Influence Diagnostics in the Capital Asset Pricing Model Under Elliptical Distributions

Manuel Galea
Departamento de Estadística, Universidad de Valparaíso, Chile
José A. Díaz-García
Departamento de Estatística y Cálculo, Universidad Autónoma Agraria Antonio Narro, Saltillo, México
and
Filidor Vilca
Departamento de Estatística, Universidade Estadual de Campinas, Brasil

Resumen

In this paper we consider the Capital Asset Pricing Model under Elliptical (symmetric) Distributions. This class of distributions, which contain the normal distribution, $\boldsymbol{t}$, contaminated normal and power exponential, among others, offers a more flexible framework for modelling asset prices or returns. In order to analyze the sensibility to possible outliers and/or atypical returns of the maximum likelihood estimators, the local influence method was implemented. The results are illustrated by using a set of shares from companies who trade in the Chilean Stock Market. Our main conclusion is that symmetric distributions having heavier tails than those of the normal distribution, especially the $\boldsymbol{t}$ distribution with small degrees of freedom, show a better fit and allow the reduction of the influence of atypical returns in the maximum likelihood estimators.
Key Words: Robust Estimation, Diagnostics, Local Influence, Elliptical Distributions.

1 Introduction

Much of the current theory of capital asset pricing is based on the assumption that excess returns (or returns) follow a multivariate normal distribution. In effect, it is usual that the Capital Asset Pricing Model, $CAPM$, and other financial models are based on the normal distribution, as can be observed in Broquet (1992), Elton and Gruber (1995), Van Horne (1997), Campbell *et al.* (1997), and Ross *et al.* (2001). However, it is well-known that in practice the excess returns are not normally distributed. Most financial assets exhibit excess kurtosis, that is, returns having distributions whose tails are heavier than those of the normal distribution, see Fama (1965), Blattberg and Gonedes (1974), Zhou (1993), Campbell *et al.* (1997) and Vorkink (2003).

In this paper we consider the Capital Asset Pricing Model under Elliptical (symmetric) Distributions. This class of distributions, which contains the normal distribution, $\boldsymbol{t}$, contaminated normal and power exponential, among others, has received greater interest in the literature, see Lange *et al.* (1989), Fang *et al.* (1990), Fang and Anderson (1990), Fang and Zhang (1990) and Gupta and Varga (1993). This class of elliptical distributions offers a more flexible framework for modelling asset prices or returns. It contains many distributions with heavier tails than the normal distribution allowing us to model tails which are frequently observed in financial data, especially in Latin American Markets. Results concerning the use of elliptical distributions in the $CAPM$ and in portfolio analysis can be found in Owen and Rabinovitch (1983), Ingersoll (1987), Zhou (1993), and recently in Hodgson *et al.* (2002). In particular, Owen and Rabinovitch (1983) and Ingersoll (1987) show the validity of $CAPM$ within the class of elliptical distributions.

The detection of atypical (outliers) and/or influential returns is an important stage in any econometric analysis of financial models. This is essential in order to evaluate the sensitivity (robustness) of the results obtained, using the set of data available, since the atypical returns can distort the estimators, leading to, in some cases, wrong decisions. Recently, van der Hart *et al.* (2003), showed that the outliers are one of the important factors in the selection of stocks in emerging markets.

There are various alternatives for evaluating the influence of perturbations in the data and /or in the model assumptions concerning the parameter estimators of our interest. See, for example, Cook and Weisberg (1982), Chatterjee and Hadi (1988), Cook (1986), and Barnett and Lewis (1994). The deletion of cases is a common diagnostic technique for evaluating the effect of an observation in the estimation process and in hypothesis testing. This is an analysis of global influence, since the effect of observation is quantified by eliminating it from the data set. Alternatively, Cook (1986) proposes an interesting method, called local influence, for evaluating the effect of small perturbations in the data and/or in the econometric model assumptions concerning the maximum likelihood estimators, without eliminating the observations. This method was applied by Galea *et al.* (1997, 2003), Díaz-Garcia *et al.* (2003) in elliptical linear models and by Cademartori *et al.* (2003) in the univariate $CAPM$ using the $\boldsymbol{t}$ distribution. Additional results about local influence and its applications can be found in Escobar and Meeker (1992), Zhao and Lee (1998), Galea *et al.* (2002), and Lesaffre and Verbeke (1998). However, no application of local influence has been considered for the $CAPM$ under elliptical distributions. Thus, the main objective of this paper is to apply the approach of local influence to the *Elliptical Capital Asset Pricing Model*, $ECAPM$. Several perturbation schemes are considered such as the case perturbation and market portfolio perturbation. With this we hope to expand the results in Cademartori *et al.* (2003).

The article is developed as follows. Section 2 briefly reviews $CAPM$ under elliptical distributions. Section 3 analyzes the local influence method to detect influential returns in the maximum likelihood estimators. In Section 4, model curvatures are considered for different perturbations schemes. In Section 5, the methodology is applied to a set of shares traded in the Santiago Stock Exchange Market. In Section 6 we conclude with some final comments.

2 The elliptical capital asset pricing model

In this section, we present some results and notations about the $CAPM$ under elliptical distributions. For this, we consider only the class of elliptical distributions with density.

We say that a $p \times 1$ random vector $\boldsymbol{Y} = (Y_1, ..., Y_p)^T$ has an elliptical distribution with parameters $\boldsymbol{\mu}$ (the location vector) and $\boldsymbol{\Sigma}$ (the scale matrix) of dimensions $p \times 1$ and $p \times p$, respectively, with $\boldsymbol{\Sigma} > 0$, if its density function is given by

$$f_{\boldsymbol{Y}}(\boldsymbol{y}) = |\boldsymbol{\Sigma}|^{-1/2} g[(\boldsymbol{y} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\boldsymbol{y} - \boldsymbol{\mu})], \quad \boldsymbol{y} \in \mathbb{R}^p, \tag{2.1}$$

where the function $g : \mathbb{R} \to [0, \infty)$ is such that $\int_0^\infty u^{p-1} g(u^2) du < \infty$. The function g is called the density generator and is written $\boldsymbol{Y} \sim El_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}; g)$ or simply $El_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. When the expectation and variance exist, we have that $E(\boldsymbol{Y}) = \boldsymbol{\mu}$ and $\text{Var}(\boldsymbol{Y}) = c_g \boldsymbol{\Sigma}$, where c_g is a positive constant. See, for example Fang *et al.* (1990). In the case where $\boldsymbol{\mu} = \boldsymbol{0}$ and $\boldsymbol{\Sigma} = \boldsymbol{I}_p$, we obtain the spherical family of densities. This class of distributions includes Normal ($g(u) = ce^{-u/2}$), $\boldsymbol{t}$ ($g(u) = c(\nu, p)(1 + u/\nu)^{-(\nu+p)/2}$, $\nu > 0$), Contaminated Normal ($g(u) = c\{(1-\gamma)e^{-u/2} + \frac{\gamma}{\sqrt{\phi}} e^{-u/2\phi}\}, \phi > 0, 0 \leq \gamma \leq 1$), Logistic ($g(u) = e^{-\sqrt{u}}/(1 + e^{-\sqrt{u}})^2$) and Power Exponential $g(u) = c(\lambda) e^{-u^\lambda/2}, \lambda > 0$), among other distributions.

The capital asset pricing model, $CAPM$, states that the share expected return is equal to the risk free rate return plus a prize for risk. This model was independently derived by Sharpe (1964), Lintner (1965), and Mossin (1966). Let R_i be a random variable, which denotes the return for asset i. According to the $CAPM$, the expected value of R_i is given by:

$$E[R_i] = R_f + \beta_i (E[R_m] - R_f), \tag{2.2}$$

where R_f is the risk free rate return, β_i is the systematic risk of the asset i, and R_m is the market return given by an index. Although criticisms exist about the $CAPM$ (Fama and French, 1992), and the APT has been developed as an alternative model, the $CAPM$ continues to be used in the administration of portfolios as well as in academic research (Elsas *et al.* 2003; Bartholdy and Peare, 2003).

