Stat Papers (2012) 53:665–683
DOI 10.1007/s00362-011-0371-8

REGULAR ARTICLE

On diagnostics in multivariate measurement error models under asymmetric heavy-tailed distributions

Camila B. Zeller · Rignaldo R. Carvalho · Victor H. Lachos

Received: 25 May 2010 / Revised: 2 February 2011 / Published online: 25 February 2011

Abstract In this paper, we discuss the extension of some diagnostic procedures to multivariate measurement error models with scale mixtures of skew-normal distributions (Lachos et al., Statistics 44:541–556, 2010c). This class provides a useful generalization of normal (and skew-normal) measurement error models since the random term distributions cover symmetric, asymmetric and heavy-tailed distributions, such as skew-t, skew-slash and skew-contaminated normal, among others. Inspired by the EM algorithm proposed by Lachos et al. (Statistics 44:541–556, 2010c), we develop a local influence analysis for measurement error models, following Zhu and Lee's (J R Stat Soc B 63:111–126, 2001) approach. This is because the observed data log-likelihood function associated with the proposed model is somewhat complex and Cook's well-known approach can be very difficult to apply to achieve local influence measures. Some useful perturbation schemes are also discussed. In addition, a score test for assessing the homogeneity of the skewness parameter vector is presented. Finally, the methodology is exemplified through a real data set, illustrating the usefulness of the proposed methodology.

Keywords EM algorithm · Local influence · Mahalanobis distance · Measurement error models · Scale mixtures of skew-normal distributions

C. B. Zeller (✉)
Department of Statistics, Federal University of Juiz de Fora, Juiz de Fora, Minas Gerais, CEP 36036-330, Brazil
e-mail: camila.zeller@ufjf.edu.br

R. R. Carvalho · V. H. Lachos
Department of Statistics, Campinas State University, Campinas, São Paulo, CEP 13083-859, Brazil

1 Introduction

Multivariate measurement error models (MEM) are useful concepts in many disciplines, including linear and nonlinear errors-in-variables regression models, factor analysis models, latent structural models and simultaneous equations models, among others. MEM have also been extensively used in the problem of comparing measurement devices (Vilca-Labra et al. 2010) for which the random terms are assumed to follow a normal distribution. However, normal MEM (N-MEM) suffers from the same lack of robustness against departures from distributional assumptions as other statistical models based on the Gaussian distribution and may be too restrictive to provide an accurate representation of the structure that is present in the data. To overcome this deficiency, some proposals have been made in the literature involving replacing the assumption of normality by more flexible classes of distributions. For instance, Montenegro et al. (2010) showed the advantage of using the skew-normal distribution in the context of the MEM; Lachos et al. (2010a) proposed using scale mixtures of normal in the context of MEM (SMN-MEM) and showed that it performed well in the presence of outliers; Patriota and Bolfarine (2010) adopted a general class of error distribution in MEM. They proposed a simple method for obtaining consistent estimators, based on the corrected score approach. More recently, Lachos et al. (2010c) have introduced MEM with scale mixtures of skew-normal distributions (SMSN-MEM) and presented a complete likelihood based analysis, including an efficient EM algorithm for maximum likelihood estimation.

The assessment of robustness aspects of the parameter estimates in statistical models has been an important concern of various researchers in recent decades. The deletion method, which consists of studying the impact on the parameter estimates after dropping individual observations, is probably the most employed technique to detect influential observations (see Cook and Weisberg 1982). Nevertheless, research on the influence of small perturbations in the model/data on the parameter estimates has received increasing attention in recent years. This can be achieved by performing local influence analysis (Cook 1986), a general statistical technique used to assess the stability of the estimation outputs with respect to the model inputs. On the other hand, it is a standard assumption for MEM that all the observations have equal variances. The violation of this assumption can have adverse consequences for the efficiency of estimators (Cook and Weisberg 1982), so it is important to detect the variance heterogeneity in MEM. Motived by the work of Lachos et al. (2010c), in this paper we discuss the local influence analysis and a score test statistic for assessing homogeneity of the skewness parameter, which is a parameter included in the variance, in SMSN-MEM. Since the observed log-likelihood function of SMSN models involves some integrals, a direct application of Cook's (1986) local influence approach can be very difficult to apply, because these measures involve the first and second partial derivatives of the log-likelihood function. Instead, we discuss the local influence analysis for SMSN-MEM based on Zhu and Lee's (2001) approach, by working with a Q-displacement function, which is closely related to the conditional expectation of the complete-data log-likelihood at the E-step of the EM algorithm. The results of this paper are a necessary supplement to those presented in Lachos et al. (2010c).

This paper is organized as follows. In Sect. 2, we present the SMSN-MEM introduced by Lachos et al. (2010c), including the EM algorithm for maximum likelihood (ML) estimation. In Sect. 3, we introduce the local influence approach for models with incomplete-data and develop the methodology required for the SMSN-MEM. Four different perturbation schemes are considered. In Sect. 4, we develop a score test statistic for assessing the homogeneity of skewness parameter in SMSN-MEM. The methodology is illustrated in Sect. 5 using the famous data set of Chipkevitch et al. (1996), in which MEM under skew-normal and asymmetric heavy-tailed distributions are compared according to the robustness aspects of the maximum likelihood estimates. Finally, we give some concluding remarks in Sect. 6.

2 The SMSN-MEM

In order to introduce some notations, we start with the definition of SMSM distributions. Details of this section are provided in Lachos et al. (2010c). A $p \times 1$ random vector $\mathbf{Y}$ follows a SMSN distribution with a $p \times 1$ location vector $\boldsymbol{\mu}$, a $p \times p$ positive definite dispersion matrix $\boldsymbol{\Sigma}$, and a $p \times 1$ skewness parameter vector $\boldsymbol{\lambda}$, denoted by $\mathbf{Y} \sim SMSN_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\lambda}; H)$, if its probability density function (pdf) is given by

$$f(\mathbf{y}) = 2 \int_0^\infty \phi_p(\mathbf{y}|\boldsymbol{\mu}, \kappa(u)\boldsymbol{\Sigma})\Phi(\kappa^{-1/2}(u)\mathrm{A})dH(u; \boldsymbol{\nu}), \quad \mathbf{y} \in \mathbb{R}^p, \tag{1}$$

where $\mathrm{A} = \boldsymbol{\lambda}^\top \boldsymbol{\Sigma}^{-1/2}(\mathbf{y} - \boldsymbol{\mu})$, $\phi_p(\cdot|\boldsymbol{\mu}, \boldsymbol{\Sigma})$ stands for the pdf of the p-variate normal distribution with mean vector $\boldsymbol{\mu}$ and covariate matrix $\boldsymbol{\Sigma}$, $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ say, $\Phi(\cdot)$ represents the cumulative distribution function (cdf) of the standard univariate normal distribution, $\kappa(\cdot)$ is a positive weight function and U is a positive random variable with a cdf $H(u; \boldsymbol{\nu})$, where $\boldsymbol{\nu}$ is a scalar or parameter vector indexing the distribution of U. One particular case of this distribution is the skew-normal distribution (Azzalini and Dalla-Valle 1996), for which H is degenerate, with $\kappa(u) = 1$, $u > 0$. Also, when $\boldsymbol{\lambda} = \mathbf{0}$, the SMSN distribution reduces to the scale mixtures of normal distribution (SMN) (Andrews and Mallows 1974). The asymmetrical class of SMSN distributions includes many distributions such as the skew-normal, the skew-t, the skew-slash and the skew-contaminated normal, as special cases. Note that the term $\phi_p(\cdot)$ in (1) depend on the Mahalanobis distance $d = (\mathbf{y} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\mathbf{y} - \boldsymbol{\mu})$ and in practice it provides useful diagnostic statistics for identifying sample units with outlying observations, as pointed by Pinheiro et al. (2001). We refer readers to Lachos et al. (2010b) for details and additional properties related to this class of distributions.

Following Lachos et al. (2010c), we write the MEM as

$$\mathbf{Z}_i = \mathbf{a} + \mathbf{b}x_i + \boldsymbol{\epsilon}_i, \tag{2}$$

where $\mathbf{Z}_i = (X_i, \mathbf{Y}_i^\top)^\top$, $\mathbf{a} = (0, \boldsymbol{\alpha}^\top)^\top$, $\mathbf{b} = (1, \boldsymbol{\beta}^\top)^\top$ and $\boldsymbol{\epsilon}_i = (u_i, \mathbf{e}_i^\top)^\top$ are $p \times 1$ vectors, with $\mathbf{Y}_i = (y_{i1}, \ldots, y_{ir})^\top$ the vector of responses for the i-th experimental unit, X_i the observed value of the covariate x_i, $\mathbf{e}_i = (e_{i1}, \ldots, e_{ir})^\top$ a random vector of measurement errors of dimension r and $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_r)^\top$ and $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_r)^\top$ parameter vectors of dimension $r = p-1$. The structural SMSN-MEM are defined as

$$\begin{bmatrix} x_i \\ \boldsymbol{\epsilon}_i \end{bmatrix} \overset{\text{iid}}{\sim} \text{SMSN}_{p+1}\left(\begin{pmatrix} \mu_x \\ \mathbf{0} \end{pmatrix}, D(\phi_x, \boldsymbol{\phi}), \begin{pmatrix} \lambda_x \\ \mathbf{0} \end{pmatrix}; H\right), \quad i = 1, \ldots, n, \tag{3}$$

where $D(\phi_x, \boldsymbol{\phi}) = diag(\phi_x, \phi_1, \ldots, \phi_p)$, with $\boldsymbol{\phi} = (\phi_1, \ldots, \phi_p)$. From Proposition 1 of Lachos et al. (2010c), it can be shown that

$$\mathbf{Z}_i \overset{\text{iid}}{\sim} \text{SMSN}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \bar{\boldsymbol{\lambda}}_x; H), \quad i = 1, \ldots, n, \tag{4}$$

where $\boldsymbol{\mu} = \mathbf{a} + \mathbf{b}\mu_x, \boldsymbol{\Sigma} = \phi_x \mathbf{b}\mathbf{b}^\top + D(\boldsymbol{\phi})$ and $\bar{\boldsymbol{\lambda}}_x = \frac{\lambda_x \phi_x \boldsymbol{\Sigma}^{-1/2}\mathbf{b}}{\sqrt{\phi_x + \lambda_x^2 \Lambda_x}}$, with $\Lambda_x = \phi_x / c$ and $c = 1 + \phi_x \mathbf{b}^\top D^{-1}(\boldsymbol{\phi})\mathbf{b}$.

