i = 1, \ldots, m; \\ -\frac{2}{\sigma^2} \delta_i \boldsymbol{x}_i, & i = m+1, \ldots, n; \end{cases}$$

$$\ddot{\ell}_{i\sigma\sigma} = \begin{cases} \frac{1}{\sigma^2} \left[2W(\delta_i) + \delta_i W'(\delta_i)\right] \delta_i, & i = 1, \ldots, m; \\ \frac{1}{\sigma^2} - \frac{3}{\sigma^2} \delta_i^2, & i = m+1, \ldots, n; \end{cases}$$

$\ddot{\ell}_{i\beta\sigma} = \ddot{\ell}_{i\sigma\beta}^\top$ and $W'(\cdot)$ being the derivative of $W(\cdot)$ given below Eq. (5).

Remark 3 The variance-covariance matrix of $\widehat{\boldsymbol{\theta}}$ can be obtained from the expected information matrix $\mathcal{I}(\boldsymbol{\theta}) = -\mathrm{E}[\,\ddot{\ell}\,]$ as $\mathrm{Var}[\widehat{\boldsymbol{\theta}}] = \boldsymbol{\Sigma}_{\widehat{\theta}} = \mathcal{I}(\boldsymbol{\theta})^{-1}$, where $\ddot{\ell}$ is given in Eq. (6). Instead of the expected information matrix, one may prefer using the observed information matrix $\mathcal{J}(\boldsymbol{\theta})$ to approximate the standard errors of the ML estimators by $\boldsymbol{\Sigma}_{\widehat{\theta}} \approx \mathcal{J}(\boldsymbol{\theta})^{-1}$. Now, in order to estimate these errors, one should compute the square roots of the diagonal elements of $\mathcal{J}(\boldsymbol{\theta})^{-1}$ and evaluate these roots at $\widehat{\boldsymbol{\theta}}$; for more details about this proposal, see Efron and Hinkley (1978).

Based on the asymptotic normality of the ML estimator $\widehat{\boldsymbol{\theta}}$, we can construct hypothesis tests and confidence regions for $\boldsymbol{\theta}$ by using $\widehat{\boldsymbol{\theta}} \,\dot{\sim}\, \mathrm{N}_{p+1}(\boldsymbol{\theta}, \boldsymbol{\Sigma}_{\widehat{\theta}})$. Thus, an approximate $100 \times [1-\varrho]\%$ confidence region for $\boldsymbol{\theta}$, with $0 < \varrho < 1$, is given by

$$\mathcal{R} = \left\{\boldsymbol{\theta} \in \mathbb{R}^{p+1} \colon [\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}]^\top \widehat{\boldsymbol{\Sigma}}_{\widehat{\theta}}^{-1} [\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}] \leq \chi^2_{1-\varrho}(p+1)\right\},$$

where $\chi^2_{1-\varrho}(p+1)$ is the $(1-\varrho)$th percentile of the χ^2 distribution with $p+1$ degrees of freedom and $\widehat{\boldsymbol{\Sigma}}_{\widehat{\theta}}$ is an estimate of $\boldsymbol{\Sigma}_{\widehat{\theta}}$ that can be obtained as mentioned in Remark 3.

3 Influence diagnostics

Diagnostic techniques have been proposed to detect observations that could exercise some influence on the parameter estimates. These techniques can be divided in two: (i) deletion of cases for assessing global influence and (ii) incorporation of diverse types of perturbations for assessing local influence. Next, for the tobit model, we describe the generalized Cook distance (GCD) and a residual analysis for global influence and total and conformal local influence procedures.

3.1 The generalized Cook distance

The elements of the coefficient vector $\boldsymbol{\beta}$ are of primary interest in a regression analysis. The Cook distance allows us to study the change on these estimated coefficients when a case is eliminated and it is considered as an important diagnostic technique of the global influence method; see Cook (1977) and Cook and Weisberg (1982). A generalization of the Cook distance for the tobit model based on $\boldsymbol{\theta} = [\boldsymbol{\beta}^\top, \sigma]^\top$ is given by

$$\mathrm{GCD}_i(\boldsymbol{\theta}) = \frac{1}{p+1}\left[\left[\widehat{\boldsymbol{\theta}} - \widehat{\boldsymbol{\theta}}_{(i)}\right]^\top \widehat{\boldsymbol{\Sigma}}_{\widehat{\theta}}^{-1}\left[\widehat{\boldsymbol{\theta}} - \widehat{\boldsymbol{\theta}}_{(i)}\right]\right], \quad i = 1, \ldots, n, \tag{7}$$

where p is the number of regression coefficients of the tobit model, $\widehat{\boldsymbol{\Sigma}}_{\widehat{\theta}}$ is an estimate of the variance-covariance matrix of $\widehat{\boldsymbol{\theta}}$, and $\widehat{\boldsymbol{\theta}}_{(i)}$ is the ML estimate of $\widehat{\boldsymbol{\theta}}$ without considering the ith observation. As mentioned, $\widehat{\boldsymbol{\Sigma}}_{\widehat{\theta}}$ can be approximated by $-\ddot{\ell}^{-1}$. In addition, if we use a first order approximation of the type $\widehat{\boldsymbol{\theta}} - \widehat{\boldsymbol{\theta}}_{(i)} \approx \ddot{\ell}_{(i)}^{-1}\,\dot{\ell}_{(i)}$, we obtain

$$\mathrm{GCD}_i(\boldsymbol{\theta}) \approx \frac{1}{p+1}\left[\dot{\ell}_{(i)}^\top\,\ddot{\ell}_{(i)}^{-1}[-\ddot{\ell}]\,\ddot{\ell}_{(i)}^{-1}\,\dot{\ell}_{(i)}\right], \quad i = 1, \ldots, n, \tag{8}$$

where $\dot{\ell}_{(i)}$ and $\ddot{\ell}_{(i)}$ are the score vector and the Hessian matrix associated with the tobit model defined in Eqs. (5) and (6), respectively, without considering the ith case, evaluated at $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}$. If we consider the approximation given in Eq. (8), we avoid the calculation of $\widehat{\boldsymbol{\theta}}$ eliminating the ith observation, for each $i = 1, \ldots, n$, and so this estimator is only computed once.

Remark 4 A high value for GCD_i indicates that the ith observation has a high impact on the ML estimate of $\boldsymbol{\theta}$. Considering the approximation for $\widehat{\boldsymbol{\Sigma}}_{\widehat{\theta}}$ indicated in Remark 3 and Theorem 2 given in Zhu and Zhang (2004), it follows that the mean of the GCD_i's is approximately $1/n$, such that we can use $2/n$ as benchmark for the GCD_i's, with $i = 1, \ldots, n$.