Let $\boldsymbol{Y}_k = (Y_{1k}, ..., Y_{pk})^T$ be the vector of excess returns for p assets (or portfolios of assets). For these p assets, the excess returns can be described using the Excess-return Market Model (Elton and Gruber, 1995; Campbell *et al.*, 1997),

$$\mathbf{Y}_k = \boldsymbol{\alpha} + \boldsymbol{\beta} x_k + \boldsymbol{\epsilon}_k, \tag{2.3}$$

where x_k is the time period k market portfolio excess return, $\boldsymbol{\beta} = (\beta_1, ..., \beta_p)^T$ is the $p \times 1$ vector of betas, $\boldsymbol{\alpha} = (\alpha_1, ..., \alpha_p)^T$ is the $p \times 1$ vector of asset return intercepts, and $\boldsymbol{\epsilon}_k$ is the disturbances, $k = 1, ..., n$. The elliptical model is obtained considering

$$\boldsymbol{\epsilon}_k \sim El_p(\mathbf{0}, \boldsymbol{\Sigma}; g), \quad k = 1, ..., n, \tag{2.4}$$

where $\boldsymbol{\Sigma} > 0$ and the density function of $\boldsymbol{Y}_k$ is given by (2.1), with $\boldsymbol{\mu} = \boldsymbol{\mu}_k = \boldsymbol{\alpha} + \boldsymbol{\beta} x_k$. This model will be called the Elliptical Excess-return Market Model. If $g(u) = ce^{-u/2}$, $u \geq 0$, we have the normal model considered by Campbell *et al.* (1997). Recently, Hodgson *et al.* (2002) proposed a semi-parametric approach for testing the $CAPM$ efficienty based on elliptical distributions. Here we consider influence diagnostics in the $ECAPM$.

The log-likelihood function that corresponds to the model (2.3)-(2.4) is given by,

$$\ell(\boldsymbol{\theta}) = \sum_{k=1}^{n} l_k(\boldsymbol{\theta}), \tag{2.5}$$

where $l_k(\boldsymbol{\theta}) = -\dfrac{1}{2}\log|\boldsymbol{\Sigma}| + \log g(d_k)$, with $d_k = (\boldsymbol{Y}_k - \boldsymbol{\alpha} - \boldsymbol{\beta} x_k)^T \boldsymbol{\Sigma}^{-1} (\boldsymbol{Y}_k - \boldsymbol{\alpha} - \boldsymbol{\beta} x_k)$, $k = 1, ..., n$ and $\boldsymbol{\theta} = (\boldsymbol{\alpha}^T, \boldsymbol{\beta}^T, v(\boldsymbol{\Sigma})^T)^T$, where $v(\boldsymbol{\Sigma})$ is the $p(p+1)/2$ vector obtained from $vec(\boldsymbol{\Sigma})$ by deleting from it all of the elements that are above the diagonal of $\boldsymbol{\Sigma}$. If g is a continuous and decreasing function, then maximum likelihood estimators of $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, and $\boldsymbol{\Sigma}$ are obtained as solution to the equations,

$$\hat{\boldsymbol{\alpha}} = \overline{\boldsymbol{y}}_v - \hat{\boldsymbol{\beta}} \overline{x}_v, \qquad \hat{\boldsymbol{\beta}} = \frac{\sum_{k=1}^{n} v(d_k) x_k (\boldsymbol{y}_k - \overline{\boldsymbol{y}}_v)}{\sum_{k=1}^{n} v(d_k) x_k (x_k - \overline{x}_v)} \quad \text{and} \quad \hat{\boldsymbol{\Sigma}} = \frac{1}{n} \sum_{k=1}^{n} v(d_k) \boldsymbol{e}_k \boldsymbol{e}_k^T, \tag{2.6}$$

where $\boldsymbol{e}_k = \boldsymbol{y}_k - \hat{\boldsymbol{\alpha}} - \hat{\boldsymbol{\beta}} x_k$, $\overline{\boldsymbol{y}}_v = \sum_{k=1}^{n} v(d_k) \boldsymbol{y}_k / \sum_{k=1}^{n} v(d_k)$ *and* $\overline{x}_v = \sum_{k=1}^{n} v(d_k) x_k / \sum_{k=1}^{n} v(d_k)$, with $v(d_k) = -2W_g(d_k)$, $W_g(u) = g'(u)/g(u)$, $u \geq 0$ and d_k is as in (2.5).

Note that for the normal model $v(d_k) = 1$, $k = 1, ..., n$ and the maximum likelihood estimators above correspond to the normal case. Under the elliptical model, the exact marginal distribution of $\hat{\boldsymbol{\alpha}}$, $\hat{\boldsymbol{\beta}}$, and $\hat{\boldsymbol{\Sigma}}$ are particularly difficult to obtain, but under normal distribution the estimators of $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, and $\boldsymbol{\Sigma}$ have exact marginal distributions, see Campbell *et al.* (1997).

3 Local influence

Let $l(\boldsymbol{\theta})$ denote the log-likelihood function from the postulated model, where $\boldsymbol{\theta}$ is as in (2.5), and let $\boldsymbol{\omega}$ be an $q \times 1$ vector of perturbations restricted to some open subset of $\mathbb{R}^q$. The perturbations are made on the log-likelihood, such that it takes the form $l(\boldsymbol{\theta}/\boldsymbol{\omega})$. Let $\boldsymbol{\omega}_0$ the vector of no perturbation such that $l(\boldsymbol{\theta}) = l(\boldsymbol{\theta}/\boldsymbol{\omega}_0)$. The idea of the local influence method is to investigate how much the maximum likelihood estimates are affected by the corresponding perturbations. To asses influence of the perturbations on the maximum likelihood estimate $\hat{\boldsymbol{\theta}}$, we consider the likelihood displacement $LD(\boldsymbol{\omega}) = 2[l(\hat{\boldsymbol{\theta}}) - l(\hat{\boldsymbol{\theta}}_\omega)]$, where $\hat{\boldsymbol{\theta}}_\omega$ denotes the maximum likelihood estimates under the model $l(\boldsymbol{\theta} \mid \boldsymbol{\omega})$. The $LD(\boldsymbol{\omega})$ is useful in measuring the distance between $\hat{\boldsymbol{\theta}}$ and $\hat{\boldsymbol{\theta}}_\omega$. Cook (1986) proposes studying the local behavior of $LD(\boldsymbol{\omega})$ around $\boldsymbol{\omega}_0$. The procedure consists in selecting a unit direction $\boldsymbol{l}$ ($\| \boldsymbol{l} \| = 1$), and then to considering the plot of $LD(\boldsymbol{\omega}_0 + a\boldsymbol{l})$ against a, where $a \in \mathbb{R}$. This plot is called *lifted line*. Each lifted line can be characterized by considering the normal curvature $C_{\boldsymbol{l}}(\boldsymbol{\theta})$ around $a = 0$. The suggestion is to consider the directions $\boldsymbol{l}_{max}$ corresponding to the largest curvature $C_{\boldsymbol{l}_{max}}(\boldsymbol{\theta})$. The index plot of the $\boldsymbol{l}_{max}$ 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

$$C_{\boldsymbol{l}}(\boldsymbol{\theta}) = 2 \mid \boldsymbol{l}^T \boldsymbol{\Delta}^T \ddot{\boldsymbol{L}}^{-1} \boldsymbol{\Delta} \boldsymbol{l} \mid, \tag{3.1}$$

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 $\Delta_{ij} = \partial^2 l(\boldsymbol{\theta}/\boldsymbol{\omega})\partial\theta_i\partial\omega_j$, evaluated at $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}$ and $\boldsymbol{\omega} = {}_0$, $i = 1, ..., p^*$; $j = 1, ..., q$, with $p^* = p(p+5)/2$. Therefore, the maximization of (3.1) is equivalent to finding the largest absolute eigenvalue of the matrix $\boldsymbol{B} = \boldsymbol{\Delta}^T \ddot{\boldsymbol{L}}^{-1} \boldsymbol{\Delta}$, and $\boldsymbol{l}_{max}$ is the corresponding eigenvector. When a subset $\boldsymbol{\theta}_1$ of $\boldsymbol{\theta} = (\boldsymbol{\theta}_1^T, \boldsymbol{\theta}_2^T)^T$, is of special interest the influence diagnostics can be based on, see Cook (1986), $\boldsymbol{B} = \boldsymbol{\Delta}^T(\ddot{\boldsymbol{L}}^{-1} - \boldsymbol{B}_{22})\boldsymbol{\Delta}$ where, $\boldsymbol{B}_{22} = Diag(\boldsymbol{0}, \ddot{\boldsymbol{L}}_{22}^{-1})$, and $\ddot{\boldsymbol{L}}_{22}$ is obtained from the partition of $\ddot{\boldsymbol{L}}$ according to the partition of $\boldsymbol{\theta}$.

