Local Influence In Comparative Calibration Models Under Elliptical t-Distributions

Manuel Galea,
Departamento de Estadística, Universidad de Valparaíso, Chile
Heleno Bolfarine,
Departamento de Estatística, Universidade de São Paulo, Brasil
and
Filidor Vilca,
Departamento de Estatística, Universidade Estadual de Campinas, Brasil

Abstract

In this paper we consider applications of local influence (Cook, 1986) to evaluate small perturbations in the model or data set in the context of structural comparative calibration (Bolfarine and Galea, 1995) assuming that the measurements obtained follow a multivariate elliptical distribution. Different perturbation schemes are investigated and an application is considered to a real data set, using the elliptical $t-$distribution.
Key Words: Local influence, Comparative calibration, Elliptical distribution.

1 Introduction

The main object of this paper is the study of local influence and diagnostic in comparative calibration models designed to compare the efficiency of several measuring devices (or instruments) when measuring the same unknown quantity x in a common group of individuals or experimental units. It is assumed that the observed measurements follow a multivariate elliptical distribution. Moreover, the comparative calibration model can be seen as a special case of the general multivariate measurement error model (Fuller, 1987).

Comparing measuring devices which varies in pricing, fastness and other features, such as efficiency, has been of growing interest in several areas like engineering, medicine, psychology and agriculture. Grubbs (1948, 1973) reports data on an experiment designed for comparing three cronometers and Barnett (1969) reports on the comparison of four combinations of two instruments and two operators for measuring vital capacity. Several other examples in the medical area are reported in the literature specially in Kelly (1984, 1985), Chipkevitch et al. (1996) and Lu et al. (1997). Examples in agriculture are considered in Fuller (1987) and in psychology and education in Dunn (1992). Outliers and detection of influent observations is an important step in the analysis of a data set. There are several ways of evaluating the influence of perturbations in the data set and in the model given the parameter estimates. Important reviews can be found in the books by Cook and Weisberg (1982) and Chatterjee and Hadi (1988) and in the paper by Cook (1986). On the other hand, there are just a few works in the literature for diagnostic and influence of observations in models with measurement errors. Kelly (1984) considered a diagnostic procedure in the structural linear model based on the influence function. Tanaka et al. (1991) also consider the influence function introduced by Hampel for evaluating the influence of observations in the analysis of covariance structures. Recently, Zhao and Lee (1998) define leverage of one observation and Cook's distance in a simultaneous equation model. Rather than eliminating cases, the approach proposed by Cook (1986) is a general method for evaluating, under the maximum likelihood estimators, the influence of small perturbations in the model or data set.

Additional results on local influence and applications in linear regression and mixed models can be found in Beckman et al. (1987), Lawrance (1988), Thomas and Cook (1990), Tsai and Wu (1992), Paula (1993), Galea et al. (1997) and Lesaffre and Verbeke (1998). Zhao and Lee (1998) and Kwan and Fung (1998) apply the local influence approach for factor analysis and simultaneous equations under the normality assumption. Recently, Galea et al. (2002a) apply the local influence method in functional and structural comparative calibration models also under the normal distribution assumption. However, no application of local influence has been considered for comparative calibration under elliptical models. Thus, the main object of this paper is to apply the approach of local influence to elliptical measurement error models. As typically considered in the literature, the relevance of using the t-distribution is related to its capability of downweighting influent observations. See, for example, Lange et al. (1989). Several perturbation schemes are considered such as case perturbation and response perturbation. In Section 2 the elliptical structural comparative calibration model is considered and in Section 3 the main concepts of local influence are revised. In Section 4 model curvatures are considered for different perturbation schemes and in Section 5 an illustration of the methodology is presented for a real data set.

2 Elliptical comparative calibration models

Suppose that we have at our disposal $p \geq 2$ instruments for measuring a characteristic of interest x in a group of n experimental units. Let x_i the true (unknown) value in unit i and y_{ij} the measured value obtained with instrument j in unit i, $i = 1, \ldots, n$ and $j = 1, \ldots, p$. A model typically considered in the literature (see, Jaech (1964), Cochran (1968), Barnett (1969), Williams (1969) and Shyr and Gleser (1986)), for such situation is given, in matrix notation, by

$$\begin{aligned} \mathbf{Y}_i &= \boldsymbol{\alpha} + \boldsymbol{\beta} x_i + \mathbf{e}_i \qquad (2.1) \\ &= \boldsymbol{\alpha} + \mathbf{B}\mathbf{U}_i, \end{aligned}$$

where $\boldsymbol{\alpha}=(\alpha_1, \ldots, \alpha_p)^\top$, $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_p)^\top$ are $p \times 1$ vectors, $\mathbf{B} = (\boldsymbol{\beta}, \mathbf{I}_p)$ is a $p \times (p+1)$ matrix, $\mathbf{Y}_i = (y_{i1}, \ldots, y_{ip})^\top$ and $\mathbf{e}_i = (e_{i1}, \ldots, e_{ip})^\top$ are $p \times 1$ random vectors $\mathbf{U}_i = (x_i, \mathbf{e}_i^\top)^\top$ is of dimension $(p+1) \times 1$ and $\mathbf{I}_p$ denotes the identity matrix of dimension p, $i = 1, \ldots, n$. The x_i can be considered as unknown parameters in which case the model is called functional or, it can be considered as independent and identically distributed random variables, in which case the model is called structural. In this paper, we consider the structural version, which is free of incidental parameters. Finally, it is considered that the random vectors $\mathbf{U}_1, \ldots, \mathbf{U}_n$

are independent and identically distributed with a distribution which we denote by $Q_u(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, where

$$\boldsymbol{\mu} = \begin{pmatrix} \mu_x \\ 0 \end{pmatrix} \quad \text{and} \quad \boldsymbol{\Sigma} = \begin{pmatrix} \phi_x & 0 \\ 0 & D(\boldsymbol{\phi}) \end{pmatrix}, \tag{2.2}$$

with $D(\boldsymbol{\phi}) = \mathrm{diag}(\phi_1, \ldots, \phi_p)$. Typically Q_u is considered to be the multivariate normal distribution with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. However, in this paper we consider it to denote a member of the elliptical family of distributions with the normal distribution as an special case. Thus, $\mathbf{Y}_1, \ldots, \mathbf{Y}_n$ are independent and identically distributed with according to the $Q_y(\boldsymbol{\mu}, \mathbf{V})$, where

$$\boldsymbol{\mu} = \boldsymbol{\alpha} + \boldsymbol{\beta}\mu_x = \boldsymbol{\mu}(\boldsymbol{\theta}) \quad \text{and} \quad \mathbf{V} = D(\boldsymbol{\phi}) + \phi_x \boldsymbol{\beta}\boldsymbol{\beta}' = \mathbf{V}(\boldsymbol{\theta}), \tag{2.3}$$

with $\boldsymbol{\theta} = (\mu_x, \boldsymbol{\alpha}^\top, \boldsymbol{\beta}^\top, \phi_x, \boldsymbol{\phi}\top)^\top$.

Let $\overline{\boldsymbol{\theta}} = (\overline{\mu}_x, \overline{\boldsymbol{\alpha}}^\top, \overline{\boldsymbol{\beta}}^\top, \overline{\phi}_x, \overline{\boldsymbol{\phi}}^\top)^\top$, where $\overline{\mu}_x = (\mu_x + d)/c$, $\overline{\boldsymbol{\alpha}} = \boldsymbol{\alpha} - d\boldsymbol{\beta}$, $\overline{\boldsymbol{\beta}} = c\boldsymbol{\beta}$ and $\overline{\phi}_x = \phi_x/c^2$, $c, d \neq 0 \in I\!R$. It can be verified that

$$Q_y(\boldsymbol{\mu}(\boldsymbol{\theta}), \mathbf{V}(\boldsymbol{\theta})) = Q_y(\boldsymbol{\mu}(\overline{\boldsymbol{\theta}}), \mathbf{V}(\overline{\boldsymbol{\theta}})).$$

This result implies that if Q_u (or Q_y) follows an elliptical distribution, then $\boldsymbol{\theta}$ is not identifiable. See Bolfarine and Galea-Rojas (1996). One way of dealing with this problem is to impose restrictions on $\boldsymbol{\theta}$. It is possible (see Barnett, 1969) to consider that there is a reference instrument (denoted instrument 1, without loss of generality) which measures without bias (additive or multiplicative) the quantity of interest. Hence, corresponding to instrument 1, it is considered that $\alpha_1 = 0$ and $\beta_1 = 1$ and under this assumption the model defined by (2.1)-(2.2) becomes identifiable. Thus, the model proposed by Barnett (1969) is given by

$$\boldsymbol{Y}_i = \begin{pmatrix} 0 \\ \boldsymbol{\alpha} \end{pmatrix} + \begin{pmatrix} 1 \\ \boldsymbol{\beta} \end{pmatrix} x_i + \boldsymbol{\varepsilon}_i, i = 1, \ldots, n, \tag{2.4}$$

with $\boldsymbol{\alpha} = (\alpha_2, \ldots, \alpha_p)^\top$ and $\boldsymbol{\beta} = (\beta_2, \ldots, \beta_p)^\top$.