As the observed log-likelihood function $\ell(\boldsymbol{\theta}|\mathbf{z})$ involves complex expressions, it is very difficult to work directly with $\ell(\boldsymbol{\theta}|\mathbf{z})$, either using ML estimation or in the context of local influence analysis. For the SMSM-MEM, an EM algorithm has been developed by Lachos et al. (2010c) to perform the ML estimation. In their estimation procedure $\mathbf{x}$, $\mathbf{u}$ and $\mathbf{t}$ are treated as hypothetical missing data and the vector of complete data are given by $\mathbf{z}_c = (\mathbf{z}, \mathbf{x}, \mathbf{t}, \mathbf{u})$, where $\mathbf{z} = (\mathbf{z}_1^\top, \ldots, \mathbf{z}_n^\top)^\top$ corresponding to the vector of observable response for the n units, $\mathbf{x} = (x_1, \ldots, x_n)^\top$, $\mathbf{u} = (u_1, \ldots, u_n)^\top$ and $\mathbf{t} = (t_1, \ldots, t_n)^\top$. The EM algorithm is applied to the complete-data log-likelihood function $\ell_c(\boldsymbol{\theta}|\mathbf{z}_c)$ (see Lachos et al. 2010c), where $\boldsymbol{\theta} = (\boldsymbol{\theta}_1^\top, \boldsymbol{\theta}_2^\top)^\top$, with $\boldsymbol{\theta}_1 = (\boldsymbol{\alpha}^\top, \boldsymbol{\beta}^\top, \boldsymbol{\phi}^\top)^\top$ and $\boldsymbol{\theta}_2 = (\mu_x, \phi_x, \lambda_x)^\top$. Let $\widehat{\boldsymbol{\theta}}^{(k)}$ denote the estimates of $\boldsymbol{\theta}$ at the k-th iteration. Given the current estimate $\widehat{\boldsymbol{\theta}}^{(k)}$, the E-step calculates $Q(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}}^{(k)}) = \sum_{i=1}^n Q_{1i}(\boldsymbol{\theta}_1|\widehat{\boldsymbol{\theta}}^{(k)}) + \sum_{i=1}^n Q_{2i}(\boldsymbol{\theta}_2|\widehat{\boldsymbol{\theta}}^{(k)})$, with

$$\begin{aligned} Q_{1i}(\boldsymbol{\theta}_1|\widehat{\boldsymbol{\theta}}^{(k)}) = &-\frac{1}{2}\log(|D(\boldsymbol{\phi})|) - \frac{1}{2}\Big[\widehat{u}_i^{(k)}(\mathbf{z}_i - \mathbf{a})^\top D^{-1}(\boldsymbol{\phi})(\mathbf{z}_i - \mathbf{a}) \\ &-2\widehat{ux}_i^{(k)}(\mathbf{z}_i - \mathbf{a})^\top D^{-1}(\boldsymbol{\phi})\mathbf{b} + \widehat{ux^2}_i^{(k)}\mathbf{b}^\top D^{-1}(\boldsymbol{\phi})\mathbf{b}\Big], \end{aligned} \tag{5}$$

$$\begin{aligned} Q_{2i}(\boldsymbol{\theta}_2|\widehat{\boldsymbol{\theta}}^{(k)}) = &-\frac{1}{2}\log(\nu_x^2) - \frac{1}{2\nu_x^2}\Big[\widehat{ux^2}_i^{(k)} + \mu_x^2\widehat{u}_i^{(k)} + \tau_x^2\widehat{ut^2}_i^{(k)} \\ &-2\widehat{ux}_i^{(k)}\mu_x - 2\widehat{utx}_i^{(k)}\tau_x + 2\mu_x\tau_x\widehat{ut}_i^{(k)}\Big], \end{aligned} \tag{6}$$

where $\nu_x^2 = \phi_x(1 - \delta_x^2)$, $\tau_x = \phi_x^{1/2}\delta_x$, with $\delta_x = \lambda_x/(1 + \lambda_x^2)^{1/2}$ and $\widehat{u}_i^{(k)}, \widehat{ut}_i^{(k)}$, $\widehat{ut_i^2}^{(k)}, \widehat{ux}_i^{(k)}, \widehat{ux^2}_i^{(k)}$ and $\widehat{utx}_i^{(k)}$ are all as given in Lachos et al. (2010c). As recommended by Lucas (1997), who pointed out difficulties in estimating $\boldsymbol{\nu}$ due to problems of unbounded and local maxima the likelihood function, we consider the value of $\boldsymbol{\nu}$ to be known. The motivation for employing the EM algorithm is that it can be used advantageously to obtain closed-form equations for the M-step. As our main focus is not on ML estimation, we refer the interested reader to see Lachos et al. (2010c) for a detailed discussion of the EM algorithm. They also considered empirical Bayes estimates for the latent variable (see Lachos et al. 2010c, Eq. 16). Hence, replacing $\boldsymbol{\theta}$ e x_i with their current estimates, we obtain the following decomposition for the Mahalanobis distance, $d_i = d_i(\boldsymbol{\theta}) = (\mathbf{z}_i - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\mathbf{z}_i - \boldsymbol{\mu})$, under the class of SMSN distributions

$$\widehat{d}_i = d_i(\widehat{\boldsymbol{\theta}}) = (\mathbf{z}_i - \widehat{\boldsymbol{\mu}})^\top \widehat{\boldsymbol{\Sigma}}^{-1}(\mathbf{z}_i - \widehat{\boldsymbol{\mu}}) = \widehat{d}_{\mathbf{e}i} + \widehat{d}_{xi}, \tag{7}$$

where $\widehat{d}_{\mathbf{e}i} = \widehat{\mathbf{e}}_i^\top D^{-1}(\widehat{\boldsymbol{\phi}})\widehat{\mathbf{e}}_i$, $\widehat{d}_{xi} = \frac{1}{\widehat{\phi}_x}\widehat{\mu}_{xi}^2$ and $\widehat{\mathbf{e}}_i = \mathbf{z}_i - \widehat{\boldsymbol{\mu}} - \widehat{\mathbf{b}}\widehat{\mu}_{xi}$ is the estimated residual, with $\widehat{\mu}_{xi} = \widehat{\Lambda}_x \widehat{a}_i$ and $\widehat{a}_i = (\mathbf{z}_i - \widehat{\boldsymbol{\mu}})^\top D^{-1}(\widehat{\boldsymbol{\phi}})\widehat{\mathbf{b}}$. Equation 7 provides a simple way to compute the Mahalanobis distance. The estimated distances d_i, d_{x_i} and $d_{\mathbf{e}i}$ provide useful diagnostic statistics for identifying sample units with outlying observations. Thus, the SMSN-MEM given in (2)–(3) considers two sources of variation, one due to an error component and the other due to a latent variable, which may be sensitive to outliers (Zeller et al. 2011).

From Branco and Dey (2001), we have that $d_i \sim \chi_p^2$ (the chi-square with p degrees of freedom), for the skew-normal (SN) case and thus, we can use as cutoff points the quantile $\upsilon = \chi_p^2(\xi)$, where $0 < \xi < 1$, to identify outliers. Also, it can be shown that the distribution of d_i is the same as that under SMN distributions (see Branco and Dey 2001; Proposition 5.2). From Zeller et al. (2011), we have $d_i \sim pF(p, \nu)$ for skew-t (ST) case, consequently $Pr(d_i \leq \upsilon) = Pr(\chi_p^2 \leq \upsilon) - \frac{2^\nu \Gamma(\nu + p/2)}{\upsilon^\nu \Gamma(p/2)} Pr(\chi_{2\nu+p}^2 \leq \upsilon)$ for skew-slash (SSL) and $Pr(d_i \leq \upsilon) = \nu Pr(\chi_p^2 \leq \gamma\upsilon) + (1-\nu) Pr(\chi_p^2 \leq \upsilon)$ for skew-contaminated normal (SCN). In the next section, we discuss influence diagnostics with emphasis on the local influence approach proposed by Zhu and Lee (2001).

3 Influence diagnostics

Cook (1986) proposed a unified approach for assessment of local influence in minor perturbations of a statistical model, which can be viewed as a generalization of the robustness concept to study and detect the influential subsets of data. Since a direct application of this approach involves extensive algebraic manipulation for SMSN-MEM, in this article we will apply the general approach of Zhu and Lee (2001) to achieve diagnostic measures for local influence analysis.

3.1 Local influence

Consider a perturbation vector $\boldsymbol{\omega} = (\omega_1, \ldots, \omega_g)^\top$ varying in an open region $\boldsymbol{\Omega} \subset \mathbb{R}^g$. Let $\ell_c(\boldsymbol{\theta}, \boldsymbol{\omega}|\mathbf{z}_c)$, $\boldsymbol{\theta} \in \mathbb{R}^h$, be the complete-data log-likelihood of the perturbed model. We assume there is a $\boldsymbol{\omega}_0$ such that $\ell_c(\boldsymbol{\theta}, \boldsymbol{\omega}_0|\mathbf{z}_c) = \ell_c(\boldsymbol{\theta}|\mathbf{z}_c)$ for all $\boldsymbol{\theta}$. Let $\widehat{\boldsymbol{\theta}}(\boldsymbol{\omega})$ denote the maximum of the function $Q(\boldsymbol{\theta}, \boldsymbol{\omega}|\widehat{\boldsymbol{\theta}}) = \mathrm{E}[\ell_c(\boldsymbol{\theta}, \boldsymbol{\omega}|\mathbf{z}_c)|\mathbf{z}, \widehat{\boldsymbol{\theta}}]$. The influence graph is defined as $\boldsymbol{\alpha}(\boldsymbol{\omega}) = (\boldsymbol{\omega}^\top, f_Q(\boldsymbol{\omega}))^\top$, where $f_Q(\boldsymbol{\omega})$ is the Q-displacement function, defined as $f_Q(\boldsymbol{\omega}) = 2\left[Q\left(\widehat{\boldsymbol{\theta}}|\widehat{\boldsymbol{\theta}}\right) - Q\left(\widehat{\boldsymbol{\theta}}(\boldsymbol{\omega})|\widehat{\boldsymbol{\theta}}\right)\right]$. Following the approach developed in Cook (1986) and Zhu and Lee (2001), the normal curvature $C_{f_Q,\mathbf{d}}$ of $\boldsymbol{\alpha}(\boldsymbol{\omega})$ at $\boldsymbol{\omega}_0$ in the direction of some unit vector $\mathbf{d}$ can be used to summarize the local behavior of the Q-displacement function. It can be shown that (Zhu and Lee 2001) $C_{f_Q,\mathbf{d}} =$

$-2\mathbf{d}^\top \ddot{Q}_{\boldsymbol{\omega}_o}\mathbf{d}$ and $-\ddot{Q}_{\boldsymbol{\omega}_0} = \boldsymbol{\Delta}_{\boldsymbol{\omega}_o}^\top \left\{-\ddot{Q}(\widehat{\boldsymbol{\theta}})\right\}^{-1} \boldsymbol{\Delta}_{\boldsymbol{\omega}_0}$, where $\ddot{Q}(\widehat{\boldsymbol{\theta}}) = \frac{\partial^2 Q(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}})}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}^\top}|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}}$ and $\boldsymbol{\Delta}_{\boldsymbol{\omega}} = \frac{\partial^2 Q(\boldsymbol{\theta},\boldsymbol{\omega}|\widehat{\boldsymbol{\theta}})}{\partial\boldsymbol{\theta}\partial\boldsymbol{\omega}^\top}|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}(\boldsymbol{\omega})}$.