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

4 Curvature derivative for ECAPM

In this section we present the observed information matrix and the $\boldsymbol{\Delta}$ matrix for different perturbation approaches.

4.1 The observed information matrix

The observed information matrix is given by $-\ddot{\boldsymbol{L}}$, where $\ddot{\boldsymbol{L}} = [(L_{\boldsymbol{\gamma\tau}})]$, $\boldsymbol{\gamma}, \boldsymbol{\tau} = \boldsymbol{\alpha}, \boldsymbol{\beta}, v(\boldsymbol{\Sigma})$ with

$$L_{\boldsymbol{\gamma\tau}} = \sum_{k=1}^{n} \frac{\partial^2 l_k(\boldsymbol{\theta})}{\partial \boldsymbol{\gamma} \partial \boldsymbol{\tau}^T}, \tag{4.1}$$

$l_k(\boldsymbol{\theta})$ is given by (2.5) and, see Appendix,

$$\frac{\partial^2 l_k(\boldsymbol{\theta})}{\partial \boldsymbol{\alpha} \partial \boldsymbol{\alpha}^T} = 2W_g(d_k)\boldsymbol{\Sigma}^{-1} + 4W_g'(d_k)\boldsymbol{\Sigma}^{-1}\boldsymbol{\epsilon}_k\boldsymbol{\epsilon}_k^T\boldsymbol{\Sigma}^{-1}, \qquad \frac{\partial^2 l_k(\boldsymbol{\theta})}{\partial \boldsymbol{\alpha} \partial \boldsymbol{\beta}^T} = 2x_k \frac{\partial^2 l_k(\boldsymbol{\theta})}{\partial \boldsymbol{\alpha} \partial \boldsymbol{\alpha}^T}, \tag{4.2}$$

$$\frac{\partial^2 l_k(\boldsymbol{\theta})}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}^T} = x_k^2 \frac{\partial^2 l_k(\boldsymbol{\theta})}{\partial \boldsymbol{\alpha} \partial \boldsymbol{\alpha}^T}, \qquad \frac{\partial^2 l_k(\boldsymbol{\theta})}{\partial v(\boldsymbol{\Sigma}) \partial \boldsymbol{\beta}^T} = x_k \frac{\partial^2 l_k(\boldsymbol{\theta})}{\partial v(\boldsymbol{\Sigma}) \partial \boldsymbol{\alpha}^T}, \tag{4.3}$$

$$\frac{\partial^2 l_k(\boldsymbol{\theta})}{\partial v(\boldsymbol{\Sigma}) \partial \boldsymbol{\alpha}^T} = 4\boldsymbol{D}_p^T(W_g(d_k)(\boldsymbol{\Sigma}^{-1} \otimes \boldsymbol{\Sigma}^{-1}\boldsymbol{\epsilon}_k) + W_g'(d_k)vec(\boldsymbol{\Sigma}^{-1}\boldsymbol{\epsilon}_k\boldsymbol{\epsilon}_k^T\boldsymbol{\Sigma}^{-1})\boldsymbol{\epsilon}_k^T\boldsymbol{\Sigma}^{-1}), \tag{4.4}$$

$$\begin{aligned}\frac{\partial^2 l_k(\boldsymbol{\theta})}{\partial v(\boldsymbol{\Sigma}) \partial v^T(\boldsymbol{\Sigma})} &= \boldsymbol{D}_p^T[\frac{1}{2}(\boldsymbol{\Sigma}^{-1} \otimes \boldsymbol{\Sigma}^{-1}) + 2W_g(d_k)(\boldsymbol{\Sigma}^{-1}\boldsymbol{\epsilon}_k\boldsymbol{\epsilon}_k^T\boldsymbol{\Sigma}^{-1} \otimes \boldsymbol{\Sigma}^{-1}) \\ &\quad + W_g'(d_k)vec(\boldsymbol{\Sigma}^{-1}\boldsymbol{\epsilon}_k\boldsymbol{\epsilon}_k^T\boldsymbol{\Sigma}^{-1})vec^T(\boldsymbol{\Sigma}^{-1}\boldsymbol{\epsilon}_k\boldsymbol{\epsilon}_k^T\boldsymbol{\Sigma}^{-1})]\boldsymbol{D}_p,\end{aligned} \tag{4.5}$$

where $W_g(u) = g'(u)/g(u)$, $W_g'(u) = dW_g(u)/du$ and $\boldsymbol{D}_p$ the duplication matrix, see Magnus and Neudecker (1988). All of these expressions should be evaluated in $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}$.

4.2 Perturbation of case weights

Consider the weight vector $\boldsymbol{\omega} = (\omega_1, ..., \omega_n)^T$ so that the perturbed log-likelihood is denoted by

$$l(\boldsymbol{\theta}/\boldsymbol{\omega}) = \sum_{k=1}^{n} \omega_k l_k(\boldsymbol{\theta}), \tag{4.6}$$

where $l_k(\boldsymbol{\theta})$ is given by (2.5). The vector of no-perturbation is denoted by $\boldsymbol{\omega}_0 = \mathbf{1}_n = (1, ..., 1)^T$. In this case $\boldsymbol{\Delta} = (\boldsymbol{\Delta}_1(\boldsymbol{\theta}), ..., \boldsymbol{\Delta}_n(\boldsymbol{\theta}))$ is the $p^* \times n$ matrix, where $\boldsymbol{\Delta}_k(\boldsymbol{\theta}) = \partial l_k(\boldsymbol{\theta})/\partial \boldsymbol{\theta} \mid_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}}$, $k = 1, ..., n$, with elements $\partial l_k(\boldsymbol{\theta})/\partial \boldsymbol{\gamma}$, $\boldsymbol{\gamma} = \boldsymbol{\alpha}, \boldsymbol{\beta}, v(\boldsymbol{\Sigma})$ that are given by

$$\frac{\partial l_k(\boldsymbol{\theta})}{\partial \boldsymbol{\alpha}} = -2W_g(d_k)\boldsymbol{\Sigma}^{-1}\boldsymbol{\epsilon}_k, \qquad \frac{\partial l_k(\boldsymbol{\theta})}{\partial \boldsymbol{\beta}} = x_k \frac{\partial l_k(\boldsymbol{\theta})}{\partial \boldsymbol{\alpha}} \tag{4.7}$$

and

$$\frac{\partial l_k(\boldsymbol{\theta})}{\partial v(\boldsymbol{\Sigma})} = \boldsymbol{D}_p^T vec[-\frac{1}{2}\boldsymbol{\Sigma}^{-1} - W_g(d_k)\boldsymbol{\Sigma}^{-1}\boldsymbol{\epsilon}_k\boldsymbol{\epsilon}_k^T\boldsymbol{\Sigma}^{-1}]. \tag{4.8}$$

4.3 Perturbation Market Return

In this section, the market returns x_k, are perturbed considering additive perturbation schemes, thus $x_{k\omega} = x_k + \omega_k$, $k = 1, ..., n$ and the perturbed log-likelihood is constructed with x_k, which in the following function is replaced by $x_{k\omega}$, that is

$$l(\boldsymbol{\theta}/\boldsymbol{\omega}) = \sum_{k=1}^{n} l_k(\boldsymbol{\theta}/\omega_k), \tag{4.9}$$

where $l_k(\boldsymbol{\theta}/\omega_k) = -\frac{1}{2}\log|\boldsymbol{\Sigma}| + \log g(d_{k\omega})$, with $d_{k\omega} = (\boldsymbol{Y}_k - \boldsymbol{\alpha} - \boldsymbol{\beta}x_{k\omega})^T\boldsymbol{\Sigma}^{-1}(\boldsymbol{Y}_k - \boldsymbol{\alpha} - \boldsymbol{\beta}x_{k\omega})$, $k = 1, ..., n$ and $\boldsymbol{\omega} = (\omega_1, ..., \omega_n)^T$. The no perturbation case follows by taking $\boldsymbol{\omega} = (0, ..., 0)^T$. In this case $\boldsymbol{\Delta} = (\boldsymbol{\Delta}_1(\boldsymbol{\theta}), ..., \boldsymbol{\Delta}_n(\boldsymbol{\theta}))$, evaluated in $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}$, where the elements of the $\boldsymbol{\Delta}_k(\boldsymbol{\theta})$, $k = 1, .., n$ are $\boldsymbol{\Delta}_{\boldsymbol{\gamma}k}$, $\boldsymbol{\gamma} = \boldsymbol{\alpha}, \boldsymbol{\beta}, v(\boldsymbol{\Sigma})$, that after some algebraic manipulation, it follows that