When the interest is just on the $p \times 1$ vector of regression coefficients, $\boldsymbol{\beta} = [\beta_1, \ldots, \beta_p]^\top$, then Eq. (7) is transformed to

$$\mathrm{GCD}_i(\boldsymbol{\beta}) = \frac{1}{p}\left[\left[\widehat{\boldsymbol{\beta}} - \widehat{\boldsymbol{\beta}}_{(i)}\right]^\top \widehat{\boldsymbol{\Sigma}}_{\widehat{\beta}}^{-1}\left[\widehat{\boldsymbol{\beta}} - \widehat{\boldsymbol{\beta}}_{(i)}\right]\right], \quad i = 1, \ldots, n. \tag{9}$$

Thus, we can use the approximation given in Eq. (8) for avoiding the calculation of $\widehat{\boldsymbol{\beta}}_{(i)}$ for each $i = 1, \ldots, n$ and, once again, we can also use the value $2/n$ as benchmark.

Remark 5 Alternatively, $\ddot{\ell}_{(i)}$ given in Eq. (8) can be replaced by $\ddot{\ell}$ in order to facilitate the computation of GCD_i; for more details, see Eq. (5) in Xie and Wei (2007).

3.2 Residual analysis

In order to study departures from the error assumptions as well as the presence of atypical observations, we consider two kinds of residuals: deviance component (DC) and martingale-type (MT), which can be revised in McCullagh and Nelder (1989), for the DC residual, in Barlow and Prentice (1988) and Therneau et al. (1990), for the MT residual, and both of them in Ortega et al. (2003).

3.2.1 Deviance component residual

The DC residual is defined as

$$r_{\mathrm{DC}_i} = \operatorname{sign}(y_i - \widehat{\mu}_i)\sqrt{2\,[\ell_i(\widehat{\boldsymbol{\theta}}_{\mathrm{s}}) - \ell_i(\widehat{\boldsymbol{\theta}})]}, \quad i =, 1, \ldots, n,$$

where $\widehat{\boldsymbol{\theta}}_{\mathrm{s}}$ is the ML estimate of $\boldsymbol{\theta}$ under the saturated model (with n parameters), $\widehat{\boldsymbol{\theta}}$ is the ML estimate of $\boldsymbol{\theta}$ under the model of interest (with $p+1$ parameters), $\ell_i(\cdot)$ is given in Eq. (4), $\widehat{\mu}_i = \widehat{\mathrm{E}[Y_i]}$ (with Y_i given in Eq. (2)), and sign(z) denotes the sign of z. Davison and Gigli (1989) defined the DC residual for censored data as

$$r_{\mathrm{DC}_i} = \operatorname{sign}(y_i - \widehat{\mu}_i)\sqrt{-2\log\left(\widehat{S}(y_i)\right)}, \quad i = 1, \ldots, n, \tag{10}$$

where $\widehat{S}(\cdot)$ is the ML estimate of the survival function of Y given in Eq. (2); see Leiva et al. (2007).

3.2.2 Martingale-type residual

Therneau et al. (1990) introduced the DC residual in counting process basically using martingale residuals, which are skewed and have minimum and maximum values at $-\infty$ and $+1$, respectively. For censored data, the martingale residual can be given by

$$r_{\mathrm{M}_i} = \eta_i + \log\left(\widehat{S}(y_i)\right), \quad i = 1, \ldots, n,$$

where $\eta_i = 0, 1$ indicates whether the observation is censored or not, respectively, and once gain $\widehat{S}(\cdot)$ is the ML estimate of the survival function of Y given in Eq. (2); see Klein and Moeschberger (1997, p. 359) and Ortega et al. (2003).

The DC residual proposed by Therneau et al. (1990) is a transformation of the martingale residual to attenuate the skewness. This transformation was motivated by the DC residuals found in generalized linear models. In particular, as in Leiva et al. (2007), we use for the tobit model a MT residual given by

$$r_{\mathrm{MT}_i} = \operatorname{sign}(r_{\mathrm{M}_i})\sqrt{-2\left[r_{\mathrm{M}_i} + \eta_i \log(\eta_i - r_{\mathrm{M}_i})\right]}, \quad i =, 1, \ldots, n,$$

where r_{M_i} is the martingale residual.

Remark 6 Although the MT residuals are not deviance components of the tobit model, we use these residuals as a transformation of the martingale residuals to have symmetrically distributed residuals around zero.

3.2.3 Standardized residuals

Ortega et al. (2003) suggested to standardize the DC and MT residuals for censored data as

$$r^*_{\mathrm{DC}_i} = \frac{r_{\mathrm{MT}_i}}{\sqrt{1-\mathrm{GL}_{ii}}} \quad \text{and} \quad r^*_{\mathrm{MT}_i} = \frac{r_{\mathrm{MT}_i}}{\sqrt{1-\mathrm{GL}_{ii}}}, \quad i = 1, \ldots, n,$$

respectively, with GL_{ii} being the ith principal diagonal element of the generalized leverage (GL) matrix given by

$$\mathrm{GL}(\boldsymbol{\theta}) = \boldsymbol{D}_\theta \left[-\ddot{\ell}\right]^{-1} \ddot{\ell}_{\theta y}, \tag{11}$$

where $\boldsymbol{D}_\theta = \partial\boldsymbol{\mu}/\partial\boldsymbol{\theta}^\top$, with $\boldsymbol{\mu} = \mathrm{E}[\boldsymbol{Y}]$ and $\boldsymbol{Y}$ given above Eq. (3), $\ddot{\ell}$ is the Hessian matrix given in Eq. (6) and

$$\ddot{\ell}_{\theta y} = \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}\partial\boldsymbol{y}^\top} = \begin{bmatrix} \dfrac{\partial^2 \ell(\boldsymbol{\theta})}{\partial\boldsymbol{\beta}\partial\boldsymbol{y}^\top} \\[2ex] \dfrac{\partial^2 \ell(\boldsymbol{\theta})}{\partial\sigma\partial\boldsymbol{y}^\top} \end{bmatrix} = \begin{bmatrix} \ddot{\ell}_{\beta y} \\ \ddot{\ell}_{\sigma y} \end{bmatrix},$$

with

$$\ddot{\ell}_{\beta y} = \begin{cases} 0, & i = 1, \ldots, m; \\ \frac{\boldsymbol{x}_i}{\sigma^2}, & i = m+1, \ldots, n; \end{cases} \quad \text{and} \quad \ddot{\ell}_{\sigma y} = \begin{cases} 0, & i = 1, \ldots, m, \\ \frac{2\delta_i}{\sigma^2}, & i = m+1, \ldots, n; \end{cases}$$

which must be evaluated at $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}$.

