European University Institute
Max Weber Programme

EUI Working Papers

MWP 2008/05

Indirect Estimation of Elliptical Stable Distributions

Marco J. Lombardi David Veredas

EUROPEAN UNIVERSITY INSTITUTE
MAX WEBER PROGRAMME

Indirect Estimation of Elliptical Stable Distributions

MARCO J. LOMBARDI DAVID VEREDAS

EUI Working Paper MWP No. 2008/05

This text may be downloaded for personal research purposes only. Any additional reproduction for other purposes, whether in hard copy or electronically, requires the consent of the author(s), editor(s). If cited or quoted, reference should be made to the full name of the author(s), editor(s), the title, the working paper or other series, the year, and the publisher.

The author(s)/editor(s) should inform the Max Weber Programme of the EUI if the paper is to be published elsewhere, and should also assume responsibility for any consequent obligation(s).

ISSN 1830-7728

© 2008 Marco J. Lombardi David Veredas

Printed in Italy
European University Institute
Badia Fiesolana
I – 50014 San Domenico di Fiesole (FI)
Italy

http://www.eui.eu/
http://cadmus.eui.eu/

Abstract

We present an indirect estimation approach for elliptical stable distributions which relies on the use of a multivariate Student-t distribution as auxiliary model. This distribution is also elliptical and we show that its parameters have a one-to-one relationship with those of the elliptical stable, therefore making the proposed indirect approach particularly suitable. We analyze the finite sample behaviour of the estimators via a comprehensive Monte Carlo study. An application to 27 emerging markets stock indexes concludes the paper.

Keywords: Stable, elliptical, high dimension, multivariate, indirect inference.

JEL classification: C13, C15, G11.

Indirect Estimation of Elliptical Stable Distributions

Marco J. LOMBARDI* and David VEREDAS†

January 2008

1 Introduction

In the recent years, the modelling of multivariate data has received increasing attention among researchers and practitioners. In the field of financial time series, the assumptions underlying multivariate modelling typically refer to the specification of the first two moments and of the distribution from which the data is assumed to be generated. Researchers have mainly focussed on conditional moments and have proposed the use of VAR and GARCH types of models, such as BEKK, CCC and DCC (cf. Bauwens, Laurent and Rombouts, 2006). As for the distribution, many financial models rely on the multivariate Gaussian distribution as a building block – for instance, the classical CAPM, factor models or the Black and Scholes option pricing equation. The reason behind this choice is twofold: on the one hand the presence of the central limit theorem in a sense justifies the appearance of a Gaussian distribution whenever the phenomenon of interest can be thought of as the aggregation of a large number of micro-contributions; on the other hand, the fact that the Gaussian family of distributions has a number of useful properties, which are very helpful in establishing theoretical results. However, using multivariate Gaussian distributions has a major shortcoming: the tails of the distribution are seldom able to accommodate for extreme gains and losses that are frequently observed on financial markets. Some alternatives have been proposed in the literature. Multivariate Student-t and its skewed version (cf. Bauwens and Laurent, 2005) are two of them. However, although they provide a clear improvement in the fit of the distribution, Student-t has the shortcoming of not belonging to a family which is close under summation; this fact makes the derivation of theoretical results much more cumbersome. An alternative is the use of copulas (cf. Patton, 2004, and references therein), which circumvents the choice of the multivariate distribution, given that the dependence structure is constructed via the specification of appropriate marginal distributions and a suitable copula function.

*Max Weber Fellow & Economics Department, EUI; email: marco.lombardi@eui.eu

†ECARES, Universit Libre de Bruxelles; email: david.veredas@ulb.ac.be. 50 Av F.D. Roosevelt CP 114, B1050 Brussels, Belgium. Phone: +32 2 6504218. Fax: +31 2 6504475. Work supported in part by the European Community's Human Potential Programme under contract HPRN-CT-2002-00232, [MICFINMA]. With the usual disclaimer, we are grateful to the editor and two anonymous referees as well as Claudia Czado, Marc Hallin, Alexander McNeil, Stefan Mittnik, Giorgio Calzolari, John Nolan, Gabriel Frahm, Davy Paindaveine, Marc Paolella, Eric Renault and Sebastien Van Bellegem for useful comments that improve the paper. Research assistantship by Delphine Cassart and David de Antonio is highly appreciated. This text presents research results of the Belgian Program on Interuniversity Poles of Attraction initiated by the Belgian State, Prime Minister's Office, Science Policy Programming. The scientific responsibility is assumed by the authors.

Among other possible heavy-tailed alternatives, the multivariate stable distribution (cf. Samorodnitsky ad Taqqu, 1994) plays a special role. It originates from a generalization of the central limit theorem in which the assumption on the finiteness of the variance of the components is replaced by a much less restrictive one concerning somewhat regular tail behaviour (cf. Ibragimov and Linnik, 1971). As a consequence, stable distributions enjoy many of the properties of the Gaussian family, including closeness under summation. Therefore, a number of theoretical results in asset allocation and option pricing are available (cf. Fama, 1965a, and 1965b; Ortobelli, Huber and Schwartz, 2002; Ortobelli and Rachev, 2005; McCulloch, 2003; and the survey by Bradley and Taqqu, 2001).

Notwithstanding the appealing properties of stable distribution, estimation has always been challenging as it is defined via the characteristic function, and the density function cannot in general be expressed in a closed form. Several techniques have been proposed for the estimation of univariate distributions (cf. the survey in Garcia, Renault and Veredas, 2006, and references therein), such as the use of characteristic functions, quantiles or approximated maximum likelihood.