$$\boldsymbol{\Delta}_{\boldsymbol{\alpha}k} = -2W_g(d_k)\boldsymbol{\Sigma}^{-1}\boldsymbol{\beta} + 4W_g'(d_k)\boldsymbol{\epsilon}_k^T\boldsymbol{\Sigma}^{-1}\boldsymbol{\beta}\boldsymbol{\Sigma}^{-1}\boldsymbol{\epsilon}_k, \tag{4.10}$$

$$\boldsymbol{\Delta}_{\boldsymbol{\beta}k} = -2W_g(d_k)\boldsymbol{\Sigma}^{-1}\boldsymbol{\epsilon}_k + x_k\boldsymbol{\Delta}_{\boldsymbol{\alpha}k}, \tag{4.11}$$

$$\boldsymbol{\Delta}_{v(\boldsymbol{\Sigma})k} = \boldsymbol{D}_p^T\text{vec}[W_g(d_k)\boldsymbol{\Sigma}^{-1}(\boldsymbol{\epsilon}_k\boldsymbol{\beta}^T + \boldsymbol{\beta}\boldsymbol{\epsilon}_k^T)\boldsymbol{\Sigma}^{-1} + 2W_g'(d_k)\boldsymbol{\epsilon}_k^T\boldsymbol{\Sigma}^{-1}\boldsymbol{\beta}\boldsymbol{\Sigma}^{-1}\boldsymbol{\epsilon}_k\boldsymbol{\epsilon}_k^T\boldsymbol{\Sigma}^{-1}]. \tag{4.12}$$

4.4 Perturbation of the scale matrix

The model in (2.3) is assumed to be homochedastic, that is, the scale matrix of random errors is assumed to be the same. In this section, we assume that the scale matrix of the random errors is given by $\boldsymbol{\Sigma}/\omega_k$, $k = 1, ..., n$. Thus, the log-likelihood corresponding to the perturbed model is given by

$$l(\boldsymbol{\theta}/\boldsymbol{\omega}) = \sum_{k=1}^{n} l_k(\boldsymbol{\theta}/\omega_k), \tag{4.13}$$

where $l_k(\boldsymbol{\theta}/\omega_k) = \frac{p}{2}\log\omega_k - \frac{1}{2}\log|\boldsymbol{\Sigma}| + \log g(\omega_k d_k)$. The no-perturbation case follows by taking $\boldsymbol{\omega} = (1, ..., 1)^T$. Thus $\boldsymbol{\Delta} = (\boldsymbol{\Delta}_1(\boldsymbol{\theta}), ..., \boldsymbol{\Delta}_n(\boldsymbol{\theta}))$, evaluated in $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}$, is the $p^* \times n$ matrix, where $\boldsymbol{\Delta}_k(\boldsymbol{\theta}) = [W_g(\omega_k d_k) + \omega_k d_k W_g'(\omega_k d_k)]\partial d_k/\partial\boldsymbol{\theta}$, where the elements of $\partial d_k/\partial\boldsymbol{\theta}$, $\partial d_k/\partial\boldsymbol{\gamma}$, $\boldsymbol{\gamma} = \boldsymbol{\alpha}, \boldsymbol{\beta}, v(\boldsymbol{\Sigma})$, $k = 1, ..., n$ are given by,

$$\frac{\partial d_k}{\partial\boldsymbol{\alpha}} = -2\boldsymbol{\Sigma}^{-1}\boldsymbol{\epsilon}_k, \qquad \frac{\partial d_k}{\partial\boldsymbol{\beta}} = x_k\frac{\partial d_k}{\partial\boldsymbol{\alpha}} \quad \text{and} \quad \frac{\partial d_k}{\partial v(\boldsymbol{\Sigma})} = -\boldsymbol{D}_p^T\text{vec}(\boldsymbol{\Sigma}^{-1}\boldsymbol{\epsilon}_k\boldsymbol{\epsilon}_k^T\boldsymbol{\Sigma}^{-1}). \tag{4.14}$$

5 Application

The data corresponds to monthly returns of shares from the Chilean Stock Market. The Selective Index of Share Prices, $IPSA$, was used as the return for the market and the interest rate in the sale of discount bonus of the Central Bank was used as the risk free rate, both monthly. The data correspond to the period, January 1990 to June 2004. We illustrate the methodology with four companies, *Concha y Toro* (CyT) of the wine industry, *Copec*, a big company with activities in different areas of the economy, *Entel*, a telecommunication company, and *Cuprum*, which manages pension funds. Table 1 presents a summary of the adjusted results of the four selected models: Normal, $\boldsymbol{t}$, Power exponential, and Contaminated Normal. Of the four, the one which shows the best fit (see column $l(\hat{\boldsymbol{\theta}})$) is the $\boldsymbol{t}$ model with four degrees of freedom, $\boldsymbol{t}_4$. For example, the maximum log-likelihood for the normal model is 621.14 and for the $\boldsymbol{t}_4$ model the maximum log-likelihood is 697.14, corresponding to likelihood ratio statistic of 152. This indicates that the $\boldsymbol{t}_4$ model fits the data significantly better than the normal model. The greatest variation in systematic estimated risks, $\hat{\beta}_j, j = 1, 2, 3, 4$, are seen in CyT, $\hat{\beta}_1$, while the most stable are those from $COPEC$, $\hat{\beta}_2$. In addition, important variations in the estimated covariance matrix, $\hat{\boldsymbol{\Sigma}}$ are observed. Figure 1 shows the dispersion diagram of the four companies. It is possible to observe in the four graphics some atypical returns that could have an influence on the maximum likelihood estimators. In these graphics the returns 14, 37, and 104 were marked since they were detected as potentially influential by the local influence method. Figures 2 and 4 present the index graphs of $\boldsymbol{l}_{max}$ for the three perturbation schemes considered in the four models selected.

*********Figure 1 about here*********

Tabla 1: Adjusted results for the 4 models selected

Distribution	Asset	$\hat{\boldsymbol{\alpha}}$	$\hat{\boldsymbol{\beta}}$	$\hat{\boldsymbol{\Sigma}}$	$l(\hat{\boldsymbol{\theta}})$
Normal	*CyT*	0.0105	0.9420	0.0122 -0.0005 -0.0006 0.0034	621.14
	Copec	0.0058	0.8857	0.0044 0.0006 0.0005	
	Entel	-0.0001	1.1090	0.0116 0.0015	
	Cuprum	0.0299	0.9686	0.0164	
$\boldsymbol{t}$ $(\nu = 4)$	*CyT*	0.0034	0.6857	0.004679 -0.000033 -0.000718 0.000479	697.14
	Copec	0.0050	0.8449	0.002788 0.000353 0.000086	
	Entel	-0.0021	1.1149	0.006998 0.000362	
	Cuprum	0.0134	0.7956	0.006429	
Power Exp.$(\lambda = 2/3)$	*CyT*	0.0060	0.7919	0.002291 -0.000048 -0.000220 0.000439	667.91
	Copec	0.0048	0.8745	0.001061 0.000143 0.000077	
	Entel	-0.0020	1.1011	0.002739 0.000249	
	Cuprum	0.0197	0.8713	0.003240	
CN$(\gamma = 0.10, \phi = 10)$	*CyT*	0.0011	0.6917	0.0053 0.0001 -0.0006 0.0008	687.72
	Copec	0.0047	0.8427	0.0033 0.0005 0.0000	
	Entel	-0.0014	1.0473	0.0080 0.0007	
	Cuprum	0.0156	0.8895	0.0076	

As can be observed from these figures, case 14 seems to be the most influential in the maximum likelihood estimators in the normal and power exponential models under the case and scale perturbation schemes. Return 104 seems to be the most influential in $\hat{\boldsymbol{\theta}}$, when using the $\boldsymbol{t}_4$ model and the case perturbation scheme, while the other perturbation schemes do not influence the maximum likelihood estimators, as can be observed in figures 3(b) and 4(b). In addition, note that the maximum likelihood estimators are quite stable as they relate to the market return perturbation scheme, figure 3, in the four models considered.