3.3 Local influence

Case deletion is a common way to assess the effect of an observation on the estimation procedure. This is a global influence analysis, since the effect of an observation is evaluated by eliminating it from the data set. Alternatively, local influence is based on a geometric differentiation instead of a observation deletion. In this case, a differential comparison of estimates is used before and after perturbing the data and/or model. There are several schemes to evaluate the local influence. Next, we describe this influence method.

3.3.1 The local influence method

The concept of local influence is based on the curvature of the plane of the log-likelihood function. In this respect, for the case of the tobit model given in Eq. (2), let $\boldsymbol{\theta} = [\boldsymbol{\beta}^\top, \sigma]^\top$ and $\ell(\boldsymbol{\theta}|\boldsymbol{\omega})$ be the parameter vector and the log-likelihood function corresponding to this model perturbed by $\boldsymbol{\omega}$, respectively. Here, $\boldsymbol{\omega}$ is a vector of perturbations belonging to a subset $\Omega \in \mathbb{R}^n$. In addition, $\boldsymbol{\omega}_0 = [1, \ldots, 1]^\top$ is an $n \times 1$ vector of ones called the non-perturbed vector, such that $\ell(\boldsymbol{\theta}|\boldsymbol{\omega}_0) = \ell(\boldsymbol{\theta})$, for all $\boldsymbol{\theta}$. The influence of the perturbations $\boldsymbol{\omega}$ on the ML estimate of $\boldsymbol{\theta}$ can be evaluated by the likelihood displacement (LD) given by $\mathrm{LD}(\boldsymbol{\omega}) = 2[\ell(\widehat{\boldsymbol{\theta}}) - \ell(\widehat{\boldsymbol{\theta}}_\omega)]$, where $\widehat{\boldsymbol{\theta}}_\omega$ denotes the ML estimate of $\boldsymbol{\theta}$ upon the perturbed tobit model. Then, the normal curvature for $\boldsymbol{\theta}$ in the direction $\boldsymbol{l}$, with $||\boldsymbol{l}|| = 1$, is expressed as

$$C_l(\boldsymbol{\theta}) = 2|\boldsymbol{l}^\top \boldsymbol{\Delta}^\top \ddot{\ell}^{-1} \boldsymbol{\Delta}\, \boldsymbol{l}|, \tag{12}$$

where $\boldsymbol{\Delta}$ is a $[p+1] \times n$ matrix of perturbations and $\ddot{\ell}$ is given in Eq. (6). The elements of $\boldsymbol{\Delta}$ are $\Delta_{ji} = \partial^2 \ell(\boldsymbol{\theta}|\boldsymbol{\omega})/\partial\theta_j \partial\omega_i$ evaluated at $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}$ and $\boldsymbol{\omega} = \boldsymbol{\omega}_0$, for $j = 1, \ldots, p+1$ and $i = 1, \ldots, n$. A local influence diagnostic is generally based on index plots. For example, the index graph of the eigenvector $\boldsymbol{l}_{\max}$ corresponding to the maximum eigenvalue of

$$\boldsymbol{B}(\boldsymbol{\theta}) = \boldsymbol{\Delta}^\top \ddot{\ell}^{-1} \boldsymbol{\Delta}, \tag{13}$$

say $C_{l_{\max}}(\boldsymbol{\theta})$, evaluated at $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}$, can revel those observations that under small perturbations exercise a great influence on $\mathrm{LD}(\boldsymbol{\omega})$.

3.3.2 Total local influence

In addition to the direction vector of maximum normal curvature, $\boldsymbol{l}_{\max}$, another direction of interest is $\boldsymbol{l}_i = \boldsymbol{e}_{in}$, which corresponds to the direction of the ith observation, where $\boldsymbol{e}_{in}$ is an $n \times 1$ vector of zeros with a value equal to one at the ith position, i.e., $\{\boldsymbol{e}_{in}, 1 \leq i \leq n\}$ is the canonical basis of $\mathbb{R}^n$. In this case, the normal curvature is given by $C_i(\boldsymbol{\theta}) = 2|b_{ii}|$, where b_{ii} is the ith diagonal element of $\boldsymbol{B}(\boldsymbol{\theta})$ given in Eq. (13), for $i = 1, \ldots, n$, evaluated at $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}$. Those cases when $C_i(\widehat{\boldsymbol{\theta}}) > 2\overline{C}(\widehat{\boldsymbol{\theta}})$, where $\overline{C}(\widehat{\boldsymbol{\theta}}) = \sum_{i=1}^n C_i(\widehat{\boldsymbol{\theta}})/n$, are considered as potentially influential. This procedure is called total local influence of the ith case; see Lesaffre and Verbeke (1998).

Remark 7 It is also possible to compute the normal curvature only for the regression coefficients, $C_l(\boldsymbol{\beta})$, or only for the scale parameter, $C_l(\sigma)$. For more details, see Galea et al. (2004).

3.3.3 Conformal local influence

Although the method based on normal curvature proposed by Cook (1986) can be of great utility, this possesses some inconvenience. For example, the normal curvature takes a unbounded value and, in addition, it is not invariant under scale uniform transformations. In order to solve these problems, Poon and Poon (1999) defined the conformal normal curvature in the direction $\boldsymbol{l}$ at the point $\boldsymbol{\omega}_0$ as

$$B_l(\boldsymbol{\theta}) = \frac{C_l(\boldsymbol{\theta})}{||2\boldsymbol{B}(\boldsymbol{\theta})||_{\mathcal{F}}}, \tag{14}$$

where $||\cdot||_{\mathcal{F}}$ denotes Fröbenius's norm defined as $||A||_{\mathcal{F}} = \sqrt{\mathrm{tr}(A^\top A)}$, with A being a matrix with appropriate dimensions, $C_l(\boldsymbol{\theta})$ is given in Eq. (12) and $\boldsymbol{B}(\boldsymbol{\theta})$ in Eq. (13). It is possible to prove that for any direction $\boldsymbol{l}$, $0 \leq B_l(\boldsymbol{\theta}) \leq 1$.

Remark 8 The conformal normal curvature allows us to compare normal curvatures in alternative models that do not require to be nested or belong to the same family or have the same support; see Osorio et al. (2007) and Barros et al. (2009). This influence measure possesses an one-to-one relationship with the normal curvature proposed by Cook (1986). The conformal influence measure assumes values in a bounded interval and it is invariant on a class of reparameterizations named conformal. This measure gives reference values that allow us to evaluate the magnitude of a certain curvature.