Theobald and Mallison (1978) make model (2.1)-(2.2) for $p \geq 3$ identifiable by considering $\mu_x = 0$ and $\phi_x = 1$, that is, the model is defined as a factor analysis model. More recently, Lu et al. (1997) consider model (2.1)-(2.2) with μ_x and ϕ_x as known so that model becomes identifiable. Indeed, μ_x and ϕ_x are estimated by considering incorporating additional information into the model. In all those references, Q_y is assumed to be the normal distribution.

Replacing $x_i = \mu_x + \sqrt{\phi_x} z_i$, with $z_i \sim (0, 1)$, that is, a random variable with location zero and scale 1, in (2.4) and (2.1) it follows that

$$\boldsymbol{Y}_i = \boldsymbol{\mu} + \boldsymbol{\lambda} z_i + \boldsymbol{\varepsilon}_i, \; i = 1, 2, \ldots, n, \tag{2.5}$$

where $\boldsymbol{\mu} = \binom{0}{\boldsymbol{\alpha}} + \binom{1}{\boldsymbol{\beta}}\mu_x$ and $\boldsymbol{\lambda} = \sqrt{\phi_x}\binom{1}{\boldsymbol{\beta}}$, for the model proposed in Barnett (1969) and $\boldsymbol{\mu} = \boldsymbol{\alpha} + \boldsymbol{\beta}\mu_x$ and $\boldsymbol{\lambda} = \sqrt{\phi_x}\boldsymbol{\beta}$, corresponding to the model considered in Lu et al. (1997).

From the above exposition, it follows that it is possible to work under the elliptical family with model (2.5) which presents computational advantages and from that to obtain results for the model proposed by Barnett (1969) and Lu et al. (1997). As mentioned above, inference for such model under the special case where Q_u is the normal distribution is considered in Barnett (1969), Theobald and Mallison (1978), Bolfarine and Galea (1995) and Lu et al. (1997). The case where Q_u is the Student-t distribution with ν degrees of freedom, has been studied in Bolfarine and Galea-Rojas (1996). In the case of $p = 2$, the model defined by (2.4) or (2.5) corresponds to the simple linear regression model with measurement errors and has been extensively treated in the literature as, for example, in Shyr and Gleser (1986), Fuller (1987), Arellano and Bolfarine (1996) and Cheng and van Ness (1999). More recently, Kwan and Fung (1998) consider the local influence approach in factor analysis models for studying the effect on the maximum likelihood estimators of small perturbation in the data or model.

The elliptical structural comparative calibration model or simply, the elliptical structural model can be defined as:

$$\left.\begin{array}{l} \boldsymbol{Y}_i = \boldsymbol{\mu} + \boldsymbol{\lambda} z_i + \boldsymbol{\varepsilon}_i, i = 1, 2, \ldots, n \\ \\ \boldsymbol{U}_i = \binom{z_i}{\boldsymbol{\varepsilon}_i}, 1 \le i \le n, \text{ are } iid\ \mathbb{E}\ell_{p+1}(\mathbf{0}, \boldsymbol{\Sigma}; g) \end{array}\right\} \tag{2.6}$$

where $\boldsymbol{\Sigma} = \begin{pmatrix} 1 & 0 \\ 0 & D(\boldsymbol{\phi}) \end{pmatrix}$.

This, we have that $\boldsymbol{Y}_1, \boldsymbol{Y}_2, \ldots, \boldsymbol{Y}_n$ are iid $\mathbb{E}\ell_p(\boldsymbol{\mu}, \boldsymbol{V}; g)$, with density given by:

$$f_{\boldsymbol{Y}}(\mathbf{y}) = |\boldsymbol{V}|^{-1/2} g((\boldsymbol{y} - \boldsymbol{\mu})^\top \boldsymbol{V}^{-1}(\boldsymbol{y} - \boldsymbol{\mu})),\ \boldsymbol{y} \in \mathbb{R}^p, \tag{2.7}$$

where $g : \mathbb{R} \mapsto [0, \infty)$, the generator density, is such that $\int_0^\infty u^{p-1} g(u^2) du < \infty$ and $\boldsymbol{V} = \boldsymbol{\lambda}\boldsymbol{\lambda}^\top + D(\boldsymbol{\phi})$ is the scale matrix.

Thus, the main object of this paper is to consider the approach of local influence in the elliptical comparative calibration model given in (2.5)-(2.7).

3 Local Influence

Let $l(\boldsymbol{\theta})$ denote the log-likelihood function from the postulated model (here $\boldsymbol{\theta} = (\boldsymbol{\mu}^\top, \boldsymbol{\lambda}^\top, \boldsymbol{\phi}^\top)^\top$) and let $\boldsymbol{\omega}$ be a $q \times 1$ vector of perturbation restricted to some open subset of $\mathbb{R}^q$. The perturbations are made in the likelihood function such that it takes form $l(\boldsymbol{\theta}|\boldsymbol{\omega})$. Denoting the vector of no perturbation by $\boldsymbol{\omega}_0$, we assume $l(\boldsymbol{\theta}|\boldsymbol{\omega}_0) = l(\boldsymbol{\theta})$. To asses the influence of the perturbations on the maximum likelihood estimate of $\boldsymbol{\theta}$, one may consider the likelihood displacement

$$LD(\boldsymbol{\omega}) = 2[l(\widehat{\boldsymbol{\theta}}) - l(\widehat{\boldsymbol{\theta}}_\omega)],$$

where $\widehat{\boldsymbol{\theta}}_\omega(\widehat{\boldsymbol{\theta}})$ denotes the maximum likelihood estimator under the model $l(\boldsymbol{\theta}|\boldsymbol{\omega})(l(\boldsymbol{\theta}))$.

In some situations, it may be of interest to assess the influence on a subset $\boldsymbol{\theta}_1$ of $\boldsymbol{\theta} = (\boldsymbol{\theta}_1^\top, \boldsymbol{\theta}_2^\top)$. For example, one may have interest on $\boldsymbol{\theta}_1 = \boldsymbol{\lambda}$ or $\boldsymbol{\theta}_1 = \boldsymbol{\phi}$. In such situations, the likelihood displacement can be defined as

$$LD(\boldsymbol{\omega}) = 2[l(\widehat{\boldsymbol{\theta}}) - l(\widehat{\boldsymbol{\theta}}_{1\omega}, \widehat{\boldsymbol{\theta}}_2(\widehat{\boldsymbol{\theta}}_{1\omega}))],$$

where $\widehat{\boldsymbol{\theta}}_{1\omega}$ can be obtained from $\widehat{\boldsymbol{\theta}}_\omega = (\widehat{\boldsymbol{\theta}}_{1\omega}^\top, \widehat{\boldsymbol{\theta}}_{2\omega}^\top)^\top$ and $\widehat{\boldsymbol{\theta}}_2(\widehat{\boldsymbol{\theta}}_{1\omega})$ is the maximum likelihood estimate of $\boldsymbol{\theta}_2$ for $\boldsymbol{\theta}_{1\omega}$ fixed in the perturbed model.

The idea of local influence (Cook, 1986) is concerned in characterizing the behavior of $LD(\boldsymbol{\omega})$ at $\boldsymbol{\omega}_0$. The procedure consists in selecting a unit direction $\boldsymbol{l}$, $||\boldsymbol{l}|| = 1$, and then to consider the plot of $LD(\boldsymbol{\omega}_0 + a\boldsymbol{l})$ against a with $a \in I\!R$. This plot is called *lifted line*. Notice that since $LD(\boldsymbol{\omega}_0) = 0$, $LD(\boldsymbol{\omega}_0 + a\boldsymbol{l})$ has a local minimum at $a = 0$. Each lifted line can be characterized by considering the normal curvature $C_l(\boldsymbol{\theta})$ around $a = 0$. The suggestion is to consider the direction $\boldsymbol{l}_{\max}$ corresponding to the largest curvature $C_{l\max}(\boldsymbol{\theta})$. The index plot of $\boldsymbol{l}_{\max}$ may reveal those observations that under small perturbations exert notable influence on $LD(\boldsymbol{\omega})$. Cook (1986) showed that the normal curvature at the direction $\boldsymbol{l}$ takes the form

$$C_l(\boldsymbol{\theta}) = 2|\boldsymbol{l}^\top \boldsymbol{\Delta}^\top \ddot{\boldsymbol{L}}^{-1} \boldsymbol{\Delta} \boldsymbol{l}|, \tag{3.1}$$