*********Figures 2-4 about here*********

In figures 5 and 7, the graphics of the total local influence for the three perturbation schemes in each one of the four models considered are presented. Note that returns 14(in the three perturbation schemes) and 37(in the market perturbation scheme) are influential in the maximum likelihood estimators, when using the normal and contaminated normal distributions, respectively. In the other two models, however, influential returns are not observed.

*********Figures 5-7 about here*********

6 Conclusions

The objective of influence diagnostics is to identify anomalous observations that may affect the adequacy of fit and/or statistical inferences under the proposed model. This is essential for evaluating the sensitivity (robustness) of the results obtained, with the data set available, since atypical returns can distort the estimators, leading to, in some cases, wrong decisions. In a recent study, van der Hart *et al.* (2003), showed that the atypical returns (outliers) are one of the important factors in the selection of stocks in emerging markets.

The class of elliptical distributions offers a more flexible framework for modelling asset returns. It contains many distributions with heavier tails than the normal distribution allowing us to model tails which are frequently observed in financial data, especially in Latin American Markets. These distributions have shape parameters that can be used for adjusting the distribution kurtosis and providing more robust procedures than the ones that use the normal distribution, with moderate additional computational effort. Closed form expressions are obtained for the observed information matrix and for the $\mathbf{\Delta}$ matrix for three perturbation schemes in the $ECAPM$. Perturbation of case weights, this scheme is intended to evaluate whether the contribution of the returns with different weights affects the maximum likelihood estimator (MLE) of $\boldsymbol{\theta}$. Perhaps, this is the method most commonly used to evaluate the influence of a small modification of the model. The perturbation scheme of market returns can be used for analyzing the sensitivity of the MLE when the rate of return has suffered small changes. Finally, the perturbation scheme of the matrix scale can be used for analyzing the sensitivity of the MLE, with respect to the departures from homochedastic assumptions. This empirical study provides new evidence on the robustness aspects of the MLE for the $\boldsymbol{t}$ distribution with small degrees of freedom, as also shown by Lange *et al.* (1989) in regression and multivariate analysis. Nevertheless, it also shows the need to use diagnostic techniques in models whose tails are heavier than those of the normal distribution.

Appendix: The observed information matrix

For the elliptical model the log-likelihood function is given by,

$$l(\boldsymbol{\theta}) = \sum_{k=1}^{n} l_k(\boldsymbol{\theta}), \qquad \text{(A.1)}$$

where $l_k(\boldsymbol{\theta}) = -\frac{1}{2}\log|\boldsymbol{\Sigma}| + \log g(d_k)$, with $d_k = (\boldsymbol{Y}_k - \boldsymbol{\alpha} - \boldsymbol{\beta} x_k)^T \boldsymbol{\Sigma}^{-1} (\boldsymbol{Y}_k - \boldsymbol{\alpha} - \boldsymbol{\beta} x_k),\ k = 1, ..., n.$ Then, using results of differentiation of matrix, see Magnus and Neudecker (1988), we have

$$\begin{aligned}
d^2 l_k(\boldsymbol{\theta}) \;=\;& dv^T(\boldsymbol{\Sigma})\boldsymbol{D}_p^T[\frac{1}{2}(\boldsymbol{\Sigma}^{-1} \otimes \boldsymbol{\Sigma}^{-1}) + 2W_g(d_k)(\boldsymbol{\Sigma}^{-1}\boldsymbol{\epsilon}_k\boldsymbol{\epsilon}_k^T\boldsymbol{\Sigma}^{-1} \otimes \boldsymbol{\Sigma}^{-1}) \\
&+ W_g'(d_k) vec(\boldsymbol{\Sigma}^{-1}\boldsymbol{\epsilon}_k\boldsymbol{\epsilon}_k^T\boldsymbol{\Sigma}^{-1}) vec^T(\boldsymbol{\Sigma}^{-1}\boldsymbol{\epsilon}_k\boldsymbol{\epsilon}_k^T\boldsymbol{\Sigma}^{-1})]\boldsymbol{D}_p dv(\boldsymbol{\Sigma}) \\
&+ dv^T(\boldsymbol{\Sigma})\boldsymbol{D}_p^T[4W_g(d_k)(\boldsymbol{\epsilon}_k^T\boldsymbol{\Sigma}^{-1} \otimes \boldsymbol{\Sigma}^{-1}) + 4W_g'(d_k) vec(\boldsymbol{\Sigma}^{-1}\boldsymbol{\epsilon}_k\boldsymbol{\epsilon}_k^T\boldsymbol{\Sigma}^{-1})\boldsymbol{\epsilon}_k^T\boldsymbol{\Sigma}^{-1}]d\boldsymbol{\alpha} \\
&+ dv^T(\boldsymbol{\Sigma})\boldsymbol{D}_p^T[4x_k W_g(d_k)(\boldsymbol{\epsilon}_k^T\boldsymbol{\Sigma}^{-1} \otimes \boldsymbol{\Sigma}^{-1}) \\
&+ 4x_k W_g'(d_k) vec(\boldsymbol{\Sigma}^{-1}\boldsymbol{\epsilon}_k\boldsymbol{\epsilon}_k^T\boldsymbol{\Sigma}^{-1})\boldsymbol{\epsilon}_k^T\boldsymbol{\Sigma}^{-1}]d\boldsymbol{\beta} \\
&+ d\boldsymbol{\alpha}^T[2W_g(d_k)\boldsymbol{\Sigma}^{-1} + 4W_g'(d_k)\boldsymbol{\Sigma}^{-1}\boldsymbol{\epsilon}_k\boldsymbol{\epsilon}_k^T\boldsymbol{\Sigma}^{-1}]d\boldsymbol{\alpha} \\
&+ d\boldsymbol{\alpha}^T[4x_k W_g(d_k)\boldsymbol{\Sigma}^{-1} + 8x_k W_g'(d_k)\boldsymbol{\Sigma}^{-1}\boldsymbol{\epsilon}_k\boldsymbol{\epsilon}_k^T\boldsymbol{\Sigma}^{-1}]d\boldsymbol{\beta} \\
&+ d\boldsymbol{\beta}^T[2x_k^2 W_g(d_k)\boldsymbol{\Sigma}^{-1} + 4x_k^2 W_g'(d_k)\boldsymbol{\Sigma}^{-1}\boldsymbol{\epsilon}_k\boldsymbol{\epsilon}_k^T\boldsymbol{\Sigma}^{-1}]d\boldsymbol{\beta},
\end{aligned}$$

with $W_g(d_k)$ and $W_g'(d_k)$ as defined in the Section 4.1; from where the expressions (4.2) to (4.5) are obtained.

Acknowledgements

The authors acknowledges the partial financial support from Fapesp, 03/02747-4, Brasil and Project Dipuv 17/2003, Universidad de Valparaíso, Chile.

References

Barnett, V. and Lewis, T. (1994). *Outliers in statistical data.* John Wiley. New York.

Bartholdy, J. and Peare, P. (2003). Unbiased estimation of expected return using. *International Review of Financial Analysis, 12, 69-81.*

Blattberg, R. C. and Gonedes, N. J. (1974). A comparison of the stable and student distribution as statistical models for stock prices. *Journal of Business, 47, 244-280.*

Broquet, C. and van den Berg, A. (1992). *Gestion de Portefeuille: Actions, obligations, options.* 2^e Edition, De Boeck Université. Bruxelles.

Cademartori, D., Romo, C., Campos, R. and Galea, M. (2003). Robust estimation of systematic risk using the distribution in the Chilean Stock Markets. *Applied Economics Letters, 10, 447-453.*

Campbell, J., Lo, A. and MacKinlay, A. (1997). *Econometrics of Financial Markets.* Princeton University Press. New Jersey.

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

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.

Díaz-García, J., Galea, M. and Leiva-Sánchez, V. (2003). Influence diagnostics for elliptical multivariate regression models. *Communications in Statistics - Theory and Methods, 32, 625-641.*

Elsas, R., El-Shaer, M. and Theissen, E. (2003). Beta and returns revisited Evidence from the German stock market. *Journal of International Financial Markets, Institutions and Money, 13, 1-18.*

Elton, E. and Gruber, M. (1995). *Modern Portfolio Theory and Investment Analysis.* John Wiley. New York.

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

Fama, E. (1965). The behavior of stock market prices. *Journal of Business, 38, 34-105.*

Fama E. and French, K. (1992). The cross-section of expected stock returns. *Journal of Finance, 47, 427-465.*

Fang, K. T. and Anderson, T. W., eds. (1990). *Statistical inference in elliptical contoured and related distributions.* Allerton Press. New York.