Specifically, based on Eq. (14), if $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of $2\boldsymbol{B}(\boldsymbol{\theta})$ with corresponding normalized eigenvectors $\boldsymbol{\gamma}_1, \ldots, \boldsymbol{\gamma}_n$, then the value of the conformal normal curvature for the ith case, $B_{l_i}(\boldsymbol{\theta})$, is equal to the normalized eigenvalue $\widetilde{\lambda}_i = \lambda_i/[\sum_{i=1}^n \lambda_i^2]^{1/2}$, for $i = 1, \ldots, n$. Since $\sum_{i=1}^n B_{l_i}^2(\boldsymbol{\theta}) = 1$, if the conformal normal curvature is the same for all the eigenvectors, then these values are equal to $1/\sqrt{n}$. Thus, this value can be used as reference to decide whether an eigenvector is influential or not. Poon and Poon (1999) also suggested to use a reference value to evaluate the effect of $B_{l_i}(\boldsymbol{\theta})$ at different levels. For this purpose, they defined the ith eigenvector as q-influential if $|B_{l_i}(\boldsymbol{\theta})| \geq q/\sqrt{n}$, establishing so an added contribution for the ith perturbation vector associated with the canonical basis of $\mathbb{R}^n$, which is called basic perturbation vector, for all the eigenvectors q-influential in $\mathbb{R}^n$. In this case, they defined $\zeta_i = |\widetilde{\lambda}_i|$ and obtained the ordered absolute values of the normalized eigenvalues in the following way: $\zeta_{\max} = \zeta_1 \geq \cdots \geq \zeta_k \geq q/\sqrt{n} > \zeta_{k+1} > \cdots > \zeta_n \geq 0$. Thus, the added contribution of the ith basic perturbation vector for all the eigenvectors q-influential in $\mathbb{R}^n$ is $m_i^{[q]} = \left[\sum_{j=1}^n a_{ij}^2 \zeta_i\right]^{1/2}$, for $i = 1, \ldots, n$, where a_{ij} is the jth coordinate of the eigenvector corresponding to ζ_i and $\sum_{i=1}^n a_{ij}^2 = 1$. Since $\sum_{i=1}^n [m_i^{[q]}]^2 = \sum_{i=1}^k \zeta_i$, if the contribution of all the basic perturbations is uniform, then each one of these vectors is equal to $\overline{m}^{[q]} = [\sum_{i=1}^k \zeta_i/n]^{1/2}$. This value is useful as reference to establish the relevance of the added contribution of the basic perturbation vector. The total contribution m_i and the conformal normal curvature $B_{l_i}(\boldsymbol{\theta})$ of the basic perturbation vector l_i are highly related. It is possible to show that $B_{l_i}(\boldsymbol{\theta}) = m_i^2$, for all $i = 1, \ldots, n$, and that, if the total contribution of all the $B_{l_i}(\boldsymbol{\theta})$ is uniform, then each one of these values is equal to $b = \operatorname{tr}\left([2\boldsymbol{B}(\boldsymbol{\theta})]^2\right) / \left[n\, [\operatorname{tr}\left([2\boldsymbol{B}(\boldsymbol{\theta})]^2\right)]^{1/2}\right]$. Poon and Poon (1999) proposed to utilize $\sqrt{2}\,\overline{m}^{[q]}$ and $2b$ as critical values for $m_i^{[q]}$ and $B_{l_i}(\boldsymbol{\theta})$, respectively.

3.3.4 Calculation of the curvatures

Consider the log-likelihood function given in Eq. (4). Next, for the indicated scheme, we obtain the respective perturbation matrix $\boldsymbol{\Delta}$ given in general in Eq. (12), which is already evaluated at the non-perturbed vector $\boldsymbol{\omega}_0$.

Case-weight perturbation Under this scheme, we want to evaluate whether the contributions of the observations with different weights affect the ML estimate of $\boldsymbol{\theta}$. This scheme is the most usual for evaluating the local influence in a model. In this scheme, the log-likelihood function of the perturbed tobit model is $\ell(\boldsymbol{\theta}|\boldsymbol{\omega}) = \sum_{i=1}^n \ell_i(\boldsymbol{\theta}|\omega_i) = \sum_{i=1}^n \omega_i \ell_i(\boldsymbol{\theta})$. Then, considering its derivative with respect to $\boldsymbol{\omega}^\top$, we obtain

$$\frac{\partial \ell(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \boldsymbol{\omega}^\top} = \sum_{i=1}^n \ell_i(\boldsymbol{\theta})\, \boldsymbol{e}_{in}^\top, \tag{15}$$

where $\boldsymbol{e}_{in}^{\top}$ is an $n \times 1$ vector defined in Sect. 3.3.2 and $\ell_i(\boldsymbol{\theta})$ is given in Eq. (4). Now, computing the derivative of Eq. (15) with respect to $\boldsymbol{\theta}$, we obtain the perturbation matrix $\boldsymbol{\Delta} = \sum_{i=1}^{n} \boldsymbol{h}_i \boldsymbol{e}_{in}^{\top} = [\boldsymbol{h}_1, \ldots, \boldsymbol{h}_n]$, where

$$\boldsymbol{h}_i = \frac{\partial \ell_i(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \begin{bmatrix} \dfrac{\partial \ell_i(\boldsymbol{\theta})}{\partial \boldsymbol{\beta}} \\ \dfrac{\partial \ell_i(\boldsymbol{\theta})}{\partial \sigma} \end{bmatrix} = \begin{cases} \begin{bmatrix} -\frac{1}{\sigma} W(\delta_i)\, \boldsymbol{x}_i^{\top} \\ -\frac{1}{\sigma} W(\delta_i)\, \delta_i \end{bmatrix}, & i = 1, \ldots, m; \\ \begin{bmatrix} \frac{1}{\sigma} \delta_i \boldsymbol{x}_i^{\top} \\ \frac{1}{\sigma}\left[\delta_i^2 - 1\right] \end{bmatrix}, & i = m+1, \ldots, n. \end{cases}$$

which must be evaluated at $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}$.