where $-\ddot{\boldsymbol{L}}$ is the observed Fisher information matrix for the postulated model ($\boldsymbol{\omega} = \boldsymbol{\omega}_0$) and $\boldsymbol{\Delta}$ is the $p^* \times q$ matrix with elements

$$\Delta_{ij} = \frac{\partial^2 l(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \theta_i \partial \omega_j},$$

evaluated at $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}$ and $\boldsymbol{\omega} = \boldsymbol{\omega}_0, i = l, ..., p^*$ and $j = 1, ..., q,\quad p^* = 3p$. Therefore, the maximization of (3.1) is equivalent to finding the largest absolute eigenvalue $C_{l\max}$ of the matrix $\boldsymbol{B} = \boldsymbol{\Delta}^\top \ddot{\boldsymbol{L}}^{-1} \boldsymbol{\Delta}$, and $\boldsymbol{l}_{\max}$ is the corresponding eigenvector. For the subset $\boldsymbol{\theta}_1$, the curvature at the direction $\boldsymbol{l}$ is given by

$$C_l(\boldsymbol{\theta}_1) = 2|\boldsymbol{l}^\top \boldsymbol{\Delta}^\top (\ddot{\boldsymbol{L}}^{-1} - \boldsymbol{B}_{22}) \boldsymbol{\Delta} \boldsymbol{l}|,$$

where

$$\boldsymbol{B}_{22} = \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \ddot{\boldsymbol{L}}_{22}^{-1} \end{pmatrix},$$

and $\ddot{\boldsymbol{L}}_{22}$ is obtained from the partition of $\ddot{\boldsymbol{L}}$ according to the partition of $\boldsymbol{\theta}$. The eigenvector $\boldsymbol{l}_{\max}$ corresponds to the largest absolute eigenvalue of the matrix $\boldsymbol{B} = \boldsymbol{\Delta}^\top (\ddot{\boldsymbol{L}}^{-1} - \boldsymbol{B}_{22}) \boldsymbol{\Delta}$.

Other important direction, according to Escobar and Meeker (1992) (see also Verbeke and Molenberghs, 2000) is $\boldsymbol{l} = \boldsymbol{e}_{in}$, which corresponds to the $i-$th position, where there is a one. In that case, the normal curvature, called the total local influence of individual i, is given by $C_i = 2|\boldsymbol{e}_{in}^\top \boldsymbol{B} \boldsymbol{e}_{in}| = 2|b_{ii}|$, where b_{ii} is the $i-$th element diagonal of $\boldsymbol{B}$, $i = 1, ..., n$. Verbeke and Molenberghs (2000) propose consider the $i-$th observation influential if C_i is larger than the cutoff value $2\sum_{i=1}^{q} C_i/q$. We use $\boldsymbol{l}_{\max}$ and C_i as diagnostics for local influence.

Recently, Fung and Kwan (1997) presented an interesting discussion on the application of the local influence for other influence measures than the likelihood displacement. They show that an influence measure, namely $\widehat{T}_\omega$, is scale invariant if $\dot{\mathbf{\Gamma}} = \partial\widehat{T}_\omega/\partial\boldsymbol{\omega}|_{\boldsymbol{\omega}=\boldsymbol{\omega}_0} = \mathbf{0}$. When this derivative is nonzero the ordering among the components of $\boldsymbol{l}_{\max}$ is not necessarily preserved under changes in the scale. In particular, for the likelihood displacement, $\dot{\mathbf{\Gamma}} = \partial l(\widehat{\boldsymbol{\theta}}_\omega)/\partial\boldsymbol{\omega}|_{\boldsymbol{\omega}=\boldsymbol{\omega}_0} = \mathbf{0}$. This property also follows, for instance, for the influence measures proposed in Thomas and Cook (1990) and Paula (1993). But this property is not shared by other influence measures, as pointed out by Fung and Kwan (1997).

4 Curvature Derivation for Elliptical Comparative Calibration Models

In this section we derive the observed information matrix and the $\boldsymbol{\Delta}$ matrix for different schemes of perturbations.

4.1 The observed information matrix

From (2.7) we have that log-likelihood function is given by:

$$l(\boldsymbol{\theta}) = \sum_{i=1}^{n} \ell_i(\boldsymbol{\theta}), \tag{4.1}$$

where $\ell_i(\boldsymbol{\theta}) = -\frac{1}{2}log|\mathbf{V}| + log(g(d_i))$ and $d_i = d_i(\boldsymbol{\theta}) = (\boldsymbol{y}_i - \boldsymbol{\mu})^\top \boldsymbol{V}^{-1}(\boldsymbol{y}_i - \boldsymbol{\mu})$, $i = 1, 2, ..., n$ and $\boldsymbol{V}$ as in (2.7). The matrix of second derivatives with respect to $\boldsymbol{\theta}$ is given by:

$$\boldsymbol{L} = \frac{\partial^2 l(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}^\top}\Big|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}} = \begin{pmatrix} L_{\boldsymbol{\mu}\boldsymbol{\mu}} & L_{\boldsymbol{\mu}\boldsymbol{\lambda}} & L_{\boldsymbol{\mu}\boldsymbol{\phi}} \\ & L_{\boldsymbol{\lambda}\boldsymbol{\lambda}} & L_{\boldsymbol{\lambda}\boldsymbol{\phi}} \\ & & L_{\boldsymbol{\phi}\boldsymbol{\phi}} \end{pmatrix} \tag{4.2}$$

where $\widehat{\boldsymbol{\theta}}$ is the estimator of maximum likelihood of $\boldsymbol{\theta}$. The elements of this matrix are given in the appendix.

4.2 Perturbation of cases weights

We considered the vector of weights $\boldsymbol{\omega} = (\omega_1, ..., \omega_n)^\top$, for weight the contribution of each case in the log-likelihood. Thus the perturbed likelihood is:

$$l(\boldsymbol{\theta}/\boldsymbol{\omega}) = \sum_{i=1}^{n} \omega_i \ell_i(\boldsymbol{\theta}) \tag{4.3}$$

where $\ell_i(\boldsymbol{\theta})$ is defined in (4.1). The vector of the no perturbation is $\boldsymbol{\omega}_0 = (1, ..., 1)^\top = \mathbf{1}$.

Let $\Delta_\gamma = (\Delta_{\gamma_1}, \Delta_{\gamma_2}, ..., \Delta_{\gamma_n})$ the submatrix of $\boldsymbol{\Delta}$ in (3.1), associated to the parameter $\boldsymbol{\gamma}$. That is, Δ_{γ_i} is the $i-$th column of $\Delta_\gamma, i = 1, ..., n$ and $\boldsymbol{\gamma} = \boldsymbol{\mu}, \boldsymbol{\lambda}, \boldsymbol{\phi}$. Consequently, for using (4.1), (4.3) and calculus of vector derivatives (Nel, 1980), we have, after of some computations:

$$\Delta_{\boldsymbol{\mu}_i} = -2W_g(d_i)\boldsymbol{V}^{-1}\boldsymbol{X}_i, \tag{4.4}$$

$$\Delta_{\boldsymbol{\lambda}_i} = -c^{-1}[\mathbf{I}_p + 2W_g(d_i)(D^{-1}(\boldsymbol{\phi})\boldsymbol{X}_i\boldsymbol{X}_i^\top - c^{-1}c_{i1}\boldsymbol{I}_p)]D^{-1}(\boldsymbol{\phi})\boldsymbol{\lambda}, \tag{4.5}$$

$$\Delta_{\boldsymbol{\phi}_i} = -\frac{1}{2}D^{-1}(\boldsymbol{\phi})[\mathbf{1} - c^{-1}D^{-1}(\boldsymbol{\phi})D(\boldsymbol{\lambda})\boldsymbol{\lambda}] + \tag{4.6}$$

$$W_g(d_i)D^{-2}(\boldsymbol{\phi})[-D(\boldsymbol{X}_i)\boldsymbol{X}_i + 2c^{-1}c_{i2}D(\boldsymbol{\lambda})\boldsymbol{X}_i - c^{-2}c_{i1}D(\boldsymbol{\lambda})\boldsymbol{\lambda}],$$

where $D(\boldsymbol{a}) = \mathrm{Diag}(a_1, ..., a_p)$, for $\boldsymbol{a} \in I\!R^p$, $D^{-k}(\boldsymbol{a}) = \mathrm{Diag}(a_1^{-k}, ..., a_p^{-k}), k = 1, 2, 3$, $\mathbf{M} = D^{-1}(\boldsymbol{\phi})\boldsymbol{\lambda}\boldsymbol{\lambda}^\top D^{-1}(\boldsymbol{\phi})$, $c = 1+\boldsymbol{\lambda}^\top D^{-1}(\boldsymbol{\phi})\boldsymbol{\lambda}$, $\mathbf{X}_i = \boldsymbol{Y}_i - \boldsymbol{\mu}$, $W_g(d_i) = \dfrac{g'(d_i)}{g(d_i)}$, $c_{i1} = \boldsymbol{X}_i^\top \boldsymbol{M}\boldsymbol{X}_i$ and $c_{i2} = \boldsymbol{X}_i^\top D^{-1}(\boldsymbol{\phi})\boldsymbol{\lambda}$, $i = 1, ..., n$. Expressions (4.4)-(4.6) are evaluated at the maximum likelihood estimators. Thus, $\boldsymbol{\Delta}$ in (3.1) takes the form

$$\boldsymbol{\Delta} = \begin{pmatrix} \Delta_{\boldsymbol{\mu}} \\ \Delta_{\boldsymbol{\lambda}} \\ \Delta_{\boldsymbol{\phi}} \end{pmatrix}, \tag{4.7}$$

which is of dimension $3p \times n$.