Fang, K. T., Kotz, S. and Ng, K. W. (1990). *Symmetric multivariate and related distributions.* Chapman and Hall. London.

Fang, K. T. and Zhang, Y. T. (1990). *Generalized multivariate analysis.* Springer. London.

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. (2002). Local influence in comparative calibration models. *Biometrical Journal, 44, 59-81.*

Galea, M., Paula, G. A. and Uribe, M. (2003). Influence diagnostics in univariate linear elliptical regression model. *Statistical Paper, 44, 23-45.*

Gupta, A. K. and Varga, T. (1993). *Elliptically contoured models in statistics.* Kluwer Academic publishers.

Hodgson, D., Linton. O. and Vorkink, K. (2002). Testing the Capital Asset Pricing Model Efficienty Under Elliptical Symmetry: A Semiparametric Approach. *Journal of Applied Econometrics, 17, 617-639.*

Ingersoll, J. (1987). *Theory of financial decision making.* Rowman and Littlefield. Maryland.

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.*

Lesaffre, E. and Verbeke, G. (1998). Local influence in linear mixed models. *Biometrics, 54, 570-582.*

Lintner, J. (1965). The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *Review of Economics and Statistics, 41, 13-37.*

Magnus, J. R. and Neudecker, H. (1988). *Matrix Differential Calculus with Applications in Statistics and Econometrics.* John Wiley. New York.

Mossin, J. (1966). Equilibrium in capital asset market. *Econometrica, 35, 768-783.*

Owen, J. and Rabinovitch, R. (1983). On the class of elliptical distributions and their applications to the theory of portfolio. *The Journal of Finance, 38, 745-752.*

Ross, S., Westerfield, R. and Jordan, Br. (2001). *Fundamentos de finanzas corporativas.* McGraw Hill. México.

Sharpe, W. (1964). Capital asset prices : A theory of markets equilibrium under conditions of risk. *Journal of Finance, 19, 425-442.*

van der Hart, J., Slagter, E. and van Dijk, D. (2003). Stock selection strategies in emerging markets. *Journal of Empirical Finance, 10, 105-132.*

van Horne, J. (1997). *Administración Financiera.* Prentice Hall. México.

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

Vorkink, K. (2003). Return distributions and improved tests of asset pricing models. *The Review of Financial Studies, 16, 845-874.*

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

Zhou, G. (1993). Asset-pricing test under alternative distributions. *The Journal of Finance, 48, 1927-1942.*

Figura 1: Scatter plots

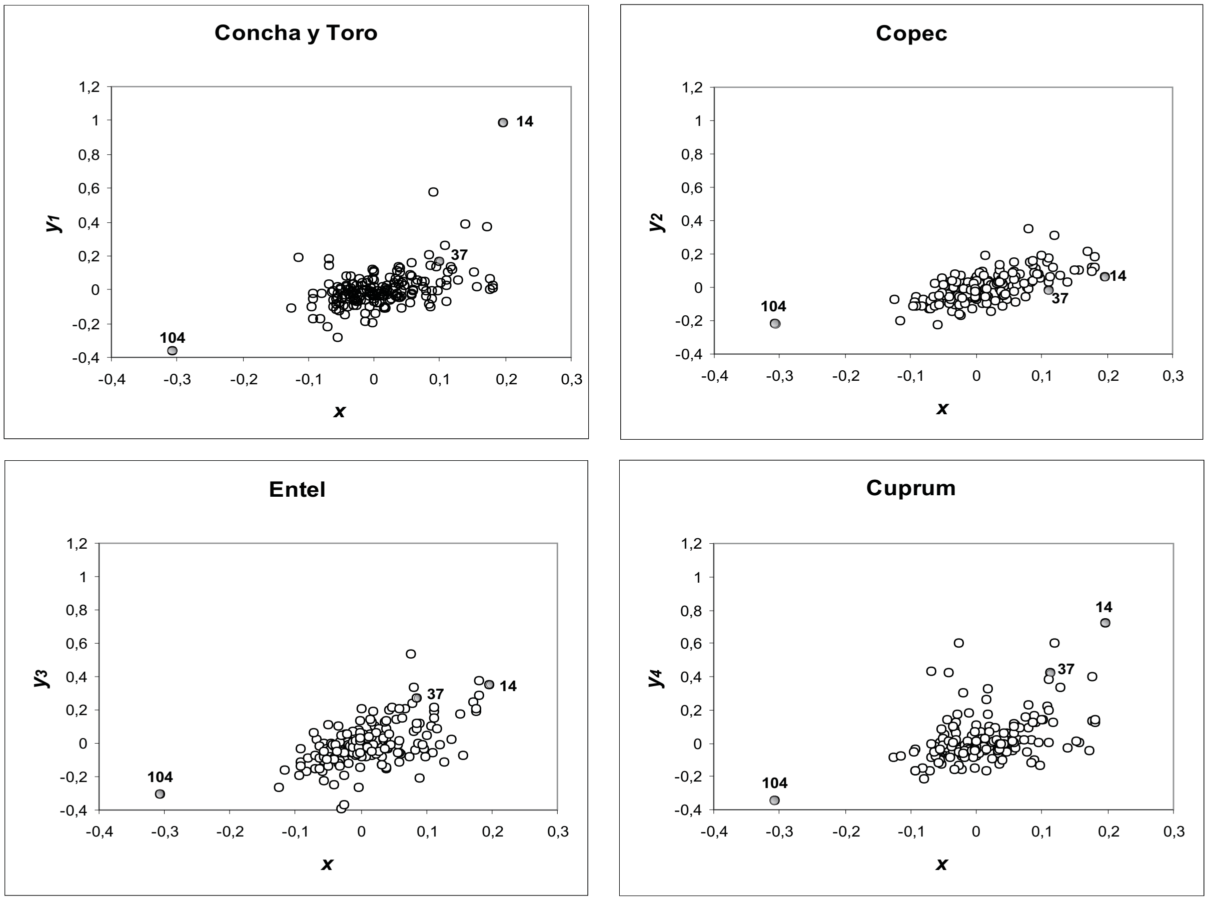

Figura 2: Index plots of $|l_{max}|$ for case perturbations in the (a) Normal (b) $\boldsymbol{t}_4$ (c) Power Exponential and (d) Contaminated Normal models