Scale perturbation Under this scheme, we want to know the behavior of the ML estimate of $\boldsymbol{\theta}$ when there is not homoscedasticity. In this case, the perturbation is given by $\varepsilon_i \sim \mathrm{N}(0, \sigma^2/\omega_i)$, for $i = 1, \ldots, n$, so that the log-likelihood function for the perturbed tobit model is $\ell(\boldsymbol{\theta}|\boldsymbol{\omega}) = \sum_{i=1}^{n} \ell_i(\boldsymbol{\theta}|\omega_i)$, where

$$\ell_i(\boldsymbol{\theta}|\omega_i) = \begin{cases} \log\left(\Phi(\sqrt{\omega_i}\, \delta_i)\right), & i = 1, \ldots, m; \\ \log\left(\sqrt{\omega_i}\right) - \frac{1}{2}\log(2\pi) - \log(\sigma) - \frac{\omega_i \delta_i^2}{2}, & i = m+1, \ldots, n. \end{cases} \tag{16}$$

Thus, considering the derivative of Eq. (16) with respect to $\boldsymbol{\omega}^{\top}$, we obtain

$$\frac{\partial \ell(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \boldsymbol{\omega}^{\top}} = \sum_{i=1}^{n} \frac{\partial \ell_i(\boldsymbol{\theta}|\omega_i)}{\partial \boldsymbol{\omega}^{\top}}, \tag{17}$$

where

$$\frac{\partial \ell_i(\boldsymbol{\theta}|\omega_i)}{\partial \boldsymbol{\omega}^{\top}} = \begin{cases} \frac{W(\sqrt{\omega_i}\, \delta_i)\, \delta_i}{2\sqrt{\omega_i}}\, \boldsymbol{e}_{in}^{\top}, & i = 1, \ldots, m; \\ \frac{1}{2\omega_i} \boldsymbol{e}_{in}^{\top} - \frac{1}{2}\delta_i^2\, \boldsymbol{e}_{in}^{\top}, & i = m+1, \ldots, n. \end{cases}$$

Now, computing the derivative of Eq. (17) with respect to $\boldsymbol{\theta}$, we obtain the perturbation matrix $\boldsymbol{\Delta}$ with elements

$$\frac{\partial^2 \ell_i(\boldsymbol{\theta}|\omega_i)}{\partial \boldsymbol{\beta}\, \partial \boldsymbol{\omega}^{\top}} = \begin{cases} -\left[W(\delta_i) + \delta_i W'(\delta_i)\right] \frac{1}{2\sigma} \boldsymbol{x}_i\, \boldsymbol{e}_{in}^{\top}, & i = 1, \ldots, m; \\ \frac{\delta_i}{\sigma} \boldsymbol{x}_i\, \boldsymbol{e}_{in}^{\top}, & i = m+1, \ldots, n; \end{cases}$$

and

$$\frac{\partial^2 \ell_i(\boldsymbol{\theta}|\omega_i)}{\partial \sigma\, \partial \boldsymbol{\omega}^{\top}} = \begin{cases} -\left[W(\delta_i) + \delta_i W'(\delta_i)\right] \frac{\delta_i}{2\sigma}\, \boldsymbol{e}_{in}^{\top}, & i = 1, \ldots, m; \\ \frac{\delta_i^2}{\sigma}\, \boldsymbol{e}_{in}^{\top}, & i = m+1, \ldots, n; \end{cases}$$

which must be evaluated at $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}$.

Response perturbation This scheme can be applied in several ways. Here, we use an additive perturbation in the tobit model given by

$$Y_{i\omega} = Y_i + \omega_i s_y, \quad i = 1, \ldots, n, \tag{18}$$

where s_y is a scale factor that could represent the standard deviation of the censored response. In this case, the perturbed log-likelihood function for the tobit model is $\ell(\boldsymbol{\theta}|\boldsymbol{\omega}) = \sum_{i=1}^{n} \ell_i(\boldsymbol{\theta}|\omega_i)$, where

$$\ell_i(\boldsymbol{\theta}|\omega_i) = \begin{cases} \log\left(\Phi(\delta_i)\right), & i = 1, \ldots, m; \\ -\frac{1}{2}\log(2\pi) - \log(\sigma) - \frac{1}{2}\delta_{i\omega}^2, & i = m+1, \ldots, n; \end{cases} \tag{19}$$

with $\delta_{i\omega}$ being expressed in a similar way to that given in Eq. (4) replacing y_i by $\mathrm{y}_{i\omega}$, with $y_{i\omega} = y_i + \omega s_y$, such as in Eq. (18). Thus, considering the derivative of Eq. (19) with respect to $\boldsymbol{\omega}^\top$, we obtain

$$\frac{\partial \ell(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \boldsymbol{\omega}^\top} = \sum_{i=1}^{n} \frac{\partial \ell_i(\boldsymbol{\theta}|\omega_i)}{\partial \boldsymbol{\omega}^\top}, \tag{20}$$

where

$$\frac{\partial \ell_i(\boldsymbol{\theta}|\omega_i)}{\partial \boldsymbol{\omega}^\top} = \begin{cases} \mathbf{0}_n^\top, & i = 1, \ldots, m; \\ -\frac{[y_i + \omega_i s_y - \boldsymbol{x}_i^\top \boldsymbol{\beta}]}{\sigma^2} s_y\, \boldsymbol{e}_{in}^\top, & i = m+1, \ldots, n; \end{cases}$$

with $\mathbf{0}_n$ being an $n \times 1$ vector of zeros. Now, computing the derivative of Eq. (20) with respect to $\boldsymbol{\theta}$, we obtain the perturbation matrix

$$\boldsymbol{\Delta} = \sum_{i=m+1}^{n} \begin{bmatrix} \frac{s_y}{\sigma^2}\boldsymbol{x}_i \\ \frac{2\delta_i}{\sigma^2} s_y \end{bmatrix} \boldsymbol{e}_{in}^\top,$$

which must be evaluated at $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}$.

Remark 9 We note that the censored part of Eq. (19) is identical to that of Eq. (4), since each censored response is replaced by a known value τ_y.

Predictor perturbation A predictor representing a continuous variable can be perturbed in several ways. Here, we use an additive perturbation given by

$$x_{it\omega} = x_{it} + s_x \omega_i, \quad i = 1, \ldots, n, \tag{21}$$

where s_x is a scale factor that could represent the standard deviation of the corresponding predictor of the tobit model. In this case, the perturbed log-likelihood function for the model is $\ell(\boldsymbol{\theta}|\boldsymbol{\omega}) = \sum_{i=1}^{n} \ell_i(\boldsymbol{\theta}|\omega_i)$, where

$$\ell_i(\boldsymbol{\theta}|\omega_i) = \begin{cases} \log(\Phi(\delta_{i\omega})), & i = 1, \ldots, m; \\ -\frac{1}{2}\log(2\pi) - \log(\sigma) - \frac{1}{2}\delta_{i\omega}^2, & i = m+1, \ldots, n; \end{cases} \tag{22}$$

with $\delta_{i\omega}$ being similarly given as in Eq. (4) replacing x_{it} by $x_{it\omega}$. In this case, the corresponding perturbation matrix is