4.3 Perturbation of the observations

In this section, the measured values obtained with the instruments are perturbed. Let $\boldsymbol{Y}_i(\boldsymbol{\omega}_i)$ the perturbation in observation $\boldsymbol{Y}_i$, where $\boldsymbol{\omega}_i = (\omega_{i_1}, ..., \omega_{i_p})^\top$ $i = 1, ..., n$. Some situations of interest in this case are:

(a) Simultaneous perturbations of the measurements of the p instruments:

$$\boldsymbol{Y}_i(\boldsymbol{\omega}_i) = \begin{cases} \boldsymbol{Y}_i + \boldsymbol{\omega}_i, & \text{additive perturbation} \\ \\ \boldsymbol{Y}_i * \boldsymbol{\omega}_i, & \text{multiplicative perturbation} \end{cases}$$

where $*$ denotes Hadamard product.

(b) Perturbing the measurements from one instrument. Suppose that it is of interest perturbing the measurements from one specific instrument, say, k, $k = 1, ..., p$. In this case

$$\boldsymbol{Y}_i(\boldsymbol{\omega}_i) = \begin{cases} \boldsymbol{Y}_i + \boldsymbol{\omega}_i * e_k; & \text{additive perturbation} \\ \\ \boldsymbol{Y}_i * \mathbf{1}_p(\omega_{i_k}), & \text{multiplicative perturbation} \end{cases}$$

where $\boldsymbol{e}_k$ is the k-th unit vector of $I\!R^p$ and $\mathbf{1}_p(\omega_{i_k}) = (1, ..., 1, \omega_{i_k}, 1, ...1)^\top$, is of dimension p.

Note that in the above perturbation schemes there exists $\boldsymbol{\omega}_{i_0}$ such that $\boldsymbol{Y}_i(\boldsymbol{\omega}_{i_0}) = \boldsymbol{Y}_i$, for example, in (a) $\boldsymbol{\omega}_{i_0} = (0, ..., 0)^\top$ in the additive case and $\boldsymbol{\omega}_{i_0} = (1, ..., 1)^\top$ in the multiplicative case.

Let $\boldsymbol{W} = (\boldsymbol{\omega}_1, ..., \boldsymbol{\omega}_n)$ a matrix $p \times n$, whose columns are $\boldsymbol{\omega}_i, i = 1, ..., n$. Denote $\boldsymbol{\omega} = \text{Vec}(\boldsymbol{W}) = (\boldsymbol{\omega}_1^\top, ..., \boldsymbol{\omega}_n^\top)^\top$.

Thus, the perturbed log-likelihood function is given by:

$$l(\boldsymbol{\theta}/\boldsymbol{\omega}) = \sum_{i=1}^{n} \ell_i(\boldsymbol{\theta}/\boldsymbol{\omega}_i), \tag{4.8}$$

where $\ell_i(\boldsymbol{\theta}/\boldsymbol{\omega}_i) = -\frac{1}{2}\ell n|V| + \ell n g(d_i(\boldsymbol{\omega}_i))$, with $d_i(\boldsymbol{\omega}_i) = (\boldsymbol{Y}_i(\boldsymbol{\omega}_i) - \boldsymbol{\mu})^\top \boldsymbol{V}^{-1}(\boldsymbol{Y}_i(\boldsymbol{\omega}_i) - \boldsymbol{\mu})$ and $\boldsymbol{Y}_i(\boldsymbol{\omega}_i)$ as defined in (a) or (b), $i = 1, ..., n$.

Differentiating $L(\boldsymbol{\theta}/\boldsymbol{\omega})$ with respect to $\boldsymbol{\omega}$ and $\boldsymbol{\theta}$ it follows that:

$$\boldsymbol{\Delta} = (\Delta_{\boldsymbol{\theta}_1}, ..., \Delta_{\boldsymbol{\theta}_n}), \tag{4.9}$$

where $\Delta_{\boldsymbol{\theta}_i} = W_g'(d_i(\boldsymbol{\omega}_i))\frac{\partial}{\partial \boldsymbol{\theta}} d_i(\boldsymbol{\omega}_i)\frac{\partial}{\partial \boldsymbol{\omega}_i^\top} d_i(\boldsymbol{\omega}_i) + W_g(d_i(\boldsymbol{\omega}_i))\frac{\partial^2 d_i(\boldsymbol{\omega}_i)}{\partial \boldsymbol{\theta} \partial \boldsymbol{\omega}_i^\top}$, with $\frac{\partial}{\partial \boldsymbol{\theta}} d_i(\boldsymbol{\omega}_i)$ as in the unperturbed case, replacing $\boldsymbol{Y}_i$ for $\boldsymbol{Y}_i(\boldsymbol{\omega}_i)$, $\frac{\partial d_i(\boldsymbol{\omega}_i)}{\partial \boldsymbol{\omega}_i^\top} = 2(\boldsymbol{Y}_i(\boldsymbol{\omega}_i) - \boldsymbol{\mu})^\top \boldsymbol{V}^{-1} \frac{\partial \boldsymbol{Y}_i(\boldsymbol{\omega}_i)}{\partial \boldsymbol{\omega}_i^\top}$, $i = 1, ..., n$ and

$$\frac{\partial^2 d_i(\boldsymbol{\omega}_i)}{\partial \boldsymbol{\theta} \partial \boldsymbol{\omega}_i^\top} = \begin{pmatrix} \dfrac{\partial^2 d_i(\boldsymbol{\omega}_i)}{\partial \boldsymbol{\mu} \partial \boldsymbol{\omega}_i^\top} \\[2ex] \dfrac{\partial^2 d_i(\boldsymbol{\omega}_i)}{\partial \boldsymbol{\lambda} \partial \boldsymbol{\omega}_i^\top} \\[2ex] \dfrac{\partial^2 d_i(\boldsymbol{\omega}_i)}{\partial \boldsymbol{\phi} \partial \boldsymbol{\omega}_i^\top} \end{pmatrix}, \quad i = 1, ..., n.$$

Note that $\Delta_{\boldsymbol{\theta}_i}$ is the matrix $3p \times p, i = 1, ..., n$. The above derivations are evaluated at $\widehat{\boldsymbol{\theta}}$ and $\boldsymbol{\omega}_0$. In the following, expressions are obtained for the matrix $\boldsymbol{\Delta}$ in cases (a) and (b).

Case (a): Let $\boldsymbol{\Delta}_{\gamma i}^a$ ($\boldsymbol{\Delta}_{\gamma i}^m$) the $i-$th submatrix of dimension $p \times p$, of Δ_γ with respect to the additive (multiplicative) perturbation scheme, $\boldsymbol{\gamma} = \boldsymbol{\mu}, \boldsymbol{\lambda}, \boldsymbol{\phi}$ e $i = 1, ..., n$. From (4.9) evaluating in $\boldsymbol{\omega}_0$;

$$\Delta_{\boldsymbol{\mu}_i} = 2(W_g'(d_i)\frac{\partial d_i}{\partial \boldsymbol{\mu}} \boldsymbol{X}_i^\top - W_g(d_i)\boldsymbol{I}_p)\boldsymbol{V}^{-1}, \tag{4.10}$$

$$\Delta_{\boldsymbol{\lambda}_i} = 2W_g'(d_i)\frac{\partial d_i}{\partial \boldsymbol{\lambda}}\boldsymbol{X}_i^\top \boldsymbol{V}^{-1} \tag{4.11}$$

$$+2c^{-1}W_g(d_i)(2\boldsymbol{V}^{-1}\boldsymbol{X}_i\boldsymbol{\lambda}^\top - D^{-1}(\boldsymbol{\phi})\boldsymbol{X}_i\boldsymbol{\lambda}^\top + c_{i2}\boldsymbol{I}_p)D^{-1}(\boldsymbol{\phi}),$$

$$\Delta_{\boldsymbol{\phi}_i} = 2W_g'(d_i)\frac{\partial d_i}{\partial \boldsymbol{\phi}}\boldsymbol{X}_i^\top \boldsymbol{V}^{-1} - 2W_g(d_i)(D^{-1}(\boldsymbol{\phi})D(\boldsymbol{X}_i)\boldsymbol{V}^{-1} \tag{4.12}$$

$$+c^{-2}D^{-2}(\boldsymbol{\phi})(D(\boldsymbol{\lambda})\boldsymbol{\lambda}\boldsymbol{X}_i^\top \boldsymbol{M} - c^{-1}c_{i2}D^{-2}(\boldsymbol{\phi})D(\boldsymbol{\lambda})),\ i = 1, ..., n.$$