As in Cook (1986), the quantity $-\ddot{Q}_{\boldsymbol{\omega}_0}$ is fundamental to detect influential observations. The assessment of influential cases is based on visual inspection of the $\{M(0)_l = B_{f_Q,\mathbf{u}_l}, l = 1,\ldots,g\}$ plotted against the index l, where $B_{f_Q,\mathbf{u}}(\boldsymbol{\theta}) = C_{f_Q,\mathbf{u}}(\boldsymbol{\theta})/tr[-2\ddot{Q}_{\boldsymbol{\omega}_0}]$ and $\mathbf{u}_l$ is a basic perturbation vector with lth entry 1 and zero elsewhere. So far there is not a general rule to judge the largeness of the influence of a specific case in the data. Recently, Lee and Xu (2004) proposed using $\overline{M}(0)+c^*SM(0)$, where c^* is a selected constant. Depending on the real application, c^* may be taken as any positive value, as discussed by Montenegro et al. (2009). In this paper, we consider $c^* = 3$ unless stated otherwise.

In the following subsections, we derive the normal curvature for the proposed SMSN-MEM. We compute $\ddot{Q}(\boldsymbol{\theta})$ and $\boldsymbol{\Delta}_{\boldsymbol{\omega}}$ by using the results of matrix differentiation described in Magnus and Neudecker (1988).

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

To obtain the diagnostic measures for local influence of a particular perturbation scheme, it is necessary to compute $\ddot{Q}(\widehat{\boldsymbol{\theta}}) = \frac{\partial^2 Q(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}})}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}^\top}$, where $\boldsymbol{\theta} = (\boldsymbol{\theta}_1^\top, \boldsymbol{\theta}_2^\top)^\top$ is the parameter vector. Hence, the Hessian matrix is given by $\ddot{Q}(\widehat{\boldsymbol{\theta}}) = \sum_{i=1}^n \ddot{Q}_i(\boldsymbol{\theta})$ with

$$\ddot{Q}_i(\boldsymbol{\theta}) = -\frac{\partial^2 Q_i(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}})}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}^\top} = \begin{pmatrix} \ddot{Q}_{11,i}(\boldsymbol{\theta}_1) & \mathbf{0} \\ \mathbf{0} & \ddot{Q}_{22,i}(\boldsymbol{\theta}_2) \end{pmatrix},$$

where $\ddot{Q}_{11,i}(\boldsymbol{\theta}_1) = -\partial^2 Q_{1i}(\boldsymbol{\theta}_1|\widehat{\boldsymbol{\theta}})/\partial\boldsymbol{\theta}_1\partial\boldsymbol{\theta}_1^\top$ and $\ddot{Q}_{22,i}(\boldsymbol{\theta}_2) = -\partial^2 Q_{2i}(\boldsymbol{\theta}_2|\widehat{\boldsymbol{\theta}})/\partial\boldsymbol{\theta}_2\partial\boldsymbol{\theta}_2^\top$ are given in Appendix A.

3.3 Perturbation schemes

In this subsection, some special perturbation are considered, which are inherent to the multivariate measurement error models. We consider three different perturbation schemes for the baseline model defined in (2)–(3). We evaluate in the sequel the matrix $\boldsymbol{\Delta}_{\boldsymbol{\omega}}$ under the following perturbation schemes for SMSN-MEM: *perturbation of case weights* for detecting observations with outstanding contribution on the log-likelihood function and that may exercise high influence on the maximum likelihood estimates; *perturbation of the measurement for one instrument* made on the observed values from the instruments used in the study, which may indicate observations with large influence on their own predicted values; and *perturbation of the variance* through *perturbation of the scale matrix* $\boldsymbol{\Sigma}$, which may reveal individuals that are most influential, in the sense of the likelihood displacement on the scale structure and consequently on the estimates of ϕ_x and $\boldsymbol{\phi}$. Before we derive the appropriate matrices for assessing the local influence for perturbation schemes, we note that the matrix $\ddot{Q}(\widehat{\boldsymbol{\theta}})$ is block-diagonal with block $\ddot{Q}_{11}(\widehat{\boldsymbol{\theta}}_1)$ and $\ddot{Q}_{22}(\widehat{\boldsymbol{\theta}}_2)$, so that for any unit vector $\mathbf{d}$,

$$C_{f_Q,\mathbf{d}} = C_{f_Q,\mathbf{d}}(\boldsymbol{\theta}_1) + C_{f_Q,\mathbf{d}}(\boldsymbol{\theta}_2),$$

where

$$C_{f_Q,\mathbf{d}}(\boldsymbol{\theta}_1) = 2\mathbf{d}^\top \boldsymbol{\Delta}_{1\boldsymbol{\omega}_o}^\top (-\ddot{Q}_{11})^{-1} \Delta_{1\boldsymbol{\omega}_o}\mathbf{d}$$

and

$$C_{f_Q,\mathbf{d}}(\boldsymbol{\theta}_2) = 2\mathbf{d}^\top \boldsymbol{\Delta}_{2\boldsymbol{\omega}_o}^\top (-\ddot{Q}_{22})^{-1} \boldsymbol{\Delta}_{2\boldsymbol{\omega}_o}\mathbf{d},$$

with $\boldsymbol{\Delta}_{1\boldsymbol{\omega}_0} = (\boldsymbol{\Delta}_{11\boldsymbol{\omega}_0}, \ldots, \boldsymbol{\Delta}_{1n\boldsymbol{\omega}_0})^\top$ and $\boldsymbol{\Delta}_{2\boldsymbol{\omega}_0} = (\boldsymbol{\Delta}_{21\boldsymbol{\omega}_0}, \ldots, \boldsymbol{\Delta}_{2n\boldsymbol{\omega}_0})^\top$, respectively, where $\boldsymbol{\Delta}_{1i\boldsymbol{\omega}_0} = \frac{\partial^2 Q_{1i}(\boldsymbol{\theta}_1,\boldsymbol{\omega}|\widehat{\boldsymbol{\theta}})}{\partial\boldsymbol{\theta}_1\partial\omega_i}|_{\boldsymbol{\omega}=\boldsymbol{\omega}_0}$ and $\boldsymbol{\Delta}_{2i\boldsymbol{\omega}_0} = \frac{\partial^2 Q_{2i}(\boldsymbol{\theta}_2,\boldsymbol{\omega}|\widehat{\boldsymbol{\theta}})}{\partial\boldsymbol{\theta}_2\partial\omega_i}|_{\boldsymbol{\omega}=\boldsymbol{\omega}_0}$, $i = 1, \ldots, n$.

3.3.1 Perturbation of case weights

First, consider the following arbitrary allocation of weights for the expected value of the complete-data log-likelihood function (perturbed Q-function), which may capture departures in general directions, given by

$$Q(\boldsymbol{\theta}, \boldsymbol{\omega}|\widehat{\boldsymbol{\theta}}) = \sum_{i=1}^{n} \omega_i \mathrm{E}[\ell_i(\boldsymbol{\theta}|\mathbf{z}_c)] = \sum_{i=1}^{n} Q_{1i}(\boldsymbol{\theta}_1, \omega_i|\widehat{\boldsymbol{\theta}}) + \sum_{i=1}^{n} Q_{2i}(\boldsymbol{\theta}_2, \omega_i|\widehat{\boldsymbol{\theta}}),$$

where $\boldsymbol{\omega} = (\omega_1, \ldots, \omega_n)^\top$ is an $n \times 1$ vector with $0 \leq \omega_i \leq 1$ for $i = 1, \ldots, n$, such that $\boldsymbol{\omega}_0 = \mathbf{1}_n$, $Q_{1i}(\boldsymbol{\theta}_1, \omega_i|\widehat{\boldsymbol{\theta}}) = w_i Q_{1i}(\boldsymbol{\theta}_1|\widehat{\boldsymbol{\theta}})$ and $Q_{2i}(\boldsymbol{\theta}_2, \omega_i|\widehat{\boldsymbol{\theta}}) = w_i Q_{2i}(\boldsymbol{\theta}_2|\widehat{\boldsymbol{\theta}})$, where $Q_{1i}(\boldsymbol{\theta}_1|\widehat{\boldsymbol{\theta}})$ and $Q_{2i}(\boldsymbol{\theta}_2|\widehat{\boldsymbol{\theta}})$ are as presented in (5)–(6) and represent the contribution from the ith experimental unity to the Q-function. Note that, for $\omega_i = 0$ and $\omega_j = 1$, $j \neq i$, the ith experimental unit is dropped from the log-likelihood function for complete data. It is possible to show that the local influence for this perturbation scheme is equivalent to deletion method (see Osorio et al. 2009). Alternatively we may think on two other sub-perturbations, defined by (i) $Q(\boldsymbol{\theta}, \boldsymbol{\omega}|\widehat{\boldsymbol{\theta}}) = \sum_{i=1}^{n} Q_{1i}(\boldsymbol{\theta}_1, \omega_i|\widehat{\boldsymbol{\theta}}) + \sum_{i=1}^{n} Q_{2i}(\boldsymbol{\theta}_2|\widehat{\boldsymbol{\theta}})$ and (ii) $Q(\boldsymbol{\theta}, \boldsymbol{\omega}|\widehat{\boldsymbol{\theta}}) = \sum_{i=1}^{n} Q_{1i}(\boldsymbol{\theta}_1|\widehat{\boldsymbol{\theta}}) + \sum_{i=1}^{n} Q_{2i}(\boldsymbol{\theta}_2, \omega_i|\widehat{\boldsymbol{\theta}})$, which are assessed based on $C_{f_Q,\mathbf{d}}(\boldsymbol{\theta}_1)$ and $C_{f_Q,\mathbf{d}}(\boldsymbol{\theta}_2)$, respectively. Under this perturbation scheme the components of the matrices $\boldsymbol{\Delta}_{1\boldsymbol{\omega}_0}$ and $\boldsymbol{\Delta}_{2\boldsymbol{\omega}_0}$ are given in Appendix B.