At the multivariate level characteristic function methods are not operational for dimensions, say, higher than three. Neither are quantile methods applicable as the concept of multivariate quantile itself is not clear-cut. As for maximum likelihood, it is a complex issue even in the univariate case, due to the absence of the density function in closed form, and this carries over and amplifies at the multivariate level. In fact, most of the available results refer solely to the estimation of the so-called spectral measure, that is, a measure that contains information on the scale and skewness of the process.[1] Two approaches have been employed for the estimation of the spectral measure: the first is based on the multivariate characteristic function (Nolan, Panorska and McCulloch, 2001, and Pivato and Seco, 2003); the second approach is based on one-dimensional projections of the multivariate process (Nolan, Panorska and McCulloch, 2001; Rachev and Xin, 1993; and Cheng and Rachev, 1995). The only paper, to our knowledge, that estimates all the parameters of the multivariate stable distribution is Nolan (2005) which extends the above-mentioned results based on projections to the location and tail index.

These estimation difficulties have hindered the use of multivariate stable distributions in applied work and call for the use of simulation-based methods. Since random numbers from stable distributions can be obtained straightforwardly, simulation-based methods such as the Indirect Inference of Gourieroux, Monfort and Renault (1993) and Efficient Method of Moments -EMM hereafter- of Smith (1993) and Gallant and Tauchen (1996) are especially appealing. These two methods will be refereed to in what follows as indirect estimation methods. In the univariate case, indirect approaches have been proposed independently by Garcia, Renault and Veredas (2006) and Lombardi and Calzolari (2008). In this paper we move a step forward, considering an indirect approach to the estimation of elliptical stable distributions.

Elliptical stable distributions, ESD hereafter, are nested in the class of elliptical distributions, introduced by Kelker (1970).[2] This class is particularly relevant as it contains important distributions, some of them already mentioned (Gaussian, Student-t and ESD), and it possesses many of the attractive properties like the Gaussian and the stable. For instance, they are invariant to affine transformations, their marginal and conditional distributions are also elliptical, and they are closed under convolution. The fact that the elliptical class of distributions includes the Student-t and the ESD suggests that an indirect estimation approach could be fruitful.

[1] As pointed out in Pivato and Seco (2003), the spectral measure should be called Feldheim measure, after Feldheim (1937), and not spectral as it is unrelated to any other "spectral measures" currently existent in statistics.

[2] See Cambanis, Huang and Simons (1981) and Fang, Kotz and Ng (1990) for further references. A good, short and concise survey is chapter one of Frahm (2004).

According to indirect methods, an auxiliary model, easy to estimate, replaces the model of interest, and simulations performed under the latter are then used to match the estimates obtained on real and simulated data. The fact that the model of interest and the auxiliary model belong to the same family and share the same structure is helpful in establishing the asymptotic properties, as the parameters have a natural one-to-one relationship.

Standard asymptotic theory of indirect estimation can be applied, as the information content on the parameters of Student-t is sufficient to identify the parameters of the ESD and the score of the Student-t distribution is asymptotically Gaussian. However, in finite sample asymptotics do not apply. The problem, highlighted in Garcia, Renault and Veredas (2006) and Lombardi and Calzolari (2008), is that as the tail index of the stable distribution approaches two, and hence the distribution approaches the Gaussian, the degrees of freedom of the Student-t are attracted by infinity. While this should not be a problem asymptotically, it entails important estimation difficulties with finite samples. To avoid it, we constraint the degrees of freedom to remain below an upper bound, therefore resorting to the constrained indirect estimation of Calzolari, Fiorentini and Sentana (2004).

A comprehensive Monte Carlo study for different values of the tail index, in two and five dimensions shows that the finite sample properties of the estimates are reasonably good, unbiased in virtually all cases and with root mean square errors that decrease with the number of indirect optimizations. The empirical application is on weekly Morgan Stanley Corporate Indexes (MSCI) of 27 emerging markets. We estimate the ESD on standardized residuals, demeaned and filtered by a GARCH(1,1) model, and we show that the tail index is below two and the estimated scatter matrix mimics the empirical correlation matrix.

The plan of the paper is as follows. Section 2 introduces elliptical distributions and, in particular, ESD and Student-t. Section 3 presents the indirect estimation methods and proves its asymptotic properties in our setting as well as the one-to-one relationship between the parameters of the two distributions. A detailed simulation study highlights the small sample properties of the estimators in Section 4. Next, we illustrate the method by applying it to 27 emerging markets indexes and Section 6 concludes and gives directions for further research.

2 Elliptical Distributions

A k dimensional random vector $\mathbf{X}$ is elliptically distributed if

$$\mathbf{X} =^d \boldsymbol{\mu} + \mathcal{R}\boldsymbol{\Lambda}\mathbf{U}^{(k)},$$

where $\boldsymbol{\mu}$ is a k dimensional vector of location parameters, $\boldsymbol{\Lambda}$ is a $k \times k$ full rank arbitrary matrix of scale parameters and $\mathbf{U}^{(k)}$ is a k dimensional random vector uniformly distributed in the unit sphere with $k-1$ dimensions

$$\mathcal{S}^{k-1} = \left\{ \mathbf{x} \in \mathbb{R}^k : \|\mathbf{x}\|_2 = 1 \right\}.$$

$\mathcal{R}$ is the so-called generating variate of $\mathbf{X}$. It is a non-negative random variable stochastically independent of $\mathbf{U}^{(k)}$. The starting point in the construction of an elliptically distributed random variable is $\mathbf{U}^{(k)}$, which is radial. It is premultiplied by $\boldsymbol{\Lambda}$, such that $\boldsymbol{\Lambda}\mathbf{U}^{(k)}$ is no longer radial but rather elliptical, with the generating variate $\mathcal{R}$ giving the thickness, or thinness, of the tails of $\mathcal{R}\boldsymbol{\Lambda}\mathbf{U}^{(k)}$. The vector $\boldsymbol{\mu}$ shifts the location of the density. If $\boldsymbol{\Lambda}$ equals the identity matrix, the density of $\mathbf{X}$ remains radial. $\boldsymbol{\Lambda}$ is a matrix such that $\boldsymbol{\Sigma} = \boldsymbol{\Lambda}\boldsymbol{\Lambda}'$ is a positive definite matrix of rank k and $\mathbf{\Sigma}$ is called the dispersion or scatter matrix of $\mathbf{X}$. We are ultimately interested in $\mathbf{\Sigma}$, though elliptical distributions are expressed in terms of $\mathbf{\Lambda}$. In fact, the decomposition of $\mathbf{\Sigma}$ in terms of $\mathbf{\Lambda}$ is itself irrelevant as $\mathbf{\Lambda}$ is not identified.[3]