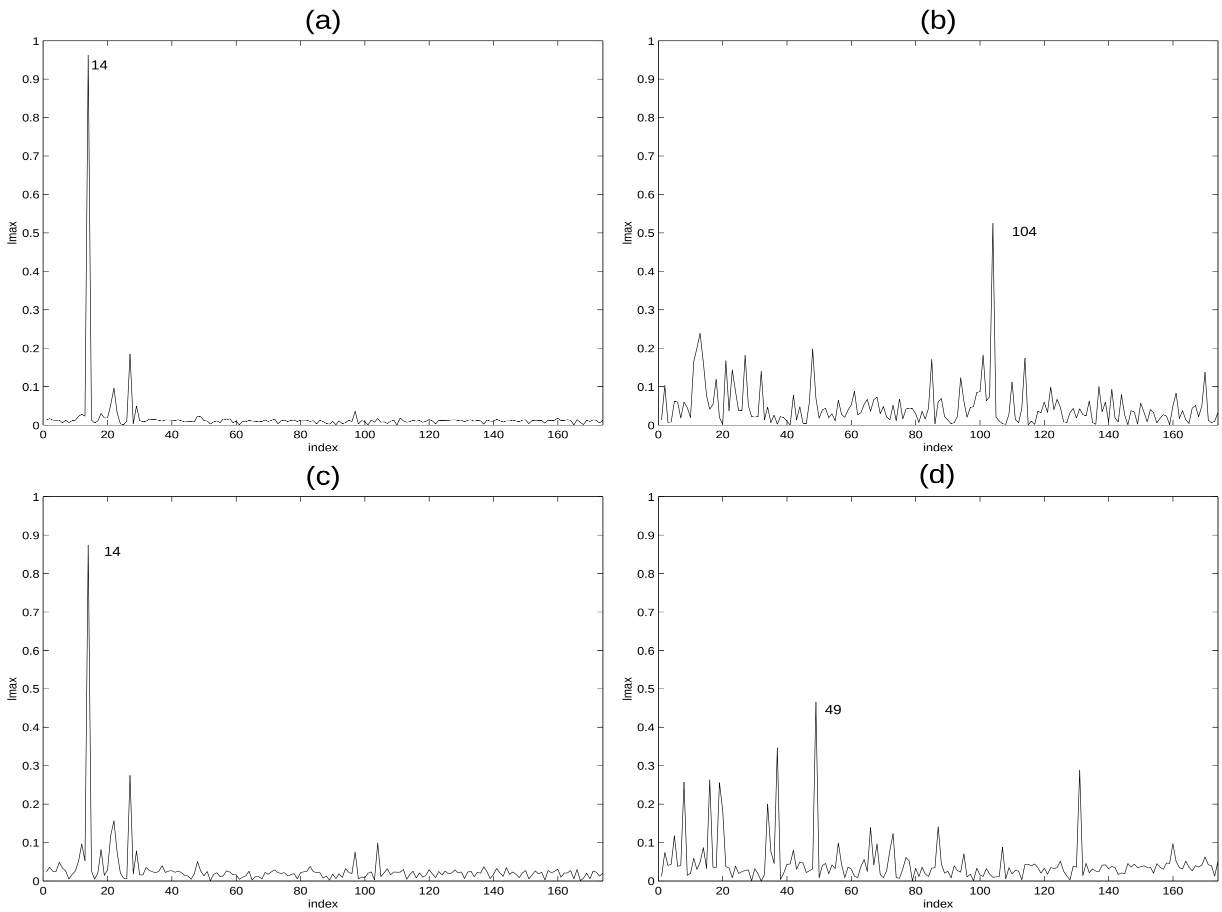

Figura 3: Index plots of $|l_{max}|$ for market perturbations in the (a) Normal (b) $\boldsymbol{t}_4$ (c) Power Exponential and (d) Contaminated Normal models

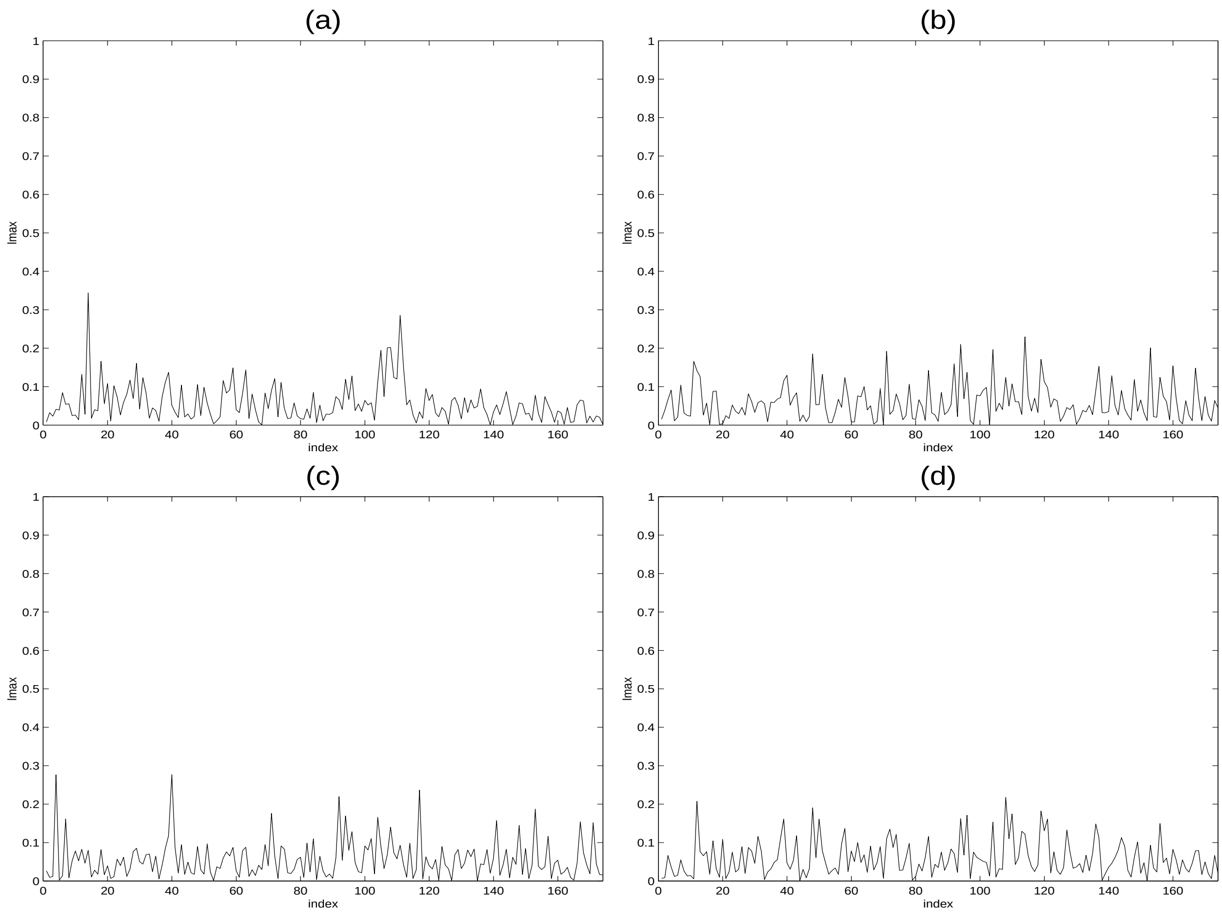

Figura 4: Index plots of $|l_{max}|$ for scale perturbations in the (a) Normal (b) $\boldsymbol{t}_4$ (c) Power Exponential and (d) Contaminated Normal models

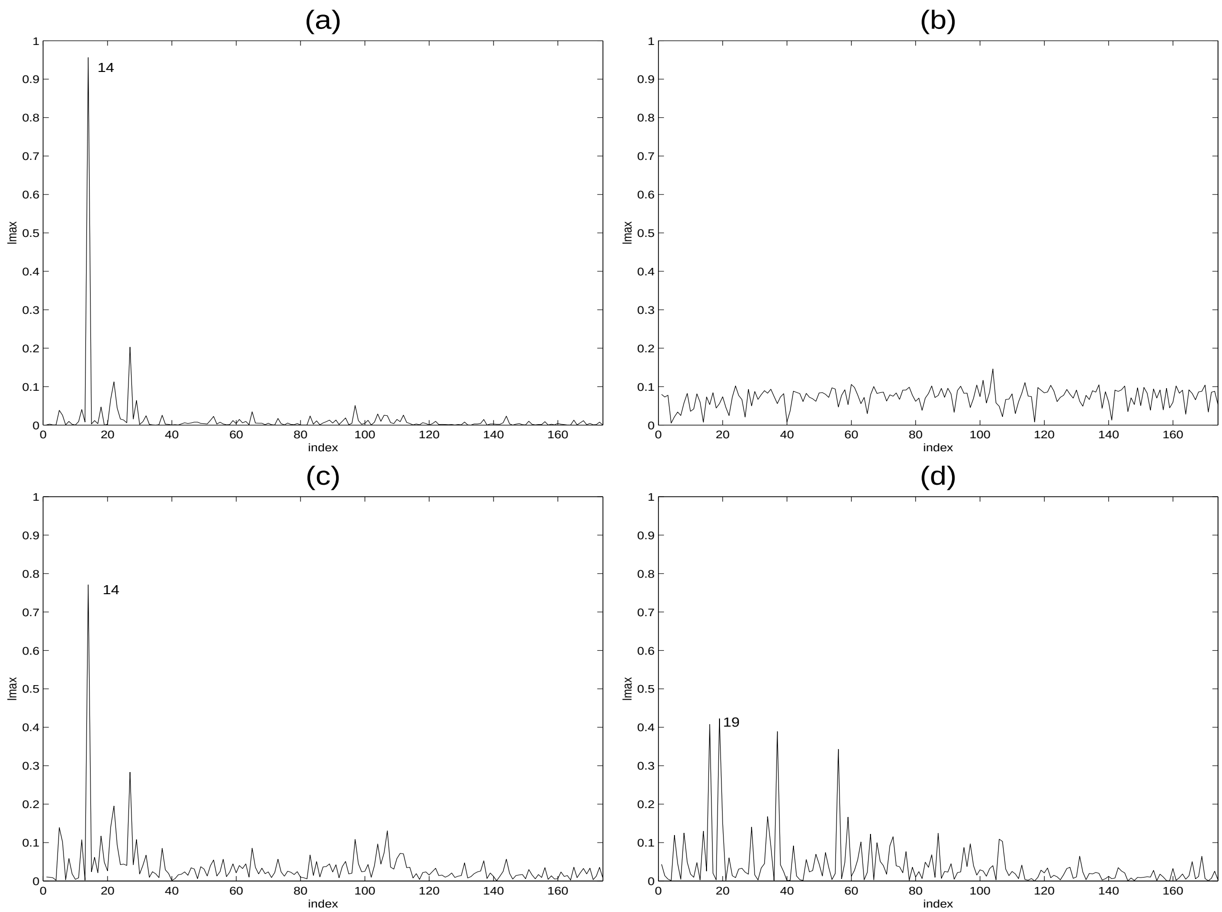

Figura 5: Index plots of C_k for case perturbations in the (a) Normal (b) $\boldsymbol{t}_4$ (c) Power Exponential and (d) Contaminated Normal models