3.3.2 Perturbation of the measurement for one instrument

In this case the measurements are obtained when one instrument is modified considering additive and multiplicative perturbation schemes. Let $\mathbf{Z}_{mi}(\omega_i)$ denote the perturbed measurement of the mth instrument for the ith experimental unit, with $m = 1, \ldots, p$ and $\boldsymbol{\omega} = (\omega_1, \ldots, \omega_n)^\top$. The following perturbation schemes will be evaluated:

(1) Additive perturbation: $\mathbf{Z}_{mi}(\omega_i) = \mathbf{Z}_i + \omega_i \mathbf{e}_m$, where $\mathbf{e}_m$ is a $p-$dimensional null vector with one in the mth position. In this case $\boldsymbol{\omega}_0 = \mathbf{0}$.

(2) Multiplicative perturbation: $\mathbf{Z}_{mi}(\omega_i) = \mathbf{Z}_i \boxdot \mathbf{1}_m(\omega_i)$, where $\mathbf{1}_m(\omega_i)$ of dimension p denoting a vector of ones having the mth component replaced by ω_i and $\boxdot$ denotes the Hadamard (elementwise) product. Here $\boldsymbol{\omega}_0 = \mathbf{1}$.

For both cases, the perturbed log-likelihood function for the complete data is given by

$$Q(\boldsymbol{\theta}, \boldsymbol{\omega}|\widehat{\boldsymbol{\theta}}) = \sum_{i=1}^{n} Q_{1i}(\boldsymbol{\theta}_1, \omega_i|\widehat{\boldsymbol{\theta}}) + \sum_{i=1}^{n} Q_{2i}(\boldsymbol{\theta}_2|\widehat{\boldsymbol{\theta}}),$$

where $Q_{1i}(\boldsymbol{\theta}_1, \omega_i|\widehat{\boldsymbol{\theta}})$ is as given in equation (5), switching $\mathbf{z}_{mi}(\omega_i)$ with $\mathbf{z}_i$. Under this perturbation scheme, the procedure of local influence is based on $C_{f_Q,\mathbf{d}}(\boldsymbol{\theta}_1)$ and the components of the matrix $\boldsymbol{\Delta}_{1\boldsymbol{\omega}_0}$ are given in Appendix B.

3.3.3 *Perturbation of the scale matrix* $\boldsymbol{\Sigma}$

Note that the model defined in (2)–(3) is assumed to be homoskedastic. So, in order to study the effects of departures from the homogeneity assumption, we propose to analyze the sensitivity of the MLE regarding the scale matrix $\boldsymbol{\Sigma}$. Indeed, we consider that the scale matrix of the $\mathbf{Z}_i$ is given by $\boldsymbol{\Sigma}^{(\omega_i)} = \phi_x^{(\omega_i)}\mathbf{b}\mathbf{b}^\top + D^{(\omega_i)}(\boldsymbol{\phi})$, which corresponds to considering that the distribution of $\mathbf{Z}_i$ is heteroscedastic (Xie et al. 2009). In this case, we have that

$$\mathbf{Z}_i \sim SMSN_p\left(\boldsymbol{\mu}, \boldsymbol{\Sigma}^{(\omega_i)}, \overline{\boldsymbol{\lambda}}_x; H\right), \quad i = 1, \ldots, n,$$

where $\boldsymbol{\Sigma}^{(\omega_i)} = \boldsymbol{\Sigma}/\omega_i$, $\boldsymbol{\mu}$, $\boldsymbol{\Sigma}$ and $\overline{\boldsymbol{\lambda}}_x$ are as in (4). Thus, note that we consider a simultaneous perturbation $D(\boldsymbol{\phi})^{(\omega_i)} = D(\boldsymbol{\phi})/\omega_i$ and $\phi_x^{(\omega_i)} = \phi_x/\omega_i$ since these terms are involved linearly in the scale matrix $\boldsymbol{\Sigma}$. Under this perturbation scheme, the non-perturbed model is obtained when $\boldsymbol{\omega}_o = (1, \ldots, 1)^\top$. Moreover, the perturbed Q-function has the form

$$Q(\boldsymbol{\theta}, \boldsymbol{\omega}|\widehat{\boldsymbol{\theta}}) = \sum_{i=1}^{n} Q_{1i}(\boldsymbol{\theta}_1, \omega_i|\widehat{\boldsymbol{\theta}}) + \sum_{i=1}^{n} Q_{2i}(\boldsymbol{\theta}_2, \omega_i|\widehat{\boldsymbol{\theta}}).$$

Clearly, the perturbed Q-function is as in (5)–(6), switching $D(\boldsymbol{\phi})^{(\omega_i)}$ and $\phi_x^{(\omega_i)}$ with $D(\boldsymbol{\phi})$ and ϕ_x, respectively. Using the perturbation scheme above, one can evaluate the homoskedasticity only of errors based on the perturbed Q-function $\sum_{i=1}^{n} Q_{1i}(\boldsymbol{\theta}_1, \omega_i|\widehat{\boldsymbol{\theta}})$. To evaluate the homoskedasticity of latent variable x_i we use the perturbed Q-function $\sum_{i=1}^{n} Q_{2i}(\boldsymbol{\theta}_2, \omega_i|\widehat{\boldsymbol{\theta}})$. The components of the matrix $\boldsymbol{\Delta}_{1\boldsymbol{\omega}_0}$ and $\boldsymbol{\Delta}_{2\boldsymbol{\omega}_0}$ under this perturbation scheme are given in Appendix B.

Note that it is not possible to give details for all the perturbation schemes that are of interest. However, as long as we can find an appropriate $\boldsymbol{\omega}$, and as long as the perturbed complete data log-likelihood function $\ell_c(\boldsymbol{\theta}, \boldsymbol{\omega}|\mathbf{z}_c)$ is smooth enough, so that the required derivatives in the diagnostic measures are well defined, we can conduct the local influence analysis without much difficulty. Recently, Zhu et al. (2007) proposed to use a metric tensor to select an appropriate perturbation to a statistical model. However, there are many issues related to this method, such as the calculation of the influence measures and metric tensor under different situations. An extension of this method to MEM is of great interest but is beyond the scope of this paper.

In the next section, for simplicity we consider the diagnostics for the marginal skewness parameter in SMSN-MEM. However, the method proposed here can be used to test for homogeneity of any parameter involved in the variance, as discussed by Xie et al. (2009).

4 Score test for homogeneity in SMSN-MEM

In the SMSN-MEM, the variance of the ith observation is $Var\{\mathbf{Z}_i\} = E_2\boldsymbol{\Sigma} - E_1^2\boldsymbol{\Delta}\boldsymbol{\Delta}^\top$, where $E_m = E\{\kappa^{-m/2}(U)\}$, $m = 1, 2$, $\boldsymbol{\Delta} = \boldsymbol{\Sigma}^{1/2}\boldsymbol{\delta}_x$ and $\boldsymbol{\delta}_x = \bar{\boldsymbol{\lambda}}_x/(1 + \bar{\boldsymbol{\lambda}}_x^\top\bar{\boldsymbol{\lambda}}_x)^{1/2}$, with $\boldsymbol{\Sigma}$ and $\bar{\boldsymbol{\lambda}}_x$ as in (4), in which the skewness parameter $\bar{\boldsymbol{\lambda}}_x$ is constant. However, the actual skewness parameter may be related to the ith observation $\mathbf{z}_i$ and thus its variance is nonconstant. Then, one cannot make any inference for the model without further assumptions, since there are too many unknown parameters involved. In view of that, it is necessary to test homogeneity of skewness parameters. This section concentrates on this problem in SMSN-MEM. Following Xie et al. (2009), we suppose that the skewness parameter can be modeled by

$$\bar{\boldsymbol{\lambda}}_i = \bar{\boldsymbol{\lambda}}_x m(\boldsymbol{v}_i, \boldsymbol{\gamma}) = \bar{\boldsymbol{\lambda}}_x m_i,$$

for $i = 1, \ldots, n$, where $\bar{\boldsymbol{\lambda}}_x$ is an unknown parameter, $\boldsymbol{\gamma}$ is a $q \times 1$ unknown vector which introduces heterogeneity in the skewness parameter, $\boldsymbol{v}_i$'s are covariates of appropriates dimension, and m is a known differentiable weight function of skewness parameter in $\boldsymbol{\gamma}$. It is assumed that there is a unique value $\boldsymbol{\gamma}_0$ of $\boldsymbol{\gamma}$ such that $m(\boldsymbol{v}_i, \boldsymbol{\gamma}_0) = 1$ for all i. Hence the test for homogeneity of skewness parameter is equivalent to testing the following hypothesis:

$$H_0 : \boldsymbol{\gamma} = \boldsymbol{\gamma}_0 \text{ vs. } H_1 : \boldsymbol{\gamma} \neq \boldsymbol{\gamma}_0. \tag{8}$$

Let $\boldsymbol{\xi}_2$ denote $(\boldsymbol{\gamma}^\top, \boldsymbol{\theta}^\top)^\top$. Then for the hypothesis (8), $\boldsymbol{\gamma}$ is the parameter of interest and $\boldsymbol{\theta}$ is a nuisance parameter. The log-likelihood $l(\boldsymbol{\xi}_2|\mathbf{z}) = l(\boldsymbol{\xi}_2)$ is obtained by switching K_i^{**} with K_i (see Lachos et al. 2010c, Eq. 11), where

$$K_i^{**} = \int_0^\infty \kappa^{-p/2}(u_i)\exp\left(-\frac{1}{2}\kappa^{-1}(u_i)d_i\right)\Phi_1(\kappa^{-1/2}(u_i)\mathrm{A}_i^{**})dH(u_i;\boldsymbol{\nu}),$$

with $A_i^{**} = m_i \bar{\boldsymbol{\lambda}}_x^\top \boldsymbol{\Sigma}^{-1/2}(\mathbf{z}_i - \boldsymbol{\mu}) = m_i A_i$. Based on log-likelihood, the second-order derivatives of $l(\boldsymbol{\xi}_2)$ with respect to the parameter $\boldsymbol{\gamma}$ and $\boldsymbol{\theta}$ can be easily obtained. Thus, the observed information matrix of $\mathbf{Z}$ for $\boldsymbol{\xi}_2$ under H_0 is given by

$$\mathbf{J}_{\mathbf{Z}}(\boldsymbol{\xi}_2)|_{\boldsymbol{\xi}_2=\widehat{\boldsymbol{\xi}}_2^0} = \begin{pmatrix} \mathbf{J}_{\boldsymbol{\gamma}\boldsymbol{\gamma}} & \mathbf{J}_{\boldsymbol{\gamma}\boldsymbol{\theta}} \\ \mathbf{J}_{\boldsymbol{\gamma}\boldsymbol{\theta}}^\top & \mathbf{J}_{\boldsymbol{\theta}\boldsymbol{\theta}} \end{pmatrix}|_{\boldsymbol{\xi}_2=\widehat{\boldsymbol{\xi}}_2^0},$$