Some important multivariate densities belong to the class of elliptical distributions: Gaussian, Student-t and ESD among others.[4] We obtain a Gaussian distribution if $\mathcal{R} = \sqrt{\chi^2_k}$. Similarly, a Student-t is obtained if $\mathcal{R} = \sqrt{\nu\chi^2_k/\chi^2_\nu}$ where χ^2_k and χ^2_ν are stochastically independent. Finally, we obtain an ESD if $\mathcal{R} = \sqrt{\chi^2_k}\sqrt{S_{\alpha/2}}$, where $S_{\alpha/2}$ is a positive, and hence totally skewed to the right, $\alpha/2$ stable distributed random variable and χ^2_k and $S_{\alpha/2}$ are stochastically independent.

From these examples it is evident that there is a close connection between Student-t and ESD. The location and scale parameters play the same role in both distributions. The tail parameter, either α or ν, enters in both cases through the generating variate. This leads to the intuitive idea, proven in the next section, that if the true data generating process is stable but the assumed distribution is Student-t, a change in the location of the elliptical stable process will lead to a change in the location in Student-t, and likewise for a change in the scale.

The class of elliptical distributions possesses a number of useful properties, among which we highlight its closeness to affine transformations, conditional, and marginal distributions being also elliptical and closeness to aggregation (cf. Fang, Kotz and Ng, 1990, for further details.). As for the last property, it is worth stressing the difference between elliptical and stable distributions. Indeed, the sum of *i.i.d.* elliptically distributed random vectors remains elliptical in the sense that the resulting distribution belongs to the elliptical class, but not necessarily to the same family as that of their addends. The latter property is instead possessed by stable distributions.

Another important property of elliptical distributions is that the density function can be expressed in terms of the density function of the generating variate. More precisely, the *pdf* of $\mathbf{X}$ is given by

$$f_{\mathbf{X}}(\mathbf{x}) = \sqrt{|\mathbf{\Sigma}^{-1}|}g_{\mathcal{R}}\left((\mathbf{x}-\boldsymbol{\mu})'\mathbf{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right) \tag{1}$$

where $|\cdot|$ denotes the determinant,

$$g_{\mathcal{R}}(t) = \frac{\Gamma\left(\frac{k}{2}\right)}{(2\pi)^{k/2}}\sqrt{t}^{-(k-1)}f_{\mathcal{R}}(\sqrt{t})$$

and $f_{\mathcal{R}}$ is the *pdf* of the generate variate. For instance in the case of Student-t

$$f_{\mathcal{R}}(t) = \frac{2t}{k}f_F\left(\frac{t^2}{k}\right),$$

where f_F represents the *pdf* of a $F_{k,\nu}$ distributed random variable and hence $f_{\mathcal{R}}(t)$ is the *p.d.f.* of the random variable $\sqrt{kF_{k,\nu}}$. After some arrangements the *pdf* of $\mathbf{X}$ takes the form

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{\Gamma\left(\frac{k+\nu}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)}\left(\frac{|\mathbf{\Sigma}^{-1}|}{(\nu\pi)^k}\right)^{1/2}\left(1+\frac{(\mathbf{x}-\boldsymbol{\mu})'\mathbf{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})}{\nu}\right)^{-\frac{k+\nu}{2}}, \tag{2}$$

[3]Indeed, let $\mathbf{T}$ be an orthonormal matrix. Then $\mathbf{\Sigma} = \mathbf{\Lambda}\mathbf{\Lambda}' = \mathbf{\Lambda}\mathbf{T}\mathbf{T}'\mathbf{\Lambda}' = \mathbf{\Lambda}^*\mathbf{\Lambda}^{*'}$ and therefore $\mathbf{\Lambda}$ and $\mathbf{\Lambda}^*$ generate the same scatter matrix.

[4]Hereafter we will skip the term multivariate. Nonetheless, the reader should always keep in mind that $\mathcal{R}$ is a random variable and $\mathbf{U}^{(k)}$ is a random vector, thus $\mathbf{X}$ is a random vector as well.

which is a Student-t density with ν degrees of freedom and $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ are the location and scale parameters. Unfortunately, an equivalent closed form expression of (2) for ESD does not exist. However, the fact that Student-t and ESD are elliptical paves the road to the use of indirect methods.

3 Indirect Estimation

Let $\mathbf{x}$ be a sample of T *i.i.d.* copies from an ESD. Given (1), its log likelihood is[5]

$$\ln \ell^{\star}(\boldsymbol{\theta}, \mathbf{x}) = \frac{1}{2} \ln |\boldsymbol{\Sigma}^{-1}| + \ln g_{\mathcal{R}} \left((\mathbf{x} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu}) \right),$$

where $\boldsymbol{\theta} = (\alpha, \boldsymbol{\Sigma}, \boldsymbol{\mu}) \in \boldsymbol{\Theta} =]0, 2[\times \mathbb{R}_{++}^{k \times k} \times \mathbb{R}^k$, α is the tail index, $\boldsymbol{\Sigma}$ a $k \times k$ positive definite scatter matrix, $\boldsymbol{\mu}$ a $k \times 1$ location parameter vector and $g_{\mathcal{R}}$ is the generating variate of $\sqrt{\chi_k^2} \sqrt{S_{\alpha/2}}$. This is the *model of interest*. However, as previously noted, this log likelihood does not admit a closed-form expression and is therefore difficult to evaluate.[6] Instead, we assume, mistakenly but on purpose, that $\mathbf{x}$ follows a Student-t distribution with density (2) and therefore we can easily maximize its log likelihood:

$$\ln \tilde{\ell}(\boldsymbol{\zeta}, \mathbf{x}) = \ln \frac{\Gamma\left(\frac{\nu+k}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} + \frac{1}{2} \ln \left(\frac{|\boldsymbol{\Psi}^{-1}|}{(\nu\pi)^k} \right) - \frac{k+\nu}{2} \ln \left(1 + \frac{\left((\mathbf{x} - \boldsymbol{\delta})' \boldsymbol{\Psi}^{-1} (\mathbf{x} - \boldsymbol{\delta}) \right)}{\nu} \right),$$

where $\boldsymbol{\zeta} = (\nu, \boldsymbol{\Psi}, \boldsymbol{\delta}) \in \mathbf{Z} =]0, \infty[\times \mathbb{R}_{++}^{k \times k} \times \mathbb{R}^k$, ν is the tail index, $\boldsymbol{\Psi}$ a $k \times k$ positive definite scatter matrix and $\boldsymbol{\delta}$ a $k \times 1$ location parameter vector. This is the *auxiliary model*.[7] Since this model is misspecified, the estimators that maximize the above log-likelihood, $\hat{\boldsymbol{\zeta}}(\mathbf{x})$, are not necessarily consistent. The central idea of indirect methods is to exploit simulations under the model of interest to find parameter values that match the estimates of the auxiliary parameters obtained on actual data.

Let $\mathbf{x}_s(\boldsymbol{\theta})$, $s = 1, \ldots, S$, be a simulated sample of T *i.i.d.* copies from an ESD and for a given arbitrary parameter vector $\boldsymbol{\theta}$. And let

$$\hat{\boldsymbol{\zeta}}_s(\boldsymbol{\theta}) = \arg \max_{\boldsymbol{\zeta} \in \mathbf{Z}} \ln \tilde{\ell}(\boldsymbol{\zeta}; \mathbf{x}_s(\boldsymbol{\theta}))$$

be the maximum likelihood estimator of the Student-t distribution. Furthermore let

$$\hat{\boldsymbol{\zeta}}_S(\boldsymbol{\theta}) = \frac{1}{S} \sum_{s=1}^{S} \hat{\boldsymbol{\zeta}}_s(\boldsymbol{\theta}).$$

The Indirect Inference estimates, $\hat{\boldsymbol{\theta}}(\mathbf{x})$, are the values for which the following distance is minimized:

$$\left[\hat{\boldsymbol{\zeta}}(\mathbf{x}) - \hat{\boldsymbol{\zeta}}_S(\boldsymbol{\theta}) \right] \boldsymbol{\Omega} \left[\hat{\boldsymbol{\zeta}}(\mathbf{x}) - \hat{\boldsymbol{\zeta}}_S(\boldsymbol{\theta}) \right],$$

[5] From this section on, we slightly change the notation, denoting differently the location and scatter parameters of each law.

[6] It should be emphasized that difficult does not imply impossible. Nolan (2005) shows how to compute the log-likelihood numerically.

[7] Admittedly, Student-t is not the only *good* candidate for the auxiliary model. For instance, the symmetric generalized hyperbolic distribution is also appropriate. The choice of Student-t motivated by Demarta and McNeil (2005), Frahm, Junker and Szimayer (2005), and Frahm (2006) who suggest this distribution as a reference model for elliptically contoured distributions.

where $\mathbf{\Omega}$ is a symmetric non-negative matrix defining the metric.[8] Alternatively, EMM considers directly the score of Student's t

$$\sum_{t=1}^{T} \frac{\partial \ln \tilde{\ell}(\boldsymbol{\zeta}; \mathbf{x})}{\partial \boldsymbol{\zeta}}.$$

The EMM estimates, $\check{\boldsymbol{\theta}}(\mathbf{x})$, are the values for which the following distance is minimized:

$$\left\{ \sum_{s=1}^{S} \frac{\partial \ln \tilde{\ell}\left(\boldsymbol{\zeta}; \mathbf{x}_s(\boldsymbol{\theta})\right)}{\partial \boldsymbol{\zeta}} \right\}' \mathbf{\Upsilon} \left\{ \sum_{s=1}^{S} \frac{\partial \ln \tilde{\ell}\left(\boldsymbol{\zeta}; \mathbf{x}_s(\boldsymbol{\theta})\right)}{\partial \boldsymbol{\zeta}} \right\},$$

where $\mathbf{\Upsilon}$ is a symmetric non-negative definite matrix. Gourieroux, Monfort and Renault (1993) shown that, choosing the weighting matrices appropriately, the two methods are asymptotically equivalent in the sense that their estimators have the same asymptotic distribution.

In order to identify $\boldsymbol{\theta}$ it is necessary for the dimension of $\boldsymbol{\zeta}$ to be at least as big as that of $\boldsymbol{\theta}$. If both dimensions are equal, as is the case of the elliptical distributions considered in this article, $\check{\boldsymbol{\theta}}(\mathbf{x})$ does not depend on $\mathbf{\Upsilon}$ and one can choose the Indirect Inference or the EMM estimators that best suits the best for the practical problem to be analyzed. For instance, EMM is especially useful when an analytic expression for the gradient of the auxiliary model is available, since it allows us to avoid the numerical optimization routines in the estimation of the auxiliary model. [9]

The asymptotic behaviour of the log-likelihood of the auxiliary model is

$$\lim_{T \to \infty} \frac{1}{T} \ln \tilde{\ell}(\boldsymbol{\zeta}; \mathbf{x}_s(\boldsymbol{\theta})) = E_\theta \left[\ln \tilde{\ell}(\boldsymbol{\zeta}; \mathbf{x}_s(\boldsymbol{\theta})) \right],$$

and the solution of the maximization problem is

$$\mathbf{b}(\boldsymbol{\theta}) = \arg \max_{\boldsymbol{\zeta} \in Z} E_{\boldsymbol{\theta}} \left[\ln \tilde{\ell}(\boldsymbol{\zeta}; \mathbf{x}_s(\boldsymbol{\theta})) \right].$$