Thus, evaluating (4.10)-(4.12) at the maximum likelihood estimator $\widehat{\boldsymbol{\theta}}$ of $\boldsymbol{\theta}$ it follows in the additive case that

$$\Delta^a = (\Delta^a_{\boldsymbol{\theta}_1}, ..., \Delta^a_{\boldsymbol{\theta}_n}), \tag{4.13}$$

where

$$\Delta^a_{\boldsymbol{\theta}_i} = \begin{pmatrix} \Delta^a_{\boldsymbol{\mu}_i} \\ \Delta^a_{\boldsymbol{\lambda}_i} \\ \Delta^a_{\boldsymbol{\phi}_i} \end{pmatrix},\ i = 1, ..., n.$$

With respect to the multiplicative scheme, it follows that

$$\Delta^m_{\gamma_i} = \Delta^a_{\gamma_i} D(\boldsymbol{Y}_i)\ ;\ \boldsymbol{\gamma} = \boldsymbol{\mu}, \boldsymbol{\lambda}, \boldsymbol{\phi},\ i = 1, ..., n \tag{4.14}$$

Case (b): Without loss of generality, we can take $k = 1$. To obtain the matrix $\boldsymbol{\Delta}$ in the additive scheme, we can multiply $\Delta^a_{\gamma_i}, \boldsymbol{\gamma} = \boldsymbol{\mu}, \boldsymbol{\lambda}, \boldsymbol{\phi}$ given in (4.14) from the right by $e_1 = (1, 0, ..., 0)^\top \in I\!R^p$. Analogously, in the multiplicative scheme, it sufficient multiplying $\Delta^m_{\boldsymbol{\mu}_i}$, given in (4.14), from the right by the same expression. Expressions for $W_g'(d)$ and $W_g(d)$ are obtained in Galea, Paula and Bolfarine (1997) for some distributions in the elliptical family.

4.4 Perturbation of the degrees of freedom in the Student t-model

A special case of the elliptical comparative calibration model (2.6)-(2.7), is the Student-t, which follows by assuming that $\boldsymbol{U}_i \sim t_{p+1}(\boldsymbol{0}, \boldsymbol{\Sigma}; \nu),\ \nu > 0$. Then $\boldsymbol{Y}_i \sim t_p(\boldsymbol{\mu}, \boldsymbol{V}; \nu)$ with density function given by:

$$f_{\boldsymbol{Y}}(\boldsymbol{y}) = k(p, \nu)(\nu + (\boldsymbol{y} - \boldsymbol{\mu})^\top \boldsymbol{V}^{-1}(\boldsymbol{y} - \boldsymbol{\mu})^{-\frac{1}{2}(\nu+p)} \tag{4.15}$$

where $k(p, \nu) = \Gamma\left(\frac{p+\nu}{2}\right)\nu^{\nu/2}/\Gamma(\nu/2)\pi^{p/2}$. For this special model, $W_g(d) = -\frac{1}{2}(\nu + p)(\nu + d)^{-1}$ and $W_g'(d) = \frac{1}{2}(\nu + p)(\nu + d)^{-2}$.

Recently, several authors have considered the multivariate Student-t distribution as an alternative to the normal model because it can naturally accommodate outliers present in the data. Thus, the Student-t model provides a robust procedure for analysing data sets which may present outliers. Rubin (1983) obtain maximum likelihood estimators for the parameters of the multivariate Student-t model by using the EM-algorithm; Little (1988) extends the results in Rubin (1983) for the case of incomplete data sets, that is, data sets with missing data. Sutradhar and Ali (1986) consider maximum likelihood estimation in the multivariate Student-t regression model. Lange et al. (1989) discuss the use of the Student-t model in regression and in problems related to multivariate analysis. Taylor (1992) considers some other aspects of the Student-t model. More recently, Sutradhar (1993) has considered an score test aiming at testing if the covariance matrix is equal to some specified covariance matrix (diagonal, for example), using the Student-t distribution; Bolfarine and Arellano (1994) introduce Student-t functional and structural measurement error models and Bolfarine and Galea-Rojas (1996) use the Student-t distribution in structural comparative calibration models. The Student-t distribution incorporates an additional parameter ν, namely the degrees of freedom parameter, which allows adjusting for the kurthosis of the distribution. This parameter can be fixed previously and Lange et al. (1989) and Berkane et al. (1994) recommend $\nu = 4$, or, otherwise, get information for it from the data set. In this section ν is considered known and the following perturbed model is considered:

$$\boldsymbol{Y}_i \sim t_p(\boldsymbol{\mu}, \boldsymbol{V}; \nu_0 h(\omega_i)), \tag{4.16}$$

with the $\mathbf{Y}_i$ being independent, $i = 1, ..., n$, where h is a positive and differentiable and further, there exists ω_{0_i} such that $h(\omega_{0_i}) = 1$. Under the perturbed model, the log-likelihood function is given by

$$l(\boldsymbol{\theta}/\boldsymbol{\omega}) = \sum_{i=1}^{n} \ell_i(\boldsymbol{\theta}/\omega_i), \tag{4.17}$$

where $\ell_i(\boldsymbol{\theta}/\omega_i) = \ell nk(\nu_i, p) - \dfrac{1}{2}\ell n|\boldsymbol{V}| - \dfrac{1}{2}(\nu_i + p)\ell n(\nu_i + d_i)$, where $\nu_i = \nu_0 h(\omega_i)$, $k(\nu_i, p)$ as in (4.15) with ν replaced by ν_i and d_i as defined in (4.1), $i = 1, ..., n$. According to our notation, $\boldsymbol{\omega} = (\omega_1, ..., \omega_n)^\top$. Thus, following similar procedures as considered in Section 4.3, it follows that:

$$\boldsymbol{\Delta} = (\Delta_{\boldsymbol{\theta}_1}, ..., \Delta_{\boldsymbol{\theta}_n}), \tag{4.18}$$

where

$$\Delta_{\boldsymbol{\theta}_i} = \nu_0 h'(\omega_{0_i}) \left\{ \frac{1}{\nu_0 + p} W_g(d_i) + W_g'(d_i) \right\} \frac{\partial d_i}{\partial \boldsymbol{\theta}},$$

$i = 1, ..., n$. Expression (4.18) should be evaluated at the maximum likelihood estimate $\widehat{\boldsymbol{\theta}}$. The function h can be defined as in Escobar and Meeker (1992), namely, $h(\omega_i) = a^{\omega_i}$, with $a > 0$ and $\omega_i \in [-1, 1], i = 1, ..., n$. Thus, $\nu_i = \nu_0 h(\omega_i)$ takes values in the interval $[\nu_0/a, a\nu_0]$. For example, we can take $a = 2$ and $h'(\omega_{0_i}) = \ell n2$, $i = 1, ..., n$. If $\omega_i = \omega$ for all $i = 1, ..., n$; $h(\omega)$ is an scalar type function. In this case, considering a Taylor expansion of order 2 of $l(\widehat{\boldsymbol{\theta}}_\omega)$ around $\boldsymbol{\omega} = \boldsymbol{\omega}_0$ it follows that

$$LD_1(\omega) \cong \boldsymbol{\Delta}^\top(-\ \boldsymbol{L})^{-1}\boldsymbol{\Delta}(\omega - \omega_0)^2, \tag{4.19}$$

where

$$\mathbf{\Delta} = \nu_0 h'(\omega_0) \sum_{i=1}^{n} (W_g'(d_i) + \frac{W_g(d_i)}{\nu_0 + p}) \frac{\partial d_i}{\partial \boldsymbol{\theta}},$$

which is of dimension $3p \times 1$. Note that $\lim_{\nu_0 \to \infty} C_l(\boldsymbol{\theta}) = 0$, for all $\boldsymbol{l}$, $||\boldsymbol{l}|| = 1$, meaning that for large ν_0 (close to normality) there are no directions of local influence, which is reasonable since the normal model is independent of ν_0. Thus in the applications presented below, we consider only the case of small ν_0.