with elements given in Appendix C, and $\widehat{\boldsymbol{\xi}}_2^0 = (\boldsymbol{\gamma}_0^\top, \widehat{\boldsymbol{\theta}}^\top)^\top$ denotes the ML estimate of $\boldsymbol{\xi}_2$ under the null hypothesis H_0. The first-order derivative of $l(\boldsymbol{\xi}_2)$ with respect to $\boldsymbol{\gamma}$ is

$$\frac{\partial l(\boldsymbol{\xi}_2)}{\partial \boldsymbol{\gamma}} = \sum_{i=1}^{n} \frac{1}{K_i^{**}} I_i^{\phi}\left(\frac{p+1}{2}\right) A_i \frac{\partial m_i}{\partial \boldsymbol{\gamma}},$$

where $I_i^F(w) = \int_0^\infty \kappa^{-w}(u_i)\exp\{-\frac{1}{2}\kappa^{-1}(u_i)d_i\}F(\kappa^{-1/2}(u_i)A_i)dH(u_i;\boldsymbol{\nu})$, where $F(\cdot)$ is the function $\Phi(\cdot)$ or $\phi(\cdot)$. Moreover, the explicit expressions of $I_i^{\phi}(w)$ and $I_i^{\Phi}(w)$ for the skew-t, the skew-slash and the skew-contaminated normal distributions are given in Lachos et al. (2010c). Then, the score function of hypothesis (8) is $\mathbf{S}(\boldsymbol{\xi}_2)|_{\boldsymbol{\xi}_2=\widehat{\boldsymbol{\xi}}_2^0} = \mathbf{M}^\top \mathbf{h}|_{\boldsymbol{\xi}_2=\widehat{\boldsymbol{\xi}}_2^0}$, where $\mathbf{M}^\top = (\mathbf{M}_1^\top, \ldots, \mathbf{M}_n^\top)$, $\mathbf{M}_i = \frac{\partial m_i}{\partial \boldsymbol{\gamma}^\top}$, $\mathbf{h} = (h_1, \ldots, h_n)$, $h_i = \frac{1}{K_i^{**}} I_i^{\phi}\left(\frac{p+1}{2}\right) A_i$. Furthermore, the score test statistic for H_0 is

$$SC = \left\{\mathbf{S}^\top(\boldsymbol{\xi}_2)\mathbf{J}_{\mathbf{Z}}^{-1}(\boldsymbol{\xi}_2)\mathbf{S}(\boldsymbol{\xi}_2)\right\}|_{\boldsymbol{\xi}_2=\widehat{\boldsymbol{\xi}}_2^0}.$$

When H_0 is true, the statistic SC is asymptotically distributed as χ_q^2. In the next section, a real example is presented to illustrate the performance of the methodology developed.

5 Application

We illustrate our proposed methods with a data set from Chipkevitch et al. (1996). The data measure the testicular volume of 42 adolescents in a certain sequence by using five different techniques. Recently, Lachos et al. (2010c) have analyzed the Chipkevitch data set with the aim of providing additional inferences by using SMSN-MEM. As our main focus is not on ML estimation, we refer the interested reader to see Table 1 in Lachos et al. (2010c), which contains the ML estimates of the parameters from the SN, ST, SCN and SSL models, together with their corresponding standard errors calculated via the observed information matrix. In this section, we revisit this data set with the aim of applying Zhu and Lee's (2001) local influence approach to SMSN-MEM.

In order to detect outlying observations, the Mahalanobis distance has been considered (Zeller et al. 2011). Figure 1 (first row) displays such distances for the SN and ST fitted models. The cutoff lines corresponds to the quantile $\upsilon = \chi_p^2(\xi)$, where $\xi = 0.95$.

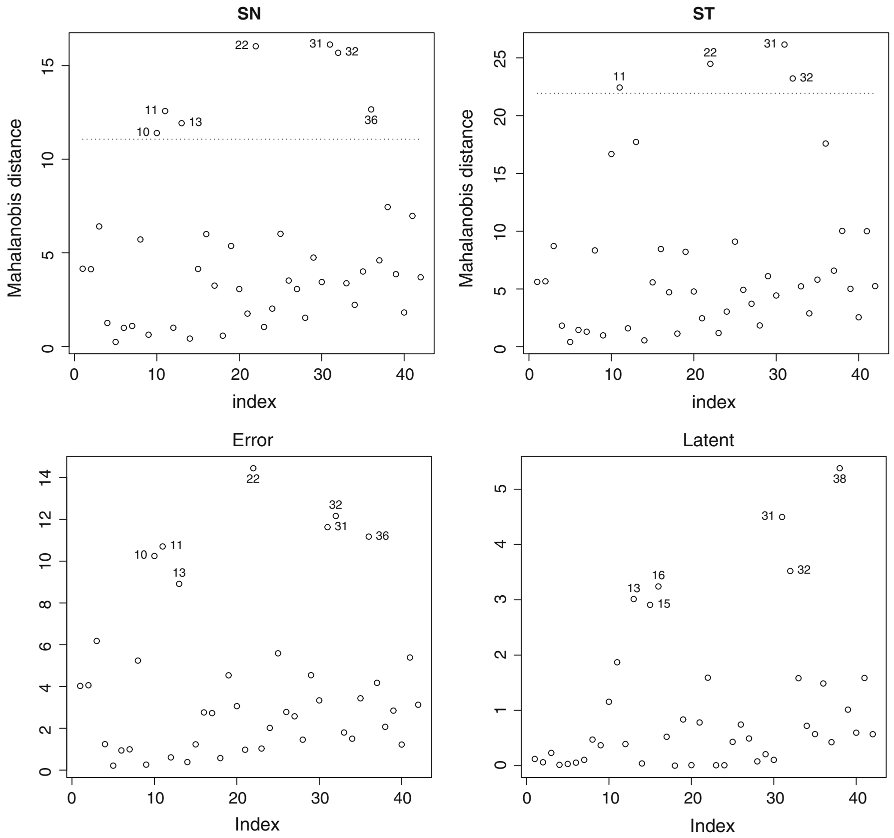

Fig. 1 Chipkevitch data set. In the first row the index plots of the Mahalanobis distances for the SN and ST fitted models. The *horizontal lines* corresponds to the quantile $\upsilon = \chi_p^2(0.95)$. The estimated $d_{\mathbf{e}i}$ (error) and d_{xi} (latent) to the skew-normal fit are given at the bottom

We can see from these figures that observations 11, 22, 31 and 32 are detected as possible outliers in the ST case, and further that observations 10, 13 and 36 are detected as possible outliers in the SN case. For the SSL and SCN cases, the results are the same as for the ST and SN cases, respectively, so they are not shown here to save space. The estimated distances $d_{\mathbf{e}i}$ (Error) and d_{xi} (Latent), defined in (7), provide useful diagnostic statistics for identifying subjects with outlying observations—see also Ho and Lin (2010). Figure 1 (second row) presents those diagnostic statistics to the SN-MEM. It can be seen that individuals 10, 11, 13, 22, 31, 32 and 36 present large values of $\widehat{d}_{\mathbf{e}i}$, suggesting they are **e**-outliers. Moreover, observations 13, 15, 16, 31, 32 and 38 present large values of $\widehat{d}_{xi}$, suggesting they are x-outliers. For cases ST-MEM, SSL-MEM and SCN-MEM, the results are the same as for SN-MEM.

From the EM-algorithm, the estimated weights $\widehat{u}_i$, $i = 1, \ldots, 42$, under ST and SSL distributions, are depicted in Figure 2 (first row). Notice that when we use distributions with tails heavier than the skew-normal ones, the EM algorithm allows accommodating such observations by attributing small weights to them in the

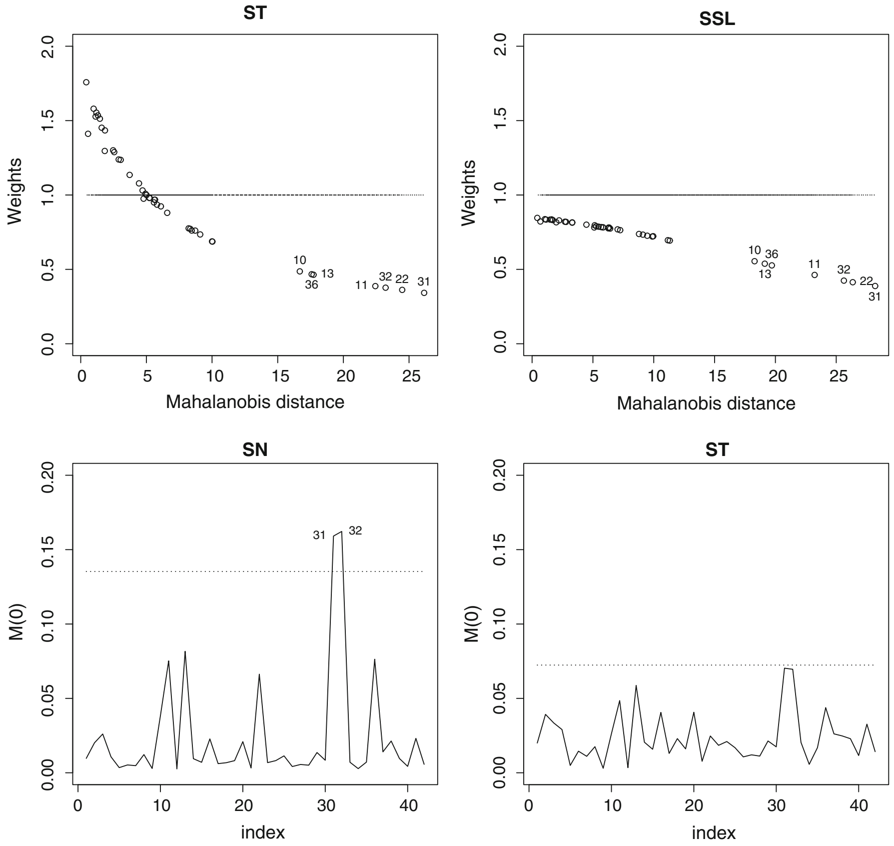

Fig. 2 Chipkevitch data set. The estimated weights $\widehat{u}_i$ for the ST and SSL models are in the first row. The index plots of M(0) under case weights perturbation for the SN and ST fitted models are given at the bottom

estimation procedure. The weights for the SN distribution ($\widehat{u}_i$) are indicated in Figure 2 (first row) as a continuous line. Therefore, this rich class of distributions may naturally attribute different weights to each observation and consequently control the influence of a single observation on the parameter estimates. These results agree with the results presented in Osorio et al. (2009), in a symmetric context. Next we identify influential observations for the Chipkevitch data set using $M(0)$ from the conformal curvature $B_{f_Q,\mathbf{u}_i}$. The perturbation schemes described in Sect. 3.3 are considered and in all cases we consider the Lee and Xu (2004) benchmark for M(0) with $c^* = 3$.