That is $\hat{\boldsymbol{\zeta}}_S(\boldsymbol{\theta})$ is a consistent estimator of $\mathbf{b}(\boldsymbol{\theta})$, the *binding function* that maps the parameters space of the true model onto the parameter space of the auxiliary model. The indirect estimator of $\boldsymbol{\theta}$ is thus based on the evaluation of the binding function at the true optimum $\boldsymbol{\theta}_0$. $\mathbf{b}(\boldsymbol{\theta})$ defines the pseudo-true value of the Student-t parameters when the true probability distribution is ESD. The fact that the model of interest and the auxiliary model belong to the same family of elliptical distributions allows us to devise a one-to-one relationship between the binding function and $\boldsymbol{\theta}$. Intuitively, the location parameters are the same for both distributions and the difference in the tail behaviour between the two generating variates is exclusively given by α and ν. Hence, $\mathbf{\Psi}$ is very informative for estimating $\mathbf{\Sigma}$. Following Garcia, Renault and Veredas (2006), we denote

$$\begin{aligned} \mathbf{b}_1(\alpha, \boldsymbol{\mu}, \mathbf{\Sigma}) &= \nu \\ \mathbf{b}_2(\alpha, \boldsymbol{\mu}, \mathbf{\Sigma}) &= \boldsymbol{\delta} \\ \mathbf{b}_3(\alpha, \boldsymbol{\mu}, \mathbf{\Sigma}) &= \mathbf{\Psi} \end{aligned}$$

The following proposition proves that the relationship is one to one.

[8] Typically, the optimal matrix is the inverse of the product of the scores.

[9] Because they are equivalent, hereafter we will use them indistinguishably although we will favour the EMM estimator for reasons that will become clear at the end of this section.

Proposition Let $\mathbf{a} \in \mathbb{R}^k$ be a k dimensional random vector sampled from an elliptical density and $\mathbf{\Delta} \in \mathbb{R}^{k\times k}$ an arbitrary matrix of full rank. Then for any k dimensional vector $\boldsymbol{\mu} \in \mathbb{R}^k$ and scatter matrix $\mathbf{\Sigma}$ such that $\mathbf{\Sigma} = \mathbf{\Lambda}\mathbf{\Lambda}'$, $\mathbf{\Lambda} \in \mathbb{R}^{k\times k}$,

$$\begin{aligned} \mathbf{b}_1(\alpha, \boldsymbol{\mu} + \mathbf{a}, \mathbf{\Delta\Sigma\Delta}') &= \nu \\ \mathbf{b}_2(\alpha, \boldsymbol{\mu} + \mathbf{a}, \mathbf{\Sigma}) &= \boldsymbol{\delta} + \mathbf{a} \\ \mathbf{b}_3(\alpha, \boldsymbol{\mu}, \mathbf{\Delta\Sigma\Delta}') &= \mathbf{\Delta\Psi\Delta}'. \end{aligned}$$

That is $\boldsymbol{\delta} \leftrightarrow \boldsymbol{\mu}$, $\mathbf{\Psi} \leftrightarrow \mathbf{\Sigma}$ and $\alpha \leftrightarrow \nu$.

Proof Consider an elliptically-distributed random vector

$$\mathbf{X} =^d \boldsymbol{\mu} + \mathcal{R}^S \mathbf{\Lambda} \mathbf{U}^{(k)},$$

where $\mathcal{R}^S = \sqrt{\chi^2_k}\sqrt{S_{\alpha/2}}$. Its characteristic function corresponds to

$$\varphi_{\mathbf{X}^S}(t) = \exp(it'\boldsymbol{\mu}) \int_0^\infty \varphi_{\mathbf{U}}(r^2 t' \mathbf{\Sigma} t) dF_{\mathcal{R}^S}(r),$$

where $\varphi_{\mathbf{U}}(\cdot)$ is the characteristic function of $\mathbf{U}^{(k)}$ and $F_{\mathcal{R}^S}$ is the cdf of $\sqrt{\chi^2_k}\sqrt{S_{\alpha/2}}$. Consider instead the integration with respect to $F_{\mathcal{R}^{St}}$, the cdf of $\sqrt{\nu\chi^2_k/\chi^2_\nu}$

$$\varphi_{\mathbf{X}^{St}}(t) = \exp(it'\boldsymbol{\mu}) \int_0^\infty \varphi_{\mathbf{U}}(r^2 t' \mathbf{\Sigma} t) dF_{\mathcal{R}^{St}}(r).$$

A change in the location and the scale $\mathbf{Y} := \mathbf{a} + \mathbf{\Delta X}$, corresponds to

$$\begin{aligned} \varphi_{\mathbf{Y}^S}(t) &= E[\exp(it'(\mathbf{a} + \mathbf{\Lambda X}))] \\ &\quad \exp(it'(\boldsymbol{\mu} + \mathbf{a})) \int_0^\infty \varphi_d(r^2 t' \mathbf{\Delta\Sigma\Delta}' t) dF_{\mathcal{R}^S}(r), \end{aligned}$$

and as well to

$$\begin{aligned} \varphi_{\mathbf{Y}^{St}}(t) &= E[\exp(it'(\mathbf{a} + \mathbf{\Lambda X}))] \\ &\quad \exp(it'(\boldsymbol{\mu} + \mathbf{a})) \int_0^\infty \varphi_d(r^2 t' \mathbf{\Delta\Sigma\Delta}' t) dF_{\mathcal{R}^{St}}(r), \end{aligned}$$

which are the characteristic functions of the elliptically distributed random vectors, $\mathbf{Y}^{St} =^d (\boldsymbol{\mu} + \mathbf{a}) + \mathcal{R}^{St}\mathbf{\Delta\Lambda U}^{(k)}$ and $\mathbf{Y}^S =^d (\boldsymbol{\mu} + \mathbf{a}) + \mathcal{R}^S\mathbf{\Delta\Lambda U}^{(k)}$, with identical location and scatter matrices.$\square$

This means that a change in the location only affects the location parameter and a scale change only affects the scatter matrix. Moreover, the generating variate of the transformed vector remains the same. Therefore even if we estimate with $\mathcal{R} = \sqrt{\nu\chi^2_k/\chi^2_\nu}$, the affine transformation does not affect the tail index. In other words, the location and scale parameters of the Student-t carry over exclusively information on the location and scale parameters of the ESD respectively: $\boldsymbol{\delta} \leftrightarrow \boldsymbol{\mu}$ and $\mathbf{\Psi} \leftrightarrow \mathbf{\Sigma}$. Hence the tail index is not modified by location and scale changes but by the stability index: $\alpha \leftrightarrow \nu$.