$$\boldsymbol{\Delta} = \begin{bmatrix} \boldsymbol{\Delta}_\beta \\ \boldsymbol{\Delta}_\sigma \end{bmatrix}, \tag{23}$$

where $\boldsymbol{\Delta}_\beta$ is a $p \times n$ matrix with elements for $j \neq t$ and $j = t$, respectively, given by

$$\boldsymbol{\Delta}_{\beta_{ij}} = \begin{cases} \frac{\beta_t s_x}{\sigma^2} x_{ij} W'(\delta_i), & i=1, \ldots, m; \\ -\frac{\beta_t s_x x_{ij}}{\sigma^2}, & i=m+1, \ldots, n; \end{cases} \quad \boldsymbol{\Delta}_{\beta_{it}} = \begin{cases} \frac{\beta_t s_x}{\sigma^2} x_{it} W'(\delta_i) - \frac{s_x}{\sigma} W(\delta_i), & i=1, \ldots, m; \\ \frac{\delta_i s_x}{\sigma} - \frac{\beta_t s_x x_{it}}{\sigma^2}, & i=m+1, \ldots, n, \end{cases}$$

and $\boldsymbol{\Delta}_\sigma$ is an $1 \times n$ vector with elements

$$\boldsymbol{\Delta}_{\sigma_i} = \begin{cases} \frac{\beta_t s_x}{\sigma^2}[\delta_i W'(\delta_i) + W(\delta_i)], & i = 1, \ldots, m; \\ -\frac{2\beta_t s_x \delta_i}{\sigma^2}, & i = m+1, \ldots, n; \end{cases}$$

which must be evaluated at $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}$.

4 Application

In this section, we apply the obtained results using the real data presented and studied in Lee (1995, 1996, p. 256), where data on female labor supply are used to exemplify the use of several type of estimators. This data set contains $n = 200$ observations from the 1987 wave of the Michigan panel study of income dynamics (PSID). The labor supply is censored at 0 (zero) reaching a 22% of censoring (i.e., $m = 44$ individuals registered a value of labor supply equal to zero). The response (Y_i) is wife's labor supply in hours per year. The following predictors are used in Lee's model: AGE (x_1), wife's age; AGE^2 (x_2), wife's age squared; EDU (x_3), wife's schooling in years; INC (x_4), other household incomes; PKID (x_5), the number of preschool children; and SKID (x_6), the number of children between the ages of 6 and 12 years old.

4.1 Estimation and hypothesis testing

Based on the data set (censored and uncensored, that we have denominated set "A"), we obtain the parameter estimates of the tobit model postulated in Eq. (2). A summary of these parameter estimates, the values of the t-statistics for testing the hypotheses H$_0$: $\beta_j = 0$, for $j = 0, 1, \ldots, 6$, and their respective p-values are presented in Table 1. These results are coherent with those indicated in Lee (1996) and Santos (2001). From Table 1, we note that β_0, β_1 and β_4 are not significant at 5% level.

Table 1 Parameter estimates of the tobit model, t-values, and p-values for Lee's data

Parameter	β_0	β_1	β_2	β_3	β_4	β_5	β_6
Estimate	−6.902	5.781	−0.109	6.640	−0.293	−46.237	−16.932
t	−0.086	1.511	−2.375	2.425	−1.415	−5.100	−2.244
p-value	0.931	0.131	0.018	0.015	0.157	0.000	0.025

4.2 Influence diagnostics

Next, we carry out global and local influence diagnostics based on the Cook distance and local influence index plots. In addition, we also conducted a residuals analysis including their envelopes.

4.2.1 Residual analysis

In order to detect the presence of potentially influential observations as well as possible violations of the assumptions that validate the tobit model given in Eq. (2), we use the MT residual discussed in Sect. 3.2. We can note from Fig. 1a that there is not indication of some systematic trend in the residual plot, although observation #31 is highlighted as an extreme case. This observation corresponds to a women of 31 years old, that studied during 17 years, whose other household incomes are US$23,000 per year, with 1 preschool children, 2 children between the ages of 6 and 12 years old, and 4742 labor supply hours per year. All these observed values for the case #31 are very different of the respective average values of each of these variables, which are: 37.4 years old, 13.0 years, US$ 33,655.5, 0.42 and 0.46 children, and 1217.5 hours, respectively. The normal probability plot of the MT residuals with generated envelopes is presented in Fig. 1b. From this figure, we note that the assumption of normality is well supported by the data.

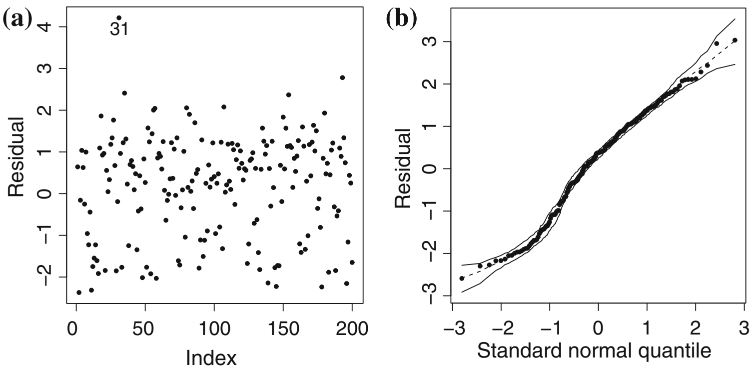

Fig. 1 Index plot (**a**) and normal QQ plot with envelopes (**b**) of MT residuals for Lee's data

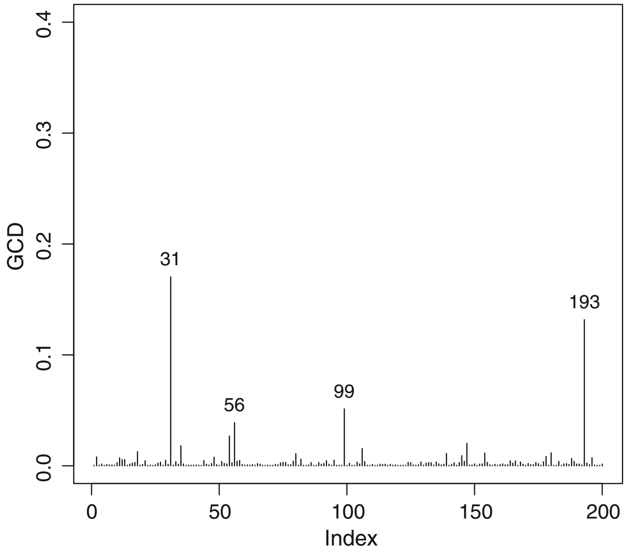

Fig. 2 Index plot of the GCD for Lee's data

4.2.2 Global influence

In order to evaluate the effect on the ML estimates of the regression coefficients of the tobit model when some observation is eliminated, we analyze the GCD index plot shown in Fig. 2. We note that four cases (#31, #56, #99, #193) are potentially influential on these estimates.