5.1 Case weights perturbation

From Figure 2 (second row) it can be seen that observations 31 and 32 are detected as outliers in Figure 1, but no observation is detected as influential under the ST-MEM and also under the SSL-MEM and SCN-MEM (not shown here).

5.2 Perturbation of the measurement for one instrument

We now examine the effects of perturbing the measurements taken by instruments II and V. Instruments II and V were chosen because they present the largest $C_{f_Q,\mathbf{dmax}}$ values. The values of $C_{f_Q,\mathbf{dmax}}$ for additive perturbations of the measurements taken by instrument II, for instance, are 2.2437 (SN), 2.6489 (ST), 2.3749 (SSL) and 2.6056 (SCN), whereas for multiplicative perturbations taken the instrument V, the $C_{f_Q,\mathbf{dmax}}$ values are 758.8672 (SN), 564.2995 (ST), 626.0935 (SSL) and 563.8614 (SCN). Figure 3 illustrates the index plot for perturbation of the measurement for one instrument. Using this perturbation scheme, we can examine the influence on the measurements taken by instruments II and V, under the additive and multiplicative cases, respectively. Figure 3 (first row) indicates some influence when the measurement of item 36 for instrument II is perturbed under SN-MEM. For cases SSL-MEM and SCN-MEM, the results are the same as for ST-MEM. In Figure 3 (second row) we see some influences when the measurement of item 32 for instrument V is perturbed under SN-MEM and SSL-MEM (not shown here). Moreover, in this perturbation case, observation 16 is only slightly prominent under SCN-MEM (not shown here). We note that for this data set, the ST model accommodates the influential observations slightly better. As expected, the influence of such observations are reduced when we consider distributions with heavier tails than the skew-normal.

5.3 Perturbation of variances

To assess the assumption of homoscedasticity of the model, we obtain Figure 4. It can be seen that observations 31 and 32 are identified as influential only under SN-MEM when compared with ST-MEM. For the SSL-MEM and SCN-MEM cases, the results are the same as for ST-MEM. Using this perturbation scheme we also can evaluate the homoscedasticity of the latent variable x. Under the SN, SSL and SCN models, observations 15 and 38 are detected as influential (not shown here). This fact is in accordance with Figure 1 (second row), where those observations were identified as x-outliers.

To test the homogeneity of skewness parameter, we assume that the weight function $m_i = \exp(\gamma \upsilon_i)$. It is easily seen that when $\gamma = 0$, then $m_i = 1$ and $\bar{\boldsymbol{\lambda}}_{x,i} = \bar{\boldsymbol{\lambda}}_x, \forall i$. Hence, the test for the homogeneity of skewness parameter becomes the test of hypothesis $H_0 : \gamma = 0$. Using the statistic given in Sect. 4 and with a little computation, we have that the $SC = 1.4066$ (p-value = 0.2356) for SN-MEM, $SC = 0.5898$ (p-value = 0.4425) for ST-MEM, $SC = 0.9223$ (p-value = 0.3369) for SSL-MEM and $SC = 0.5836$ (p-value = 0.4449) for SCN-MEM, which indicate there is no significant evidence of a varying skewness parameter and consequently of heterogeneity in the Chipkevitch's data set.

Now we use the quantities TRC and MRC to reveal the impact of the influential observations detected. These quantities are defined, respectively, by $TRC = \sum_{j=1}^{n_p} |\frac{\hat{\boldsymbol{\theta}}_j - \hat{\boldsymbol{\theta}}_{[i]j}}{\hat{\boldsymbol{\theta}}_j}|$ and $MRC = max_{j=1,\ldots,n_p} |\frac{\hat{\boldsymbol{\theta}}_j - \hat{\boldsymbol{\theta}}_{[i]j}}{\hat{\boldsymbol{\theta}}_j}|$, where n_p is the dimension of $\boldsymbol{\theta}$ and the subscript $[i]$ means the ML estimator of $\boldsymbol{\theta}$ with the i-th observation, $\mathbf{z}_i$,

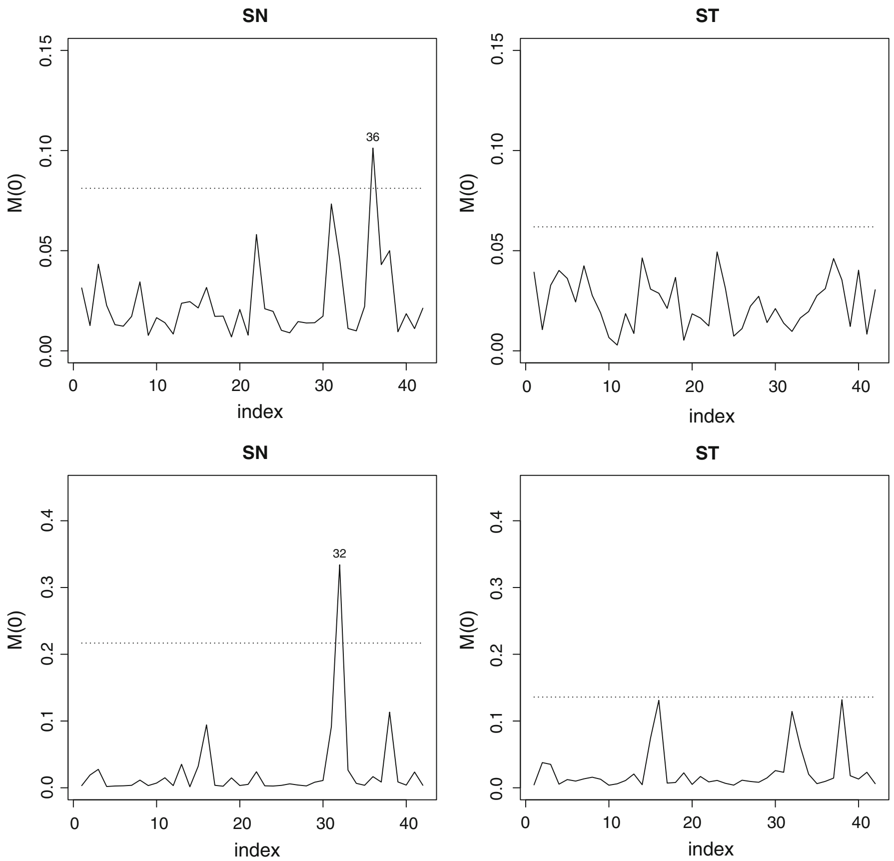

Fig. 3 Chipkevitch data set. *First row*—the index plots of M(0) under additive perturbation of the measurements taken by instrument II. *Second row*—the index plots of M(0) under multiplicative perturbation of the measurements taken by instrument V

deleted. The comparison of these measures, based on the different models, with the most influential observations 31 and 32 deleted, are given in Table 1. Notice that the greatest changes take place under the SN distribution. As expected, the results indicate that the ML estimators are less sensitive in the presence of influential observations when we use distributions with heavier tails than the SN one.

6 Final conclusion

We have presented strategies to perform influence diagnostics in multivariate measurement error models under SMSN distributions. We used the results of Lachos et al. (2010c) for obtaining parameter estimation via maximum likelihood, based on the EM algorithm, which yields closed form expressions for the equations in the M-step. Local influence methods were implemented for the SMSN-MEM in order to evaluate the consequences of model perturbations in situations where different perturbation

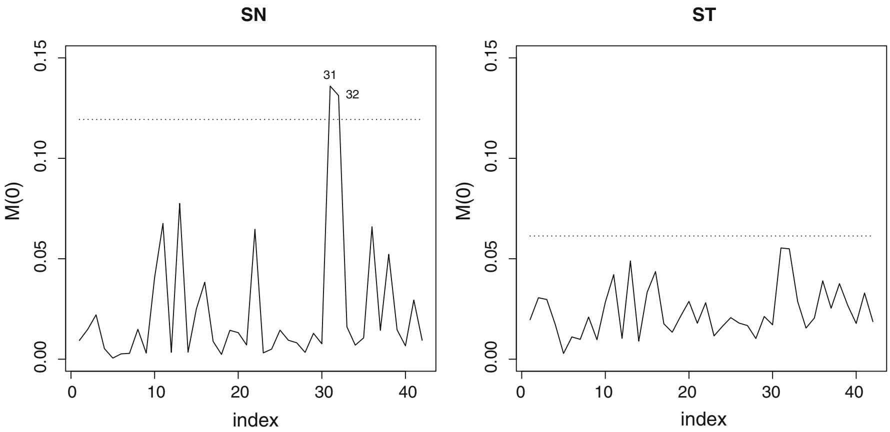

Fig. 4 Chipkevitch data set. Index plots of M(0) under scale matrix perturbation of random effects for the SN and ST fitted models

Table 1 Chipkevitch data set. Comparison of the relative changes in the ML estimators in terms of TRC and MRC for the four selected SMSN models

Observations	Distribution	TRC	MRC
31 and 32	SN	53.6917	49.7396
	ST	2.1341	1.1589
	SSL	3.1918	2.0303
	SCN	2.5972	1.6614

schemes are investigated. However, other perturbation schemes can be considered in an analogous way. The Chipkevitch et al. (1996) data set favors the use of SMSN distributions with heavy tails, specifically the use of the skew-t one. As in the symmetric case, it is important to emphasize the capacity of such models to attenuate outlying observations, by means of attributing a small weight to these observations in the estimation process.

To examine the performances and properties of the score test, formal simulations studies under several situations need to be carried out, as discussed in Xie et al. (2009), which will be reported in a follow-up paper. Due to recent advances in computational technology, it is worthwhile also to carry out Bayesian treatments via Markov chain Monte Carlo (MCMC) sampling methods in the context of SMSN-MEM. The basic idea is to explore the joint posterior distributions of the model parameters together with latent variables x_i and u_i, when informative priors are employed. Bayesian influence diagnostics can be treated via the Kullback–Leibler divergence, as proposed by Cancho et al. (2010).

Acknowledgments We are grateful to the editor and one anonymous referee for their useful comments, which substantially improved the quality of this paper. The authors acknowledge the partial financial support from Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq-Brazil) and Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP-Brazil).