Under the C regularity conditions (see Appendix), the EMM estimator $\check{\boldsymbol{\theta}}(\mathbf{x})$ is consistent for fixed S and $T \to \infty$. Furthermore, $\check{\boldsymbol{\theta}}(\mathbf{x})$ is asymptotically Gaussian for fixed S and $T \to \infty$ and the asymptotic variance-covariance matrix of $\sqrt{T}(\check{\boldsymbol{\theta}}(\mathbf{x}) - \boldsymbol{\theta}_0)$ is where

$$\mathcal{W}(S,\Upsilon) = \left(1+\frac{1}{S}\right)\left[\frac{\partial^2 E_{\boldsymbol{\theta}}[\ln\tilde{\ell}(\boldsymbol{\zeta},\mathbf{x}(\boldsymbol{\theta}))]}{\partial\boldsymbol{\zeta}\partial\boldsymbol{\theta}'}^{\prime *}\boldsymbol{\Upsilon}^{*}\frac{\partial^2 E_{\boldsymbol{\theta}}[\ln\tilde{\ell}(\zeta,\mathbf{x}(\boldsymbol{\theta}))]}{\partial\boldsymbol{\zeta}\partial\boldsymbol{\theta}'}\right]^{-1} \quad (3)$$

where

$$\boldsymbol{\Upsilon}^{*} = \lim_{T\to\infty} Var\left\{\sqrt{T}\frac{\partial\ln\tilde{\ell}\left[\mathbf{b}(\boldsymbol{\theta});\mathbf{x}\right]}{\partial\boldsymbol{\zeta}}\right\}.$$

A consistent estimator for $\check{\mathcal{W}}$ is[10]

$$\check{\mathcal{W}}(S,\boldsymbol{\Upsilon}) = \left(1+\frac{1}{S}\right)\left[\frac{\partial^2\ln\tilde{\ell}(\boldsymbol{\zeta};\mathbf{x})}{\partial\boldsymbol{\theta}\partial\boldsymbol{\zeta}'}^{\prime}\check{\Upsilon}^{*}\frac{\partial^2\ln\tilde{\ell}(\boldsymbol{\zeta};\mathbf{x})}{\partial\boldsymbol{\theta}'\partial\boldsymbol{\zeta}}\right]^{-1}$$

where

$$\check{\Upsilon}^{*} = \frac{1}{T}\left[\left.\frac{\partial\ln\tilde{\ell}(\boldsymbol{\zeta};\mathbf{x})'}{\partial\boldsymbol{\zeta}}\right|_{\boldsymbol{\zeta}=\hat{\boldsymbol{\zeta}}}\left.\frac{\partial\ln\tilde{\ell}(\boldsymbol{\zeta};\mathbf{x})}{\partial\boldsymbol{\zeta}'}\right|_{\boldsymbol{\zeta}=\hat{\boldsymbol{\zeta}}}\right].$$

Indirect estimators are asymptotically well-behaved because the information content on the parameters of Student-t is sufficient to identify the parameters of ESD and the score of Student-t distribution is asymptotically Gaussian. However, in finite samples the information content in ν is not sufficient to identify α as it approaches 2 because $\check{\nu}$ tends to infinity. To avoid this we constrain ν to remain below an upper bound $\bar{\nu}$. Let

$$\hat{\boldsymbol{\beta}}(\mathbf{x}) = \arg\max_{\boldsymbol{\beta}\in\mathbf{Z}\times\mathbb{R}}\ln\tilde{\ell}(\boldsymbol{\zeta};\mathbf{x}(\boldsymbol{\theta})) + (\nu-\bar{\nu})\rho$$

be the constrained estimator of the Student's t distribution that satisfies the inequality restriction plus the slackness restriction $(\nu-\bar{\nu})\rho = 0$. The parameter set is $\boldsymbol{\beta} = (\boldsymbol{\zeta},\rho) \in \mathbf{Z}\times\mathbb{R}$ and $\bar{\nu}$ is the upper bound for ν. Equivalently for $\hat{\boldsymbol{\beta}}(\boldsymbol{\theta})$.

Calzolari, Fiorentini and Sentana (2004) have shown that the EMM estimator in the presence of constraints is analogous to that derived by Gallant and Tauchen (1996) and the weighting matrix remains the same. The reason why theoretical results for EMM remain unchanged is that the score is taken with respect to $\boldsymbol{\zeta}$ and not with respect to ρ and hence $\boldsymbol{\theta}$ remains exactly identified. However, results change for Indirect Inference as the multiplier ρ is also minimized and therefore $\boldsymbol{\theta}$ is overidentified and an optimal weighting matrix is needed. Furthermore, this optimal matrix takes a complicated form as it accounts for the inequalities and the slackness condition. All in all, the inclusion of constraints in Student-t distribution does not change standard EMM estimation while it does in Indirect Inference. For this reason we choose the former method for the Monte Carlo study and the empirical illustration.

4 Monte Carlo study

Simulating from an ESD is fairly simple. This is due to the fact that the tail index appears only in the generating variate, which is univariate. In order to simulate random numbers from an

[10]The following expressions are specific to the *i.i.d.* case. The general expressions for serial dependence can be found in Appendix 2 of Gourieroux, Monfort and Renault (1993).