4.2.3 Conformal local influence

Based on the conformal normal curvature presented in Sect. 3.3.3 and the scheme of case-weight perturbation, we carry out influence diagnostics. The normalized eigenvalues different from zero are: 0.986, 0.174, 0.089, 0.054, 0.034, 0.028, 0.024, and 0.014. In order to determine which are the eigenvalues associated with the maximum variations of the likelihood displacement, we produce a graph of the normalized eigenvalues against the indexes of the observations, which is shown in Fig. 3a. From this figure, we note that for $q = 2$ only one eigenvalue is greater than the critical value $q/\sqrt{n}$, while for $q = 1$, it is possible to detect three cases that are greater than this critical value. Thus, there is one eigenvalue 2-influential and three 1-influential. The normalized eigenvalues associated with these eigenvalues are 0.986, 0.174, and 0.089. Therefore, the maximum conformal normal curvature $B_{l_{\max}}(\boldsymbol{\theta})$ is 0.986 and $\boldsymbol{l}_{\max}$ summarizes the maximum variations of $\mathrm{LD}(\boldsymbol{\omega})$ when the tobit model is perturbed.

Figure 3b shows the normalized eigenvalues of the added contribution of all the eigenvalues ($q = 0$). From this figure, we note that several cases are greater than the critical value suggested by Poon and Poon (1999), which value is here $2b = 0.0269$. These potentially influential cases are: #31, #35, #107, #154, and #193. The rest of

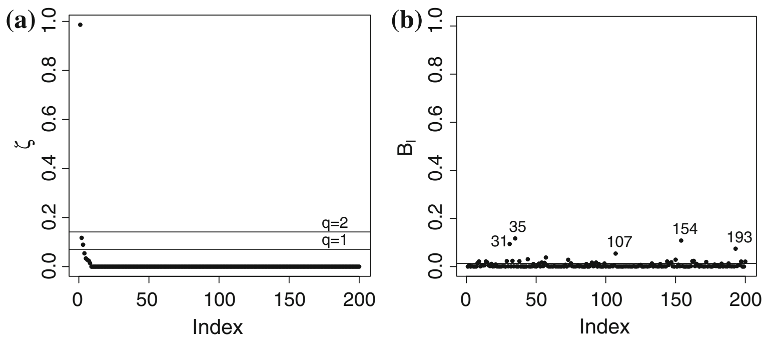

Fig. 3 Index plots of the normalized eigenvalues with the values of q (**a**) and the added contribution of all the eigenvalues (**b**) for Lee's data

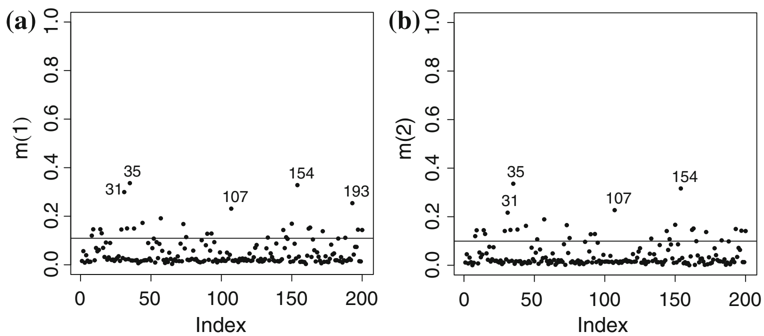

Fig. 4 Added contribution of all the eigenvalues for $q = 1$ (**a**) and $q = 2$ (**b**) for Lee's data

the observations (not highlighted in the plot) are considered as marginally influential since the difference between $B_{l_j}(\boldsymbol{\theta})$ and the critical value is near to zero.

Figure 4a and b show the added contribution of the eigenvalues associated with maximum eigenvalues corresponding to $q = 1$ and $q = 2$, respectively. From these figures, we observe the same cases detected in Fig. 3b.

In summary, global and local influence diagnostics detect as potentially influential the following seven cases: #31, #35, #56, #99, #107, #154, and #193. Three of these 7 cases (#56, #99, and #154) were also detected by the global influence analysis produced by Santos (2001).

In order to revel the impact of the seven observations that we have highlighted as potentially influential on the ML estimates of the parameters of the tobit model given in Eq. (2), we have refitted this model dropping each one of these cases. Table 2 presents the relative changes (RC) in percentage of these estimates defined by

$$\mathrm{RC}_{\beta_j} = \left| \frac{\widehat{\beta}_j - \widehat{\beta}_{j\,(\mathrm{I})}}{\widehat{\beta}_j} \right| \times 100\%, \quad j = 0, 1, \ldots, 6,$$

Table 2 RC (in %) with their respective p-values in the below line and LD for Lee's data

Set(I)	$\mathrm{RC}_{\widehat{\beta}_0}$	$\mathrm{RC}_{\widehat{\beta}_1}$	$\mathrm{RC}_{\widehat{\beta}_2}$	$\mathrm{RC}_{\widehat{\beta}_3}$	$\mathrm{RC}_{\widehat{\beta}_4}$	$\mathrm{RC}_{\widehat{\beta}_5}$	$\mathrm{RC}_{\widehat{\beta}_6}$	$\mathrm{RC}_{\widehat{\sigma}}$	$\mathrm{LD}_{\mathrm{I}}(\boldsymbol{\theta})$
A-{31}	378.22	1.13	0.11	27.84	22.76	4.87	26.02	5.36	1.90
	0.801	0.141	0.012	0.067	0.248	0.000	0.003		
A-{35}	352.94	12.46	9.63	14.68	0.52	0.46	0.80	0.66	0.39
	0.700	0.089	0.009	0.006	0.158	0.000	0.028		
A-{56}	185.46	15.69	9.85	3.47	4.64	1.45	27.51	0.57	0.46
	0.807	0.081	0.009	0.019	0.137	0.000	0.007		
A-{99}	118.10	6.84	3.67	4.05	53.61	2.51	3.28	0.06	0.62
	0.853	0.109	0.014	0.012	0.070	0.000	0.030		
A-{107}	158.50	10.90	8.31	1.33	5.68	0.06	1.02	0.38	0.14
	0.825	0.095	0.011	0.014	0.135	0.000	0.023		
A-{154}	233.60	18.64	14.20	2.04	8.88	0.65	3.83	0.70	0.33
	0.775	0.075	0.007	0.017	0.196	0.000	0.019		
A-{193}	258.38	12.33	9.14	11.79	3.72	21.79	13.88	1.41	1.71
	0.757	0.087	0.009	0.006	0.168	0.000	0.010		