Appendix A: The Hessian matrix $\ddot{Q}(\widehat{\theta})$

In this appendix, we use the following notation: $\ddot{Q}_{1i,\tau\pi} = \frac{\partial^2 Q_{1i}(\boldsymbol{\theta}_1|\widehat{\boldsymbol{\theta}})}{\partial\tau\partial\pi}$, with $\tau, \pi = \boldsymbol{\alpha}, \boldsymbol{\beta}$ or $\boldsymbol{\phi}$ and $\ddot{Q}_{2i,\tau\pi} = \frac{\partial^2 Q_2(\boldsymbol{\theta}_2|\widehat{\boldsymbol{\theta}})}{\partial\tau\partial\pi}$, with $\tau,\ \pi = \mu_x\ \phi_x$ or λ_x. The matrix Hessian has elements given by (see Magnus and Neudecker 1988)

$$
\begin{aligned}
\ddot{Q}_{1i,\alpha\alpha} &= -\widehat{u}_i \mathbb{I}_{(p)} D^{-1}(\boldsymbol{\phi})\mathbb{I}_{(p)}^\top,\\
\ddot{Q}_{1i,\alpha\beta} &= -\widehat{ux}_i \mathbb{I}_{(p)} D^{-1}(\boldsymbol{\phi})\mathbb{I}_{(p)}^\top, \quad \ddot{Q}_{1i,\beta\beta} = -\widehat{ux^2}_i \mathbb{I}_{(p)} D^{-1}(\boldsymbol{\phi})\mathbb{I}_{(p)}^\top,\\
\ddot{Q}_{1i,\phi\alpha} &= (-\widehat{u}_i D(\mathbf{z}_i - \mathbf{a}) + \widehat{ux}_i D(\mathbf{b}))\, D^{-2}(\boldsymbol{\phi})\mathbb{I}_{(p)}^\top,\\
\ddot{Q}_{1i,\phi\beta} &= \left(-\widehat{ux}_i D(\mathbf{z}_i - \mathbf{a}) + \widehat{ux^2}_i D(\mathbf{b})\right) D^{-2}(\boldsymbol{\phi})\mathbb{I}_{(p)}^\top,\\
\ddot{Q}_{1i,\phi\phi} &= \left[\frac{1}{2} D(\boldsymbol{\phi}) - \widehat{u}_i D^2(\mathbf{z}_i - \mathbf{a}) + 2\widehat{ux}_i D(\mathbf{z}_i - \mathbf{a}) D(\mathbf{b}) - \widehat{ux^2}_i D^2(\mathbf{b})\right] D^{-3}(\boldsymbol{\phi}),\\
\ddot{Q}_{2i,\mu_x\mu_x} &= -\frac{1}{\nu_x^2}\widehat{u}_i, \quad \ddot{Q}_{2i,\mu_x\phi_x} = \frac{\lambda_x(1+\lambda_x^2)^{1/2}}{2\phi_x^{3/2}}\widehat{ut}_i + \frac{(1+\lambda_x^2)}{\phi_x^2} B_{1i},\\
\ddot{Q}_{2i,\mu_x\lambda_x} &= -\frac{2\lambda_x}{\phi_x} B_{1i} - \frac{(1+2\lambda_x^2)}{\phi_x^{1/2}(1+\lambda_x^2)^{1/2}}\widehat{ut}_i, \ \ddot{Q}_{2i,\lambda_x\phi_x} = \frac{(1+2\lambda_x^2)}{2\phi_x^{3/2}(1+\lambda_x^2)^{1/2}} B_{2i} + \frac{\lambda_x}{\phi_x^2} B_{3i},\\
\ddot{Q}_{2i,\phi_x\phi_x} &= \frac{1}{2\phi_x^2} - \frac{3\lambda_x(1+\lambda_x^2)^{1/2}}{4\phi_x^{5/2}} B_{2i} - \frac{(1+\lambda_x^2)}{\phi_x^3} B_{3i},\\
\ddot{Q}_{2i,\lambda_x\lambda_x} &= \frac{(1-\lambda_x^2)}{(1+\lambda_x^2)^2} - \frac{1}{\phi_x} B_{3i} - \widehat{ut^2}_i - \frac{\lambda_x(3+2\lambda_x^2)}{\phi_x^{1/2}(1+\lambda_x^2)^{3/2}} B_{2i},
\end{aligned}
$$

where $\mathbb{I}_{(p)} = [\mathbf{0}_q, \mathbf{I}_q]$, such that $\mathbf{0}_q = (0,\ldots,0)^\top$ is a $q \times 1$ vector, $\mathbf{I}_q$ is a $q \times q$ identity matrix, $B_{1i} = \widehat{u}_i\mu_x - \widehat{ux}_i$, $B_{2i} = \widehat{ut}_i\mu_x - \widehat{utx}_i$ and $B_{3i} = \widehat{ux^2}_i + \widehat{u}_i\mu_x^2 - 2\widehat{ux}_i\mu_x$.

Appendix B: Derivatives with respect to the perturbation schemes

In this appendix, we use the following notation: $\ddot{Q}_{1i,\tau\omega_i} = \frac{\partial^2 Q_{1i}(\boldsymbol{\theta}_1,\boldsymbol{\omega}|\widehat{\boldsymbol{\theta}})}{\partial\tau\partial\omega_i}|_{\boldsymbol{\omega}=\boldsymbol{\omega}_o}$, with $\tau = \boldsymbol{\alpha}, \boldsymbol{\beta}$ or $\boldsymbol{\phi}$ and $\ddot{Q}_{2i,\tau\omega_i} = \frac{\partial^2 Q_{2i}(\boldsymbol{\theta}_2,\boldsymbol{\omega}|\widehat{\boldsymbol{\theta}})}{\partial\tau\partial\omega_i}|_{\boldsymbol{\omega}=\boldsymbol{\omega}_o}$, with $\tau = \mu_x, \quad phi_x$ or λ_x.

1. Case weights perturbation

$$
\begin{aligned}
\ddot{Q}_{1i,\alpha\omega_i} &= \mathbb{I}_{(p)} D^{-1}(\boldsymbol{\phi})(\widehat{u}_i(\mathbf{z}_i - \mathbf{a}) - \widehat{ux}_i \mathbf{b}),\\
\ddot{Q}_{1i,\beta\omega_i} &= \mathbb{I}_{(p)} D^{-1}(\boldsymbol{\phi})(\widehat{ux}_i(\mathbf{z}_i - \mathbf{a}) - \widehat{ux^2}_i \mathbf{b}),\\
\ddot{Q}_{1i,\phi\omega_i} &= \frac{1}{2}\left[-D(\boldsymbol{\phi}) + \widehat{u}_i D^2(\mathbf{z}_i - \mathbf{a}) - 2\widehat{ux}_i D(\mathbf{z}_i - \mathbf{a}) D(\mathbf{b})\right] D^{-2}(\boldsymbol{\phi})\mathbf{1}_p\\
&\quad + \frac{1}{2}\widehat{ux^2}_i D^2(\mathbf{b}) D^{-2}(\boldsymbol{\phi})\mathbf{1}_p, \quad \ddot{Q}_{2i,\mu_x\omega_i} = -\frac{1}{\nu_x^2}(B_{1i} + \widehat{ut}_i\tau_x),\\
\ddot{Q}_{2i,\phi_x\omega_i} &= -\frac{1}{2\phi_x} + \frac{\lambda_x(1+\lambda_x^2)^{1/2}}{2\phi_x^{3/2}} B_{2i} + \frac{1+\lambda_x^2}{2\phi_x^2} B_{3i},
\end{aligned}
$$

$$\ddot{Q}_{2i,\lambda_x\omega_i} = \frac{\lambda_x}{(1+\lambda_x^2)} - \frac{\lambda_x}{\phi_x}B_{3i} - \lambda_x\widehat{ut^2}_i - \frac{(1+2\lambda_x^2)}{\phi_x^{1/2}(1+\lambda_x^2)^{1/2}}B_{2i},$$

where $\mathbb{I}_{(p)} = [\mathbf{0}_q, \mathbf{I}_q]$, such that $\mathbf{0}_q = (0,\ldots,0)^\top$ is a $q\times 1$ vector, $\mathbf{I}_q$ is a $q\times q$ identity matrix, $B_{1i} = \widehat{u}_i\mu_x - \widehat{ux}_i$, $B_{2i} = \widehat{ut}_i\mu_x - \widehat{utx}_i$ and $B_{3i} = \widehat{ux^2}_i + \widehat{u}_i\mu_x^2 - 2\widehat{ux}_i\mu_x$.

2. Perturbation of the measurement for one instrument

$$\ddot{Q}_{1i,\alpha\omega_i} = \widehat{u}_i\mathbb{I}_{(p)}D^{-1}(\boldsymbol{\phi})\mathbf{p}_m, \quad \ddot{Q}_{1i,\beta\omega_i} = \widehat{ux}_i\mathbb{I}_{(p)}D^{-1}(\boldsymbol{\phi})\mathbf{p}_m,$$
$$\ddot{Q}_{1i,\phi\omega_i} = \left(\widehat{u}_i D(\mathbf{z}_i-\mathbf{a}) - \widehat{ux}_i D(\mathbf{b})\right) D^{-2}(\boldsymbol{\phi})\mathbf{p}_m,$$

where $\mathbf{p}_m = \mathbf{e}_m$ under the additive case and $\mathbf{p}_m = \mathbf{0}_m(z_{im})$ under the multiplicative case, with $\mathbf{0}_m(z_{im})$ a $p-$dimensional null vector with z_{mi} in the mth position.