4.5 Application

In this section we analyze one real data set given in Barnett (1969). Two instruments used for measuring the vital capacity of human lung and operated by skilled and unskilled operators were compared on a common group of 72 patients. We will focus on the parameter set $\boldsymbol{\theta}$, in the $t-$model, perturbation of the degrees of freedom and perturbation of cases. All computations were performed in S-Plus.

Figures 1 and 2 present graphics of local influence for the perturbation of case weights for several degrees of freedom. As expected for small degrees of freedom there are no local influent observations on the maximum likelihood estimators.

However, as ν increases (close to the normal model), some observations (23, 30, 58, 67) present significant influence on the maximum likelihood estimators, as was also verified in Galea et al. (2002a) for the normal model. This shows that the t-model with small degrees of freedom can be very useful for accommodating influent observations present on the data sets, which is not the case with normal models.

Similar results were also obtained in Galea et al. (2002b) in the structural error in variables models using a t-distribution.

Figures 3, 4 and 5 present graphics of local influence for perturbation of the degrees of freedom parameter. Note that for $\nu_0 = 1$ (Cauchy model), model perturbation with degrees of freedom around one, that is, $\nu \in [1/2, 2]$, observation 45 yields the largest local influence on the maximum likelihood estimators. On the other hand, for $\nu_0 = 4$, the maximum likelihood estimators are quite stable with respect to small perturbation on the degrees of freedom parameter. Thus, for the present data set $\nu_0 = 4$ seems to be the most adequate value of parameter degrees of freedom. This conclusion was also reached by Lange et al. (1989) for several of the data set they have analyzed using different procedures of model identification. Thus, the local influence approach can also be useful in the appropriate selection of the degrees of freedom parameter.

Figure 1: Index plot of $\boldsymbol{l}_{\max}$ for perturbation of case weights

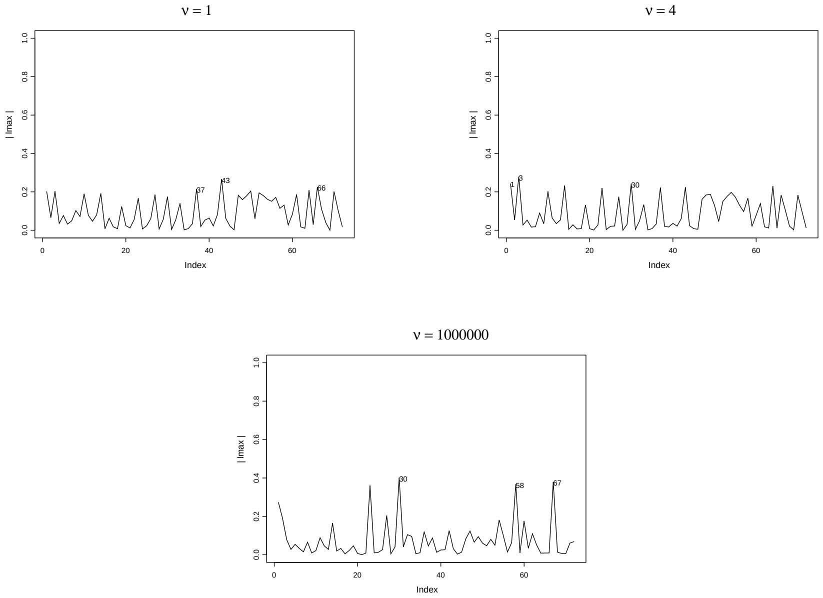


Figure 2: Index plot of C_i for perturbation of case weights

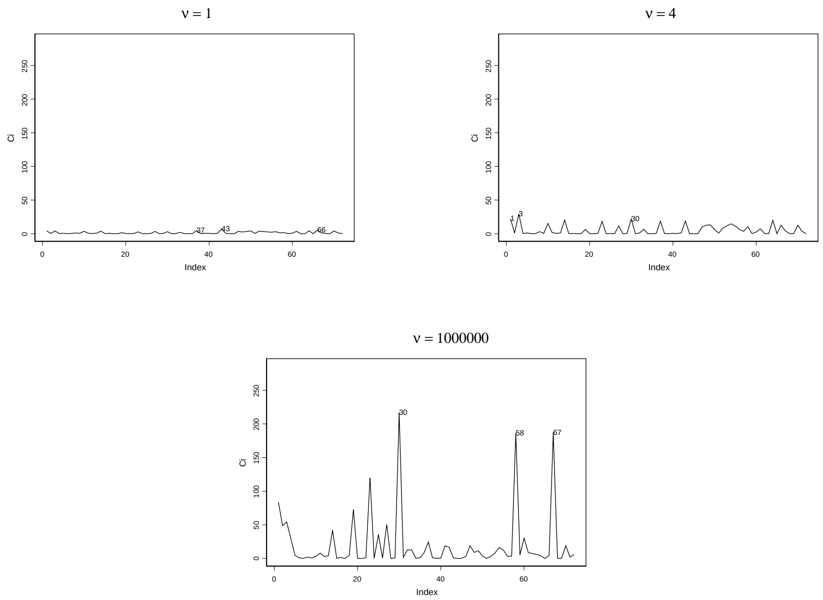


Figure 3: Index plot of $\boldsymbol{l}_{\max}$ for perturbation of the degrees of freedom

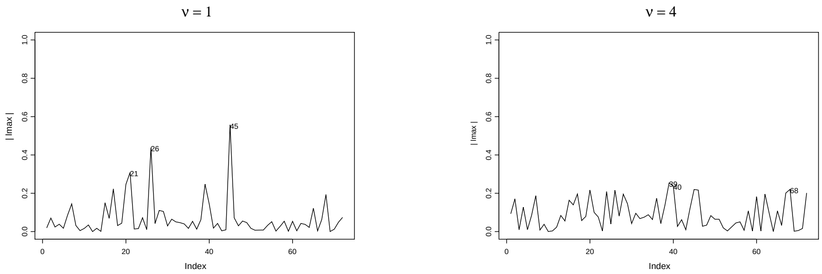


Figure 4: Index plot of C_i for perturbation of the degrees of freedom

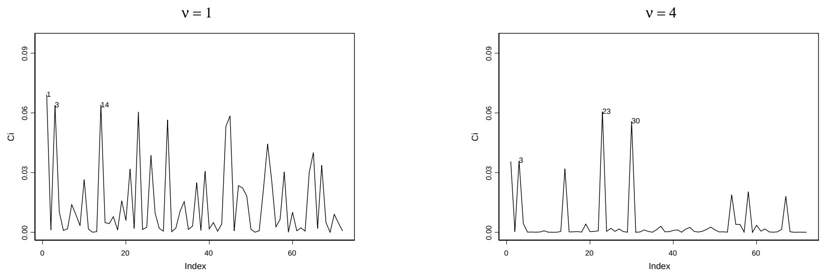
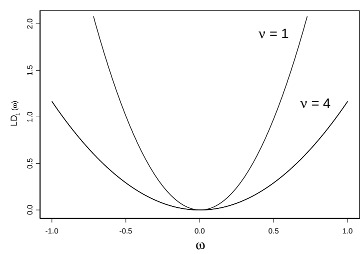


Figure 5: Plots of the likelihood displacement $LD_1(\omega)$ versus ω



Appendix: Computing the observed information matrix in the elliptical structural model

In this appendix the observed information matrix is obtained for the elliptical structural model. From (4.1), it follows that

$$\frac{\partial \ell_i(\boldsymbol{\theta})}{\partial \boldsymbol{\gamma}} = -\frac{1}{2}\frac{\partial \ell n|\boldsymbol{V}|}{\partial \boldsymbol{\gamma}} + W_g(d_i)d_{i\gamma}, \tag{A.1}$$

with $d_{i\gamma} = \dfrac{\partial d_i}{\partial \boldsymbol{\gamma}}, \boldsymbol{\gamma} = \boldsymbol{\mu}, \boldsymbol{\lambda}, \boldsymbol{\phi}$ and $d_i = \boldsymbol{X}_i^\top V^{-1}\boldsymbol{X}_i,\ \boldsymbol{X}_i = \boldsymbol{Y}_i - \boldsymbol{\mu}, i = 1, ..., n$. Further, using results in Nel (1980) related to vector derivatives it follows that,

$$\left.\begin{aligned} &\frac{\partial \ell n|\boldsymbol{V}|}{\partial \boldsymbol{\mu}} = 0, \frac{\partial \ell n|\boldsymbol{V}|}{\partial \boldsymbol{\lambda}} = 2c^{-1}D^{-1}(\boldsymbol{\phi})\boldsymbol{\lambda}, \\ &\frac{\partial \ell n|\boldsymbol{V}|}{\partial \boldsymbol{\phi}} = D^{-1}(\boldsymbol{\phi})\mathbf{1} - c^{-1}D^{-2}(\boldsymbol{\phi})D(\boldsymbol{\lambda})\boldsymbol{\lambda}, \end{aligned}\right\} \tag{A.2}$$

$$\begin{aligned} d_{i\mu} &= -2\boldsymbol{V}^{-1}\boldsymbol{X}_i, & \text{(A.3)}\\ d_{i\lambda} &= -2c^{-1}D^{-1}(\boldsymbol{\phi})\boldsymbol{X}_i\boldsymbol{X}_i^\top D^{-1}(\boldsymbol{\phi})\boldsymbol{\lambda} + 2c^{-2}c_{i1}D^{-1}(\boldsymbol{\phi})\boldsymbol{\lambda}, & \text{(A.4)}\\ d_{i\phi} &= -D^{-2}(\boldsymbol{\phi})D(\boldsymbol{X}_i)\boldsymbol{X}_i + 2c^{-1}c_{i2}D^{-2}(\boldsymbol{\phi})D(\boldsymbol{\lambda})\boldsymbol{X}_i & \text{(A.5)}\\ &\quad -c^{-2}c_{i1}D^{-2}(\boldsymbol{\phi})D(\boldsymbol{\lambda})\boldsymbol{\lambda}, i = 1, 2, ..., n \end{aligned}$$