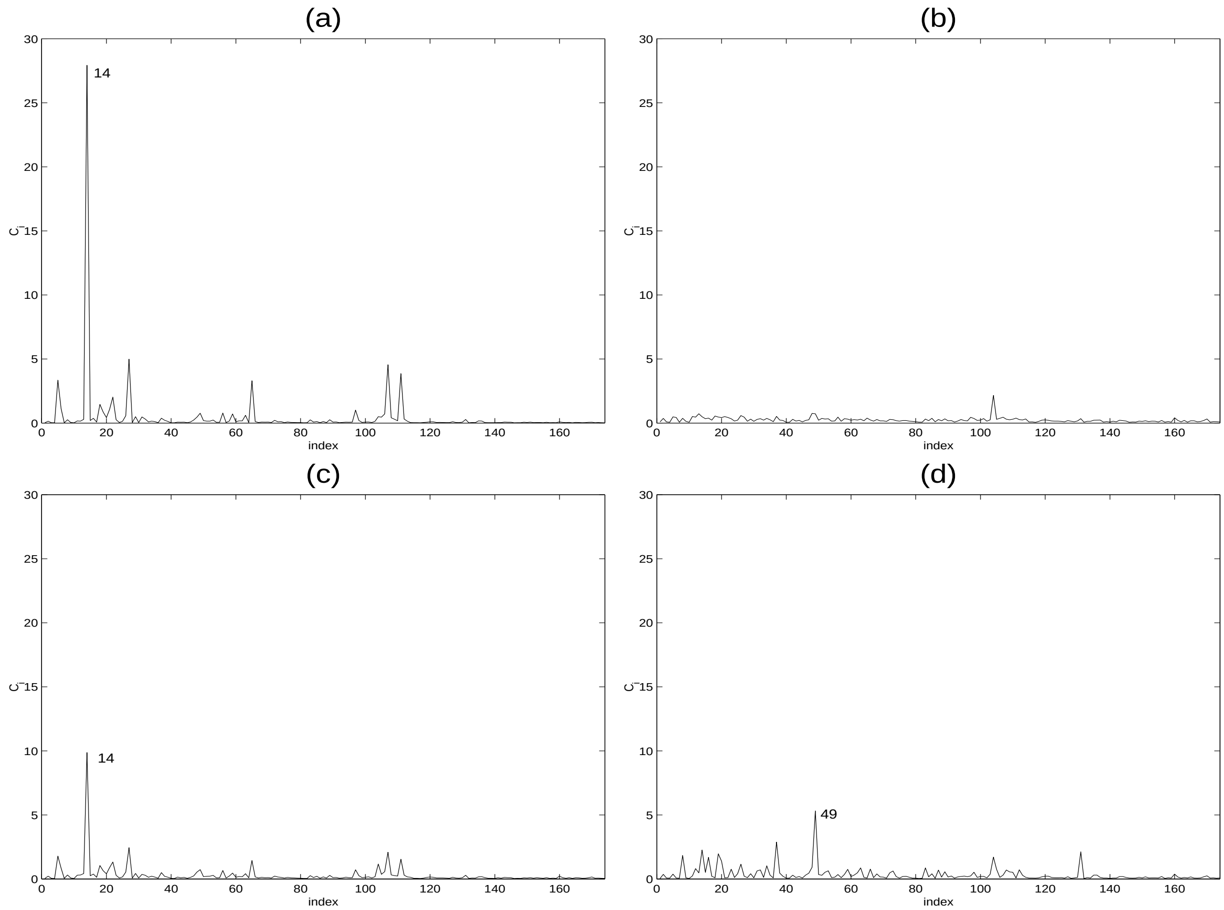

Figura 6: Index plots of C_k for market perturbations in the (a) Normal (b) $\boldsymbol{t}_4$ (c) Power Exponential and (d) Contaminated Normal models

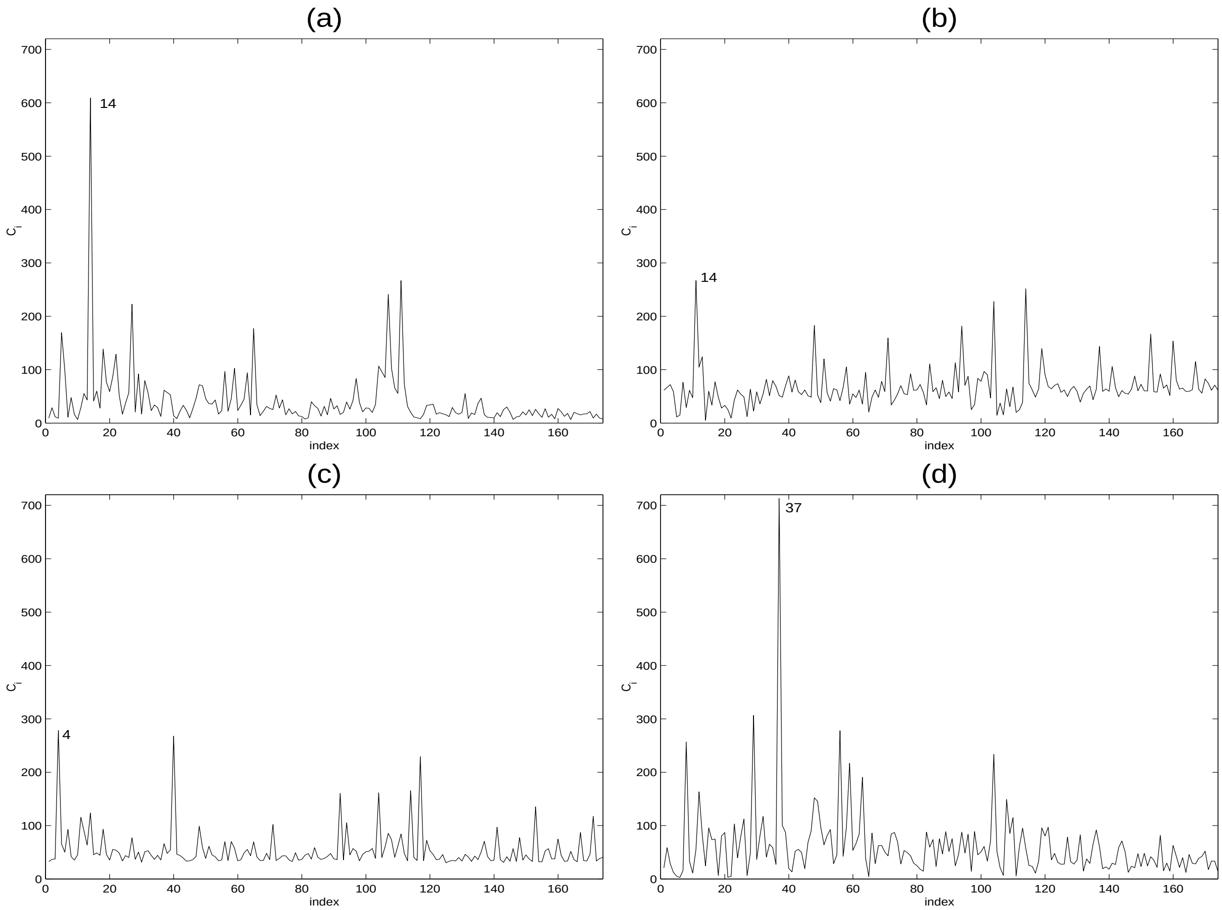

Figura 7: Index plots of C_k for scale perturbations in the (a) Normal (b) $\boldsymbol{t}_4$ (c) Power Exponential and (d) Contaminated Normal models