3. Perturbation of the scale matrix $\boldsymbol{\Sigma}$

$$\ddot{Q}_{1i,\alpha\omega_i} = \mathbb{I}_{(p)}D^{-1}(\boldsymbol{\phi})(\widehat{u}_i(\mathbf{z}_i-\mathbf{a}) - \widehat{ux}_i\mathbf{b}),$$
$$\ddot{Q}_{1i,\beta\omega_i} = \mathbb{I}_{(p)}D^{-1}(\boldsymbol{\phi})(\widehat{ux}_i(\mathbf{z}_i-\mathbf{a}) - \widehat{ux^2}_i\mathbf{b}),$$
$$\ddot{Q}_{1i,\phi\omega_i} = \frac{1}{2}\left[\widehat{u}_i D^2(\mathbf{z}_i-\mathbf{a}) - 2\widehat{ux}_i D(\mathbf{z}_i-\mathbf{a})D(\mathbf{b}) + \widehat{ux^2}_i D^2(\mathbf{b}\right] D^{-2}(\boldsymbol{\phi})\mathbf{1}_p,$$
$$\ddot{Q}_{2i,\mu_x\omega_i} = -\frac{1}{\nu_x^2}(B_{1i} + \frac{1}{2}\widehat{ut}_i\tau_x), \quad \ddot{Q}_{2i,\phi_x\omega_i} = \frac{(1+\lambda_x^2)}{2\phi_x^2}B_{3i} + \frac{\lambda_x(1+\lambda_x^2)^{1/2}}{4\phi_x^{3/2}}B_{2i},$$
$$\ddot{Q}_{2i,\lambda_x\omega_i} = -\frac{\lambda_x}{\phi_x}B_{3i} - \frac{(1+2\lambda_x^2)}{2\phi_x^{1/2}(1+\lambda_x^2)^{1/2}}B_{2i}.$$

Appendix C: Derivation of score statistic in the SMSN-MEM case

The observed information matrix $\mathbf{J_Z}(\boldsymbol{\xi}_2) = -\mathbf{L}$, where $\mathbf{L} = \begin{pmatrix} \mathbf{L}_{\boldsymbol{\gamma\gamma}} & L_{\boldsymbol{\gamma\theta}} \\ \mathbf{L}_{\boldsymbol{\gamma\theta}}^\top & \mathbf{L}_{\boldsymbol{\theta\theta}} \end{pmatrix}$ denotes the matrix of second derivatives with respect to $\boldsymbol{\xi}_2 = (\boldsymbol{\gamma}^\top, \boldsymbol{\theta}^\top)^\top$ given by

$$\mathbf{L} = \sum_{i=1}^n \frac{\partial^2 \ell_i(\boldsymbol{\xi}_2)}{\partial\boldsymbol{\varphi}\partial\boldsymbol{\varphi}^\top} = -\frac{n}{2}\frac{\partial^2 \log|\boldsymbol{\Sigma}|}{\partial\boldsymbol{\varphi}\partial\boldsymbol{\varphi}^\top} - \sum_{i=1}^n \frac{1}{K_i^{**2}}\frac{\partial K_i^{**}}{\partial\boldsymbol{\varphi}}\frac{\partial K_i^{**}}{\partial\boldsymbol{\varphi}^\top} + \sum_{i=1}^n \frac{1}{K_i^{**}}\frac{\partial^2 K_i^{**}}{\partial\boldsymbol{\varphi}\partial\boldsymbol{\varphi}^\top},$$

where $\boldsymbol{\varphi} = \boldsymbol{\gamma}$ or $\boldsymbol{\theta}$. Thus, $\mathbf{L}_{\theta\theta}$ denotes the matrix of second derivatives with respect to $\boldsymbol{\theta}$ and under $H_0 : \boldsymbol{\gamma} = \boldsymbol{\gamma}_0$ and is given in (Lachos et al. 2010c, Appendix). On the other hand, $\mathbf{L}_{\boldsymbol{\gamma\gamma}}$ denotes the matrix of second derivatives with respect to $\boldsymbol{\gamma}$ and is given by

$$\mathbf{L}_{\boldsymbol{\gamma\gamma}} = -\sum_{i=1}^n \frac{1}{K_i^{**2}}\frac{\partial K_i^{**}}{\partial\boldsymbol{\gamma}}\frac{\partial K_i^{**}}{\partial\boldsymbol{\gamma}^\top} + \sum_{i=1}^n \frac{1}{K_i^{**}}\frac{\partial^2 K_i^{**}}{\partial\boldsymbol{\gamma}\partial\boldsymbol{\gamma}^\top},$$

where $\frac{\partial K_i^{**}}{\partial \boldsymbol{\gamma}} = I_i^{\phi}(\frac{p+1}{2})\frac{\partial A_i^{**}}{\partial \boldsymbol{\gamma}}$, $\frac{\partial^2 K_i}{\partial \boldsymbol{\gamma}\partial \boldsymbol{\gamma}^\top} = -I_i^{\phi}(\frac{p+3}{2})A_i^{**}\frac{\partial A_i^{**}}{\partial \boldsymbol{\gamma}}\frac{\partial A_i^{**}}{\partial \boldsymbol{\gamma}^\top} + I_i^{\phi}(\frac{p+1}{2})\frac{\partial^2 A_i^{**}}{\partial \boldsymbol{\gamma}\partial \boldsymbol{\gamma}^\top}$, with $\frac{\partial A_i^{**}}{\partial \boldsymbol{\gamma}} = \frac{\partial m_i}{\partial \boldsymbol{\gamma}}A_i$ and $\frac{\partial^2 A_i^{**}}{\partial \boldsymbol{\gamma}\partial \boldsymbol{\gamma}^\top} = \frac{\partial^2 m_i}{\partial \boldsymbol{\gamma}\partial \boldsymbol{\gamma}^\top}A_i$. In addition, $\mathbf{L}_{\boldsymbol{\gamma}\boldsymbol{\theta}}$ denotes the matrix of second derivatives with respect to $\boldsymbol{\gamma}$ and $\boldsymbol{\theta}$, and it is given by

$$\mathbf{L}_{\boldsymbol{\gamma}\boldsymbol{\theta}} = -\sum_{i=1}^{n}\frac{1}{K_i^{**2}}\frac{\partial K_i^{**}}{\partial \boldsymbol{\gamma}}\frac{\partial K_i^{**}}{\partial \boldsymbol{\theta}^\top} + \sum_{i=1}^{n}\frac{1}{K_i^{**}}\frac{\partial^2 K_i^{**}}{\partial \boldsymbol{\gamma}\partial \boldsymbol{\theta}^\top},$$

where $\frac{\partial^2 K_i}{\partial \boldsymbol{\gamma}\partial \boldsymbol{\theta}^\top} = -\frac{1}{2}I_i^{\phi}(\frac{p+3}{2})\frac{\partial A_i^{**}}{\partial \boldsymbol{\gamma}}\frac{\partial d_i}{\partial \boldsymbol{\theta}^\top} - I_i^{\phi}(\frac{p+3}{2})A_i^{**}\frac{\partial A_i^{**}}{\partial \boldsymbol{\gamma}}\frac{\partial A_i^{**}}{\partial \boldsymbol{\theta}^\top} + I_i^{\phi}(\frac{p+1}{2})\frac{\partial^2 A_i^{**}}{\partial \boldsymbol{\gamma}\partial \boldsymbol{\theta}^\top}$, $\frac{\partial K_i^{**}}{\partial \boldsymbol{\theta}} = I_i^{\phi}(\frac{p+1}{2})\frac{\partial A_i^{**}}{\partial \boldsymbol{\theta}} - \frac{1}{2}I_i^{\Phi}(\frac{p+2}{2})\frac{\partial d_i}{\partial \boldsymbol{\theta}}$, $\frac{\partial A_i^{**}}{\partial \boldsymbol{\theta}} = m_i\frac{\partial A_i}{\partial \boldsymbol{\theta}}$, $\frac{\partial^2 A_i^{**}}{\partial \boldsymbol{\gamma}\partial \boldsymbol{\theta}^\top} = \frac{\partial m_i}{\partial \boldsymbol{\gamma}}\frac{\partial A_i}{\partial \boldsymbol{\theta}^\top}$, with $\frac{\partial A_i}{\partial \boldsymbol{\theta}}$ and $\frac{\partial d_i}{\partial \boldsymbol{\theta}^\top}$ given in Lachos et al. (2010c).

References

Andrews DF, Mallows CL (1974) Scale mixtures of normal distributions. J R Stat Soc B 36:99–102

Azzalini A, Dalla-Valle A (1996) The multivariate skew-normal distribution. Biometrika 83:715–726

Branco M, Dey DK (2001) A general class of multivariate skew-elliptical distribution. J Multivar Anal 79:93–113

Cancho VG, Dey DK, Lachos VH, Andrade MG (2010) Bayesian nonlinear regression models with scale mixtures of skew-normal distributions: estimation and case influence diagnostics. Comput Stat Data Anal 55:588–602

Chipkevitch E, Nishimura R, Tu D, Galea-Rojas M (1996) Clinical measurement of testicular volume in adolescents: comparison of the reliability of 5 methods. J Urol 156:2050–2053

Cook RD (1986) Assessment of local influence (with discussion). J R Stat Soc B 48:133–169

Cook RD, Weisberg S (1982) Residuals and influence in regression. Chapman & Hall, London

Ho HJ, Lin TI (2010) Robust linear mixed models using the skew t distribution with application to schizophrenia data. Biom J 52:449–469

Lachos VH, Abanto-Valle CA, Angolini T (2010a) On estimation and local influence analysis for measurement errors models under heavy-tailed distributions. Stat Pap. doi:10.1007/s00362-009-0270-4

Lachos VH, Ghosh P, Arellano-Valle RB (2010b) Likelihood based inference for skew-normal/independent linear mixed models. Stat Sin 20:303–322

Lachos VH, Vilca-Labra FE, Bolfarine H, Ghosh P (2010c) Robust multivariate measurement error models based on scale mixtures of the skew-normal distribution. Statistics 44:541–556

Lee SY, Xu L (2004) Influence analysis of nonlinear mixed-effects models. Comput Stat Data Anal 45: 321–341

Lucas A (1997) Robustness of the Student t based M-estimator. Commun Stat Theory Methods 26: 1165–1182

Magnus JR, Neudecker H (1988) Matrix differential calculus with applications in statistics and econometrics. Wiley, New York

Montenegro LC, Bolfarine H, Lachos VH (2009) Local influence analysis for skew-normal linear mixed models. Commun Stat Theory Methods 38:484–496

Montenegro LC, Lachos VH, Bolfarine H (2010) Inference for a skew extension of the Grubbs model. Stat Pap 51:701–715

Osorio F, Paula GA, Galea-Rojas M (2009) On estimation and influence diagnostics for the Grubbs' model under heavy-tailed distributions. Comput Stat Data Anal 53:1249–1263

Patriota AG, Bolfarine H (2010) Measurement error models with a general class of error distribution. Statistics 44:119–127

Pinheiro JC, Liu CH, Wu YN (2001) Efficient algorithms for robust estimation in linear mixed-effects models using a multivariate t-distribution. J Comput Graph Stat 10:249–276

Vilca-Labra FE, Aoki R, Zeller CB (2010) Hypotheses testing for structural calibration model. Stat Pap. doi:10.1007/s00362-009-0269-x
Xie FC, Wei BC, Lin JG (2009) Homogeneity diagnostics for skew-normal nonlinear regression models. Stat Probab Lett 79:821–827
Zeller CB, Lachos VH, Vilca-Labra FE (2011) Local influence analysis for regression models with scale mixtures of skew-normal distributions. J Appl Stat 8:343–368
Zhu HT, Lee S (2001) Local influence for incomplete-data models. J R Stat Soc B 63:111–126
Zhu HT, Ibrahim JG, Lee SY, Zhang HP (2007) Perturbation selection and influence measures in local influence analysis. Ann Stat 35:2565–2588