From (A.1) it follows that the observed, per element, information matrix is given by

$$L_i = L_i(\boldsymbol{\theta}/\boldsymbol{Y}_i) = -\begin{bmatrix} \dfrac{\partial^2 \ell_i}{\partial \boldsymbol{\mu}\partial \boldsymbol{\mu}^\top} & \dfrac{\partial^2 \ell_i}{\partial \boldsymbol{\mu}\partial \boldsymbol{\lambda}^\top} & \dfrac{\partial^2 \ell_i}{\partial \boldsymbol{\mu}\partial \boldsymbol{\phi}^\top} \\ & \dfrac{\partial^2 \ell_i}{\partial \boldsymbol{\lambda}\partial \boldsymbol{\lambda}^\top} & \dfrac{\partial^2 \ell_i}{\partial \boldsymbol{\lambda}\partial \boldsymbol{\phi}^\top} \\ & & \dfrac{\partial^2 \ell_i}{\partial \boldsymbol{\phi}\partial \boldsymbol{\phi}^\top} \end{bmatrix}, \tag{A.6}$$

$i = 1, ..., n$, where

$$\frac{\partial^2 \ell_i}{\partial \boldsymbol{\gamma}\partial \boldsymbol{\tau}^\top} = -\frac{1}{2}\frac{\partial^2 \ell n|\boldsymbol{V}|}{\partial \boldsymbol{\gamma}\partial \boldsymbol{\tau}^\top} + W_g'(di)d_{i\boldsymbol{\gamma}}d_{i\boldsymbol{\tau}^\top} + W_g(d_i)d_{i\boldsymbol{\gamma}\boldsymbol{\tau}^\top}, \tag{A.7}$$

with $d_{i\boldsymbol{\gamma}\boldsymbol{\tau}^\top} = \dfrac{\partial^2 d_i}{\partial \boldsymbol{\gamma}\partial \boldsymbol{\tau}^\top},\ i = 1, 2, ..., n$ and $\boldsymbol{\gamma}, \boldsymbol{\tau} = \boldsymbol{\mu}, \boldsymbol{\lambda}, \boldsymbol{\phi}$, where

$$\frac{\partial^2 \ell n|\boldsymbol{V}|}{\partial \boldsymbol{\mu} \partial \boldsymbol{\mu}^\top} = \frac{\partial^2 \ell n|\boldsymbol{V}|}{\partial \boldsymbol{\mu} \partial \boldsymbol{\lambda}^\top} = \frac{\partial^2 \ell n|\boldsymbol{V}|}{\partial \boldsymbol{\mu} \partial \boldsymbol{\phi}^\top} = 0, \tag{A.8}$$

$$\frac{\partial^2 \ell n|\boldsymbol{V}|}{\partial \boldsymbol{\lambda} \partial \boldsymbol{\lambda}^\top} = 2c^{-1}(\boldsymbol{V}^{-1} - c^{-1}\boldsymbol{M}), \tag{A.9}$$

$$\frac{\partial^2 \ell n|\boldsymbol{V}|}{\partial \boldsymbol{\lambda} \partial \boldsymbol{\phi}^\top} = -2c^{-1}\boldsymbol{V}^{-1}D^{-1}(\boldsymbol{\phi})D(\boldsymbol{\lambda}), \tag{A.10}$$

$$\begin{aligned} &\frac{\partial^2 \ell n|\boldsymbol{V}|}{\partial \boldsymbol{\phi} \partial \boldsymbol{\phi}^\top} = -D^2(\boldsymbol{\phi}) + 2c^{-1}D^{-3}(\boldsymbol{\phi})D^2(\boldsymbol{\lambda}) \\ &-c^{-2}D^{-1}(\boldsymbol{\phi})D(\boldsymbol{\lambda})\boldsymbol{M}D(\boldsymbol{\lambda})D^{-1}(\boldsymbol{\phi}), \end{aligned} \tag{A.11}$$

$$d_{i\boldsymbol{\mu}\boldsymbol{\mu}^\top} = 2\boldsymbol{V}^{-1}, \tag{A.12}$$

$$d_{i\boldsymbol{\mu}\boldsymbol{\lambda}^\top} = 2c^{-1}(D^{-1}(\boldsymbol{\phi})\boldsymbol{\lambda}\boldsymbol{X}_i^\top D^{-1}(\boldsymbol{\phi}) + c_{i1}(\boldsymbol{V}^{-1} - c^{-1}\boldsymbol{M})), \tag{A.13}$$

$$\begin{aligned} d_{i\boldsymbol{\mu}\boldsymbol{\phi}^\top} = {} & 2D^{-2}(\boldsymbol{\phi})D(\boldsymbol{X}_i) - 2c^{-1}D^{-1}(\boldsymbol{\phi})\boldsymbol{\lambda}\boldsymbol{X}_i^\top D(\boldsymbol{\lambda})D^{-2}(\boldsymbol{\phi}) \\ & -2c^{-1}c_{i2}D^{-2}(\boldsymbol{\phi})D(\boldsymbol{\lambda}) + 2c^{-2}\boldsymbol{M}\boldsymbol{X}_i\boldsymbol{\lambda}^\top D(\boldsymbol{\lambda})D^{-2}(\boldsymbol{\phi}), \end{aligned} \tag{A.14}$$

$$\begin{aligned} d_{i\boldsymbol{\lambda}\boldsymbol{\lambda}^\top} = {} & -2c^{-1}D^{-1}(\boldsymbol{\phi})\boldsymbol{X}_i\boldsymbol{X}_i^\top D^{-1}(\boldsymbol{\phi}) + 4c^{-2}D^{-1}(\boldsymbol{\phi})\boldsymbol{X}_i\boldsymbol{X}_i^\top \boldsymbol{M} \\ & +2c^{-2}c_{i1}(\boldsymbol{V}^{-1} - 3c^{-1}\boldsymbol{M}) + 4c^{-2}\boldsymbol{M}\boldsymbol{X}_i\boldsymbol{X}_i^\top D^{-1}(\boldsymbol{\phi}), \end{aligned} \tag{A.15}$$

$$\begin{aligned} d_{i\boldsymbol{\lambda}\boldsymbol{\phi}^\top} = {} & 2c^{-1}D^{-1}(\boldsymbol{\phi})\boldsymbol{X}_i\boldsymbol{X}_i^\top D(\boldsymbol{\lambda})D^{-2}(\boldsymbol{\phi}) && \text{(A.16)} \\ & -2c^{-2}D^{-1}(\boldsymbol{\phi})\boldsymbol{X}_i\boldsymbol{X}_i^\top \boldsymbol{M}D^{-1}(\boldsymbol{\phi})D(\boldsymbol{\lambda}) + 2c^{-1}c_{i2}D^{-2}(\boldsymbol{\phi})D(\boldsymbol{X}_i) \\ & -2c^{-2}c_{i1}D^{-2}(\boldsymbol{\phi})D(\boldsymbol{\lambda}) + 4c^{-3}c_{i1}\boldsymbol{M}D^{-1}(\boldsymbol{\phi})D(\boldsymbol{\lambda}) \\ & -4c^{-2}\boldsymbol{M}\boldsymbol{X}_i\boldsymbol{X}_i^\top D(\boldsymbol{\lambda})D^{-2}(\boldsymbol{\phi}), && \text{(A.17)} \end{aligned}$$