where $\widehat{\beta}_{j(\mathrm{I})}$ denotes the ML estimate of β_j after the set I of observations has been removed. Table 2 also presents the respective p-values (for testing the hypotheses H_0: $\beta_j = 0$, for $j = 0, 1, \ldots, 6$) in the below line and the LD given by

$$\mathrm{LD}_{\mathrm{I}}(\boldsymbol{\theta}) = 2\left[\ell(\widehat{\boldsymbol{\theta}}) - \ell(\widehat{\boldsymbol{\theta}}_{(\mathrm{I})})\right],$$

where $\widehat{\boldsymbol{\theta}}_{(\mathrm{I})}$ denotes the ML estimates of $\boldsymbol{\theta}$ after the set I of observations has been removed; see Cook et al. (1988). From Table 2, we note that the RC are important when the potentially influential cases are dropped. However, there are not changes in the significance of the coefficients of the tobit model when these cases are eliminated.

5 Concluding remarks

In this work, we have developed influence diagnostic tools for the tobit censored normal response model. In particular, the normal curvature was derived as a measure of local influence under some perturbation schemes. We have also presented some ways to perform global influence measures based on the generalized Cook distance and residual analysis. A data set from the econometrics area has been reanalyzed from the perspective of the diagnostic tools developed in the paper.

Acknowledgments The authors wish to thank the editors and referees for their helpful comments that aided in improving this article. This study was partially supported by CNPq (Casadinho 620150/2008-4, Universal-477747/2008-6, PROSUL-490429/2008-4 and INCTMat-57.33523/2008-8), Brazil, and FONDECYT (1070919 and 1080326) and DIPUV 29-2006 grants, Chile.

References

Amemiya T (1984) Tobit models: a survey. J Economet 24:3–61
Barlow WE, Prentice RL (1988) Residuals for relative risk regression. Biometrika 75:65–74
Barros M, Paula GA, Leiva V (2008) A new class of survival regression models with heavy-tailed errors: robustness and diagnostics. Lifetime Data Anal 14:316–332
Barros M, Paula GA, Leiva V (2009) An R implementation for generalized Birnbaum-Saunders distributions. Comp Stat Data Anal 53:1511–1528
Belsley D, Kuh E, Welsch R (1980) Regressions diagnostics: identifying influential data and sources collinearity. Wiley, New York
Chatterjee S, Hadi AS (1988) Sensitivity analysis in linear regression. Wiley, New York
Cohen AC (1991) Truncated and censored samples: theory and applications. Marcel Dekker, New York
Cook RD (1977) Detection of influential observation in linear regression. Technometrics 19:15–18
Cook RD (1986) Assessment of local influence. J R Stat Soc B 48:133–169
Cook RD, Weisberg S (1982) Residuals and influence in regression. Chapman and Hall, London
Cook RD, Peña D, Weisberg S (1988) The likelihood displacement: a unifying principle for influence measures. Comm Stat Theor Meth 17:623–640
Davison A, Gigli A (1989) Deviance residuals and normal scores plots. Biometrika 76:211–221
Díaz-García JA, Galea M, Leiva V (2003) Influence diagnostics for elliptical multivariate linear regression models. Comm Stat Theor Meth 32:625–641
Efron B, Hinkley D (1978) Assessing the accuracy of the maximum likelihood estimator: observed vs. expected Fisher information. Biometrika 65:457–487
Escobar LA, Meeker WQ (1992) Assessing influence in regression analysis with censored data. Biometrika 48:507–528
Galea M, Paula GA, Bolfarine H (1997) Local influence in elliptical linear regression models. Statistician 46:71–79
Galea M, Riquelme M, Paula GA (2000) Diagnostics methods in elliptical linear regression models. Braz J Prob Stat 14:167–184
Galea M, Leiva V, Paula GA (2004) Influence diagnostics in log-Birnbaum-Saunders regression models. J Appl Stat 31:1049–1064
Klein JP, Moeschberger ML (1997) Survival analysis: techniques for censored and truncated data. Springer, New York
Lawrence AJ (1998) Regression transformation diagnostics using local influence. J Am Stat Soc 84:125–141
Lee MJ (1995) A semiparametric estimation of simultaneous equations with limited dependent variables: a case study of female labor supply. J Appl Econ 10:187–200
Lee MJ (1996) Methods of moments and semiparametric econometrics for limited dependent variable models. Springer, New York
Leiva V, Barros M, Paula GA, Galea M (2007) Influence diagnostics in log-Birnbaum-Saunders regression models with censored data. Comp Stat Data Anal 51:5694–5707
Lesaffre E, Verbeke G (1998) Local influence in linear mixed models. Biometrics 54:570–582
McCullagh P, Nelder JA (1989) Generalized linear models. Chapman and Hall, London
Ortega EM, Bolfarine H, Paula GA (2003) Influence diagnostics in generalized log-gamma regression models. Comp Stat Data Anal 42:165–186
Osorio F, Paula GA, Galea M (2007) Assessment of local influence in elliptical linear models with longitudinal structure. Comp Stat Data Anal 51:4354–4368
Paula GA (1993) Assessing local influence in restricted regression models. Comp Stat Data Anal 16:63–79
Paula GA, Medeiros MJ, Vilca-Labra F (2009) Influence diagnostics for linear models with first autoregressive elliptical errors. Stat Prob Lett 79:339–346
Poon WY, Poon YS (1999) Conformal normal curvature and assessment of local influence. J R Stat Soc B 61:51–61
Santos JMC (2001) Influence diagnostics and estimation algorithms for Powell's SCLS. J Bus Econ Stat 19:55–62
Schneider H (1986) Truncated and censored samples from normal populations. Marcel Dekker, New York
Scott JL (1997) Regression models of categorical and limited dependent variables. Sage, California
Therneau TM, Grambsch PM, Fleming TR (1990) Martingale-based residuals for survival models. Biometrika 77:147–160

Tobin J (1958) Estimation of relationships for limited dependent variables. Econometrica 26:24–36
Xie FC, Wei BC (2007) Diagnostics analysis for log-Birnbaum-Saunders regression models. Comp Stat Data Anal 51:4692–4706
Zhu H, Zhang H (2004) A diagnostic procedure based on local influence. Biometrika 91:579–589