ESD it suffices to simulate from its univariate counterpart – using, for instance, the Chambers, Mallows and Stuck (1976) method. The ESD can be rewritten as

$$\mathbf{X} =^d \boldsymbol{\mu} + \sqrt{S_{\alpha/2}}\mathbf{G}, \tag{4}$$

where $\mathbf{G} = \sqrt{\chi_k^2}\boldsymbol{\Lambda}\mathbf{U}^{(k)} \sim \mathcal{N}(0, \boldsymbol{\Sigma})$. Therefore to simulate $\mathbf{X}$ we only need to simulate from a multivariate Gaussian density and from a univariate stable density. More precisely, if

$$A \sim S_{\alpha/2}\left(\left(\cos\frac{\pi\alpha}{4}\right)^{2/\alpha}, 1, 0\right)$$

and $\mathbf{G} \sim \mathcal{N}(0, \boldsymbol{\Sigma})$ independent of A, then $A^{1/2}\mathbf{G} \sim S_\alpha(\Sigma, 0, \mu)$. Notice that if α approaches 2, then $\sqrt{S_{\alpha/2}}\sqrt{\chi_k^2} \rightarrow \sqrt{\chi_k^2}$. This is a counter-intuitive result as A has a location parameter 0 and a scale that equals 0 for $\alpha = 2$. However, A converges to a Dirac delta measure. To see this, take the Laplace transform of $\mathbf{X}$ defined in (4):

$$E\left(\exp(-\gamma A)\right) = \exp(-\gamma^{\alpha/2}).$$

As $\alpha \rightarrow 2$, $\exp(-\gamma^{\alpha/2}) \rightarrow \exp(-\gamma)$, which is the Laplace transform of a Dirac delta function. That is, $\sqrt{S_{\alpha/2}}$ converges in distribution to a degenerate random variable with value 1. Because an elliptical stable random vector can be viewed as scale mixture of a normal random vector, it is also referred to as a sub Gaussian random vectors.

We use this procedure for the Monte Carlo study. To check the finite sample properties of the estimators, we choose three different values of α: 1.7, 1.9 and 1.95. The upper bound $\bar{\nu}$ is set to 100. Note that whether the bound is higher or lower is not important (cf Garcia, Renault and Veredas, 2006) as what matters is that the estimated degrees of freedom are not attracted by infinity. We simulate 500 draws of 500 observations for a grid of different values for α, two different dimensions (2 and 5) and two different values of the indirect draws (S equal to 1 and 5). The location parameter vector $\boldsymbol{\mu}$ is set to zero. For dimension two, the scatter matrix is set to

$$\boldsymbol{\Sigma} = \begin{pmatrix} 0.5 & 0.9 \\ 0.9 & 2 \end{pmatrix}$$

and for dimension five, Σ is set the following block-diagonal structure:

$$\boldsymbol{\Sigma} = \begin{pmatrix} 0.25 & 0.25 & 0.4 & 0 & 0 \\ 0.25 & 0.5 & 0.4 & 0 & 0 \\ 0.4 & 0.4 & 1 & 0 & 0 \\ 0 & 0 & 0 & 2 & 2.55 \\ 0 & 0 & 0 & 2.55 & 4 \end{pmatrix}$$

In the two-dimensional case, the off-diagonal element can be seen, loosely speaking, as a correlation (or, more precisely, a standardized covariation), as one of the diagonal elements is the inverse of the other. Likewise, for the 5-dimensional case the random variables are positively block correlated.[11] Monte Carlo results for alternative parameter configurations are available in Lombardi and Veredas (2007) and are not presented here for the sake of brevity.

[11] In fact, the off-diagonal values have been chosen such that the standardized covariation matrices are, approximately,

$$\boldsymbol{\Xi} = \begin{pmatrix} 1 & 0.7 & 0.8 & 0 & 0 \\ 0.7 & 1 & 0.6 & 0 & 0 \\ 0.8 & 0.6 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0.9 \\ 0 & 0 & 0 & 0.9 & 1 \end{pmatrix}$$

This case may correspond to a portfolio composed of a risk-free and a risky subset of assets, with positive covariation in-between but not within.

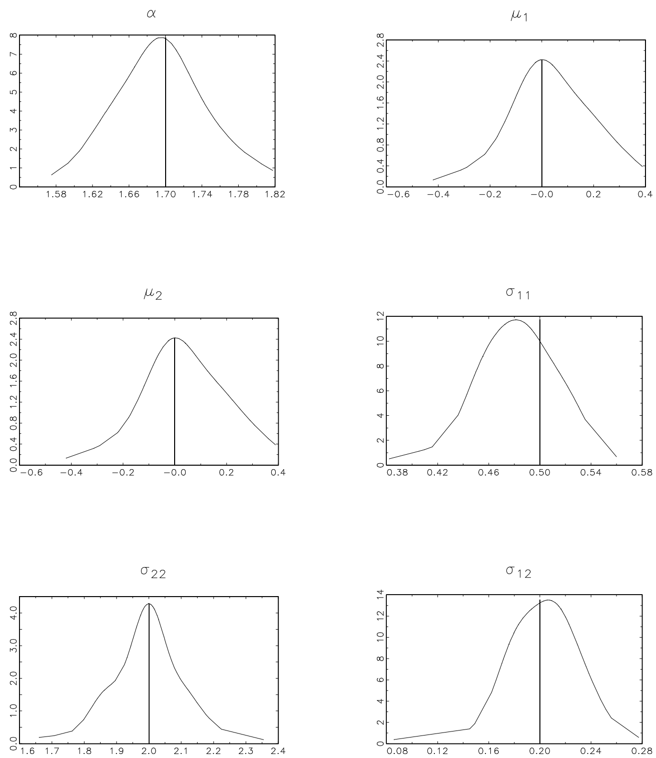

Figure 1: Densities of the estimated parameters of one the Monte Carlo scenarios for dimension 2.

Table 1 shows the results for the two dimensional case while Tables 2 and 3 display results in five dimensions. Due to space constraints we only report mean, median, root mean square error (RMSE), the mean absolute error (MAE) and, for α, the coverage rates of the asymptotic and Monte Carlo confidence sets (defined as the proportion, over 500, of the true alpha lying within the 5% confidence interval).[12] We first remark that, due to the presence of some pathological cases, the mean estimate and the RMSE are sometimes altered, and the median (and the associated MAE) proves to be a more robust alternative. In general, the results indicate that the estimators are unbiased. Furthermore, small biases present when $S = 1$ are apparently corrected when $S = 5$. Admittedly, a closer look at higher moments – not reported here – such as skewness and kurtosis, reveals that the distribution is not exactly Gaussian. Yet, this is not surprising as ours is a finite sample exercise. Nonetheless, the density of the estimators are, in some sense, well behaved. Figure 1 shows the kernel densities for one of the scenarios in dimension 2. Though not Gaussian, they do not present remarkable skewness or kurtosis.