$$\begin{aligned} d_{i\boldsymbol{\phi}\boldsymbol{\phi}^\top} = {} & 2D^{-3}(\boldsymbol{\phi})D^2(\boldsymbol{X}_i) - 4c^{-1}c_{i2}D^{-3}(\boldsymbol{\phi})D(\boldsymbol{\lambda})D(\boldsymbol{X}_i) \\ & -2c^{-1}D^{-2}(\boldsymbol{\phi})D(\boldsymbol{\lambda})\boldsymbol{X}_i\boldsymbol{X}_i^\top D(\boldsymbol{\lambda})D^{-2}(\boldsymbol{\phi}) \\ & +2c^{-2}D^{-2}(\boldsymbol{\phi})D(\boldsymbol{\lambda})\boldsymbol{X}_i\boldsymbol{X}_i^\top \boldsymbol{M}D^{-1}(\boldsymbol{\phi})D(\boldsymbol{\lambda}) \\ & +2c^{-2}c_{i1}D^{-3}(\boldsymbol{\phi})D^2(\boldsymbol{\lambda}) \\ & -2c^{-3}c_{i1}D^{-1}(\boldsymbol{\phi})D(\boldsymbol{\lambda})\boldsymbol{M}D(\boldsymbol{\lambda})D^{-1}(\boldsymbol{\phi}) \\ & +2c^{-2}D^{-1}(\boldsymbol{\phi})D(\boldsymbol{\lambda})\boldsymbol{M}\boldsymbol{X}_i\boldsymbol{X}_i^\top D(\boldsymbol{\lambda})D^{-2}(\boldsymbol{\phi}), \end{aligned} \tag{A.18}$$

$i = 1, 2, ..., n$. Thus, the complete observed information matrix is $L_{ob}(\boldsymbol{\theta}/\boldsymbol{Y}) = \sum_{i=1}^{n} L_i(\boldsymbol{\theta}/\boldsymbol{Y}_i)$. Evaluating the observed information matrix at $\widehat{\boldsymbol{\theta}}$ it follows that $L_{ob}(\widehat{\boldsymbol{\theta}}/\boldsymbol{Y}) = -\boldsymbol{L}$ given in (4.2).

Acknowledgements

The authors acknowledges the partial financial support from CNPq-Brasil and Project Fondecyt 1000424-Chile.

References

Arellano, R. and Bolarine, H. (1996). A note on the simple structural regression model. *Annals of the Institute of Statistical Mathematics*, 48, 111-125.

Barnett, V.D. (1969). Simultaneous pairwise linear structural relationships. *Biometrics*, 25, 129-142.

Bechman, R., Nachtsheim, C. and Cook, R. (1987). Diagnostics for mixed-model analysis of variance. *Technometrics*, 29, 413-426.

Berkane, M., Kano, Y. and Bentler, P.M.(1994). Pseudo maximun likelihood estimation in elliptical theory: Effects of misspecification. *Computational Statistics & Data Analysis*, 18, 255-267.

Bolfarine, H. and Arellano-Valle, R. B.(1994). Robust modeling in measurement error models using the Student-t distributions. *Brazilian Journal of Probability and Statistics*, 8, 67-84.

Bolfarine, H. and Galea, M.(1995). Structural comparative calibration using the EM algorithm. *Journal of Applied Statistics*, 22, 277-292.

Bolfarine, H. and Galea-Rojas, M.(1996). One structural comparative calibration under a t-models. *Computational Statistics*, 11, 63-85.

Chatterjee, S. Hadi, A.S. (1988). *Sensivity Analysis in Linear Regression*, John Wiley. New York.

Cheng, C. and van Ness, J. (1999). *Statistical Regression With Measurement Error*, Arnold. London.

Chipkevitch, E., Nishimura, R., Tu, D. and Galea-Rojas (1996). Clinical measurement of testicular volume in adolescents: Comparison of the reliability of 5 methods. *Journal of Urology*, 156, 2050-2053.

Cochran, W.G. (1968). Errors of measurement in statistics. *Technometrics*, 10, 637-666.

Cook, R.D. (1986). Assessment of local influence. *Journal of the Royal Statistical Society*, B, 48, 133-169.

Cook, R.D. and Weisberg, S. (1982). *Residuals and Influence in Regression*, Chapman and Hall. London.

Dunn, G. (1992). *Design and Analysis of Reliability: The statistical evaluation of measurement errors*. Edward Arnold. New York.

Escobar, E. and Meeker, W. (1992). Assessing influence in regression analysis with censored data. *Biometrics*, 48, 507-528.

Fuller, W.A. (1987). *Measurement error models*. Wiley, New York.

Fung, W. and Kwan, C.(1997). A note on local influence based on normal curvature. *Journal of the Royal Statistical Society*, B, 59, 839-843.

Galea, M., Paula, G.A. and Bolfarine, H. (1997). Local influence in elliptical linear regression models. *The Statistician*, 46, 71-79.

Galea, M., Bolfarine, H. and de Castro, M. (2002a). Local influence in comparative calibration models. *Biometrical Journal*, 44, 59-81.

Galea, M., Bolfarine, H. and Vilca, F. (2002b). Influence diagnostics for structural errors-in-variables model under the student$-t$ distribution. *Journal of Applied Statistics*. To appear.

Grubbs, F.E. (1948). On estimating precision of measuring instruments and product variability. *Journal of the American Statistical Association*, 43, 243-264.

Grubbs, F.E. (1973). Errors of measurement, precision, accuracy and the statistical comparison of measuring instruments. *Technometrics* 15, 53-66.

Jaech, J.L. (1964). A program to estimate measurement errors in nondestructive avaluation of fuel element quality. *Technometrics*, 6, 293-300.

Kaaks, R., Riboli, E., Estive, J., van Kappel and van Staveren (1994). Estimating the accuracy of questionnaire assessments: Validation in terms of strutural equation models. *Statistics in Medicine*, 13, 127-142.

Kelly, G. (1984).The influence function in the errors in variables problem. *The Annals of Statistics*, 12, 87-100.

Kelly, G. (1985). Use of the structural equations model in assessing the reliability of a new measurement technique. *Applied Statistics*, 34, 258-263.

Kwan, C. and Fung, W. (1998). Assessing local influence for specific restricted likelihood: Application to factor analysis. *Psychometrika*, 63, 35-46.

Lange, K.L., Little, R.J. and Taylor, J. (1989). Robust statistical modelling using the t-distribution. *Journal of the American Statistical Association*, 84, 881-896.

Lawrance, A.J. (1988). Regression transformation diagnostic using local influence. *Journal of the American Statistical Association*, 83, 1067-1072.

Lesaffre, E. and Verbeke, G. (1998). Local influence in linear mixed models. *Biometrics*, 54, 570-583.

Little, R.J. (1988). Robust estimation of the mean and covariance matrix from data with missing values. *Applied Statistics*, 37, 23-38.

Lu, Y., Ye, K., Mathur, A., Hui, S., Fuerst, T. and Genant, H. (1997). Comparative calibration without a gold standard. *Statistics in Medicine*, 16, 1889-1905.

Paula, G.A. (1993). Assessing local influence in restricted regressions models. *Computational Statistics* and *Data Analysis* 16, 63-79.

Rubin. D.B. (1983). Iteratively reweighted least squares. *Encyclopedia of the Statistical Sciences*, 4, 272-275.

Sutradhar, B.C. (1993). Score tests for the covariance matrix of the elliptical t-distribution. *Journal of Multivariate Analysis*, 46, 1-12.

Shyr, J. and Gleser, L. (1986). Inference about comparative precision in linear structural relationships. *Journal of Statistical Planning and Inference*, 14, 339-358.

Sutradhar, B.C. and Ali, M.M. (1996). Estimation of the parameter of a regression model with a multivariate t error variable. *Communications in Statistics - Theory and Methods*, 15, 429-450.

Tanaka, Y., Watadani, S. and Moon, S. (1991). Influence in covariance structure analysis with an application to confirmatory factor analysis. *Commutations in Statistics-Theory and Methods*, 20, 3805-3821.

Taylor, J. (1992). Properties of modelling the error distribution with an extra shape parameter. *Computational Statistics & Data Analysis*, 13, 33-46.

Theobald, C.M. and Mallison, J.R. (1978). Comparative callibration, linear structural relationship and congeneric measurements. *Biometrics*, 34, 39-45.

Thomas, W. and Cook, R.D. (1990). Assessing influence on predictions from generalized linear models. *Technometrics*, 32, 59-65.

Tsai, C.L. and Wu, X. (1992). Assessing local influence in linear regression models with first-order autoregressive or heteroscedastic error structure. *Statistics and Probability Letters*, 14, 247-252.

Verbeke, G. and Molenberghs, G. (2000). *Linear mixed models for longitudinal data.* Springer. New York.

Williams E.J. (1969). Regression methods in calibration problems. *Bulletin of the International Statistical Institute*, 43, 17-28.

Zhao, Y. and Lee, A. (1998). Influence diagnostics for simultaneous equations models. *Australian and New Zealand Journal Statistics*, 40, 345-357.