In some cases, as α approaches 2, MAE tends to decrease, but this comes at the cost of an increased RMSE. This could be a signal that a value of α close to its bound raises problems of convergence that affect the performance of RMSE. This is also related to the poor performance of the asymptotic coverage rates of 95% asymptotic confidence intervals, which tend to deteriorate as α approaches 2, meaning that asymptotic standard errors are underestimated. Instead, if one uses the Monte Carlo standard errors (i.e. the standard deviation of the Monte Carlo distribution of the estimator), the issue seems to be solved and coverage rates appear correct. The lesson is therefore that one should be very careful in

[12]More detailed results, including standard deviation, maximum and minimum are available upon request.

estimating the asymptotic standard errors when the estimated parameter is near the boundary region.

5 Illustration

We illustrate the method with an application to 27 MSCI (Morgan Stanley Composite Index) emerging markets indexes. The MSCI indexes are free float-adjusted market capitalization indexes that are designed to measure equity market performance. The emerging markets areas and countries we consider are: East Asia (Philippines, Sri Lanka, Pakistan, China, South Korea, India, Indonesia, Russia, Thailand, Taiwan and Malaysia), Eastern Europe (Poland, Czech Republic and Hungary), South America (Mexico, Colombia, Chile, Argentina, Brazil, Venezuela and Peru), Africa (South Africa, Egypt and Morocco) and Middle East (Israel, Turkey and Jordan).

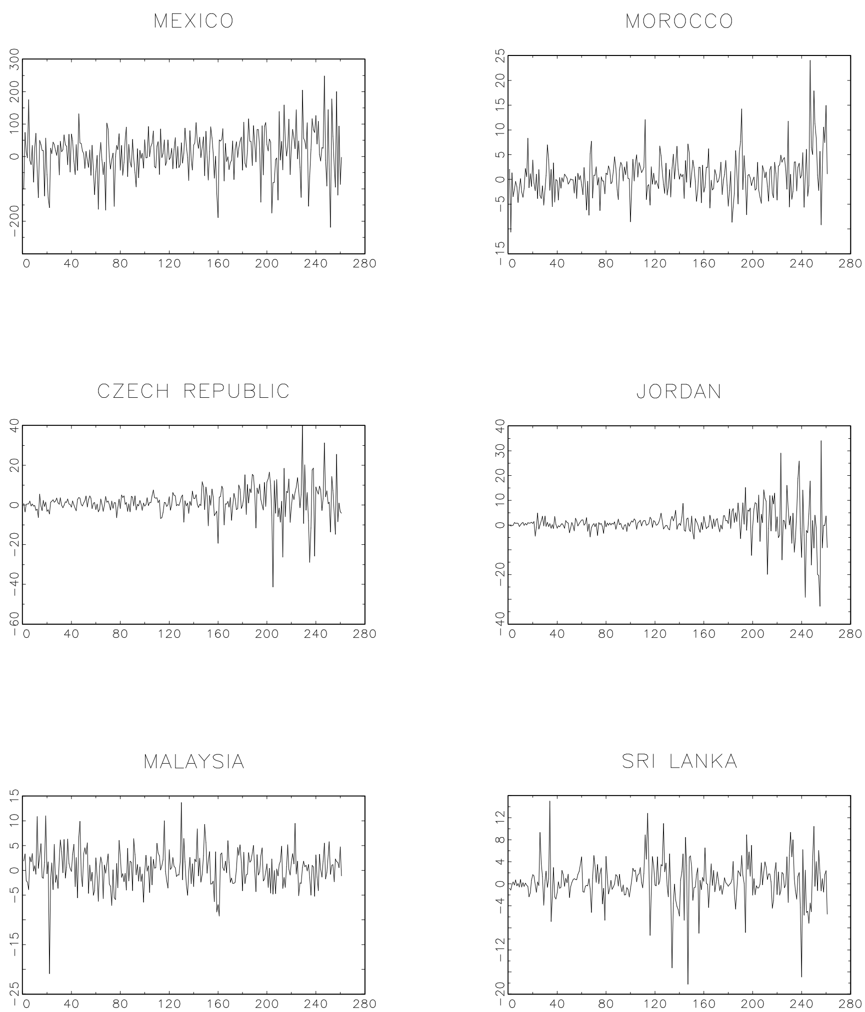

Figure 2: MSCI indexes for a selection of countries.

We use weekly returns generated by prices expressed in USD, between April 2001 to April

2006; we hence used 261 observations per country. Figure 2 shows the indexes for a sample of six countries: 2 East Asian and 1 for the other areas. Volatility behaviour is very heterogeneous. Some countries display strong clustering, like Jordan and the Czech Republic, while others present large deviations but not clusters, like Malaysia and Sri Lanka. It is known (Ghose and Kroner, 1995) that heavy tails generated by GARCH effects can be mistakenly interpreted as evidence in favour of stable distributions. To safeguard against this, we consider demeaned standardized GARCH(1,1) residuals such that the remaining heteroskedasticity is not due to dynamic conditional volatility. Denote as $\hat{\boldsymbol{\xi}}$ the QML estimated parameters with variance-covariance matrix taking the traditional sandwich form $\mathcal{J}_{\boldsymbol{\xi}}^{-1}\mathcal{I}_{\boldsymbol{\xi}}\mathcal{J}_{\boldsymbol{\xi}}^{-1}$. Lombardi and Veredas (2008) show that, despite the fact that we are estimating an elliptical stable distribution on estimated residuals, denoted by $\mathbf{X}(\hat{\boldsymbol{\xi}})$, the EMM point estimates, under the D regularity conditions, are not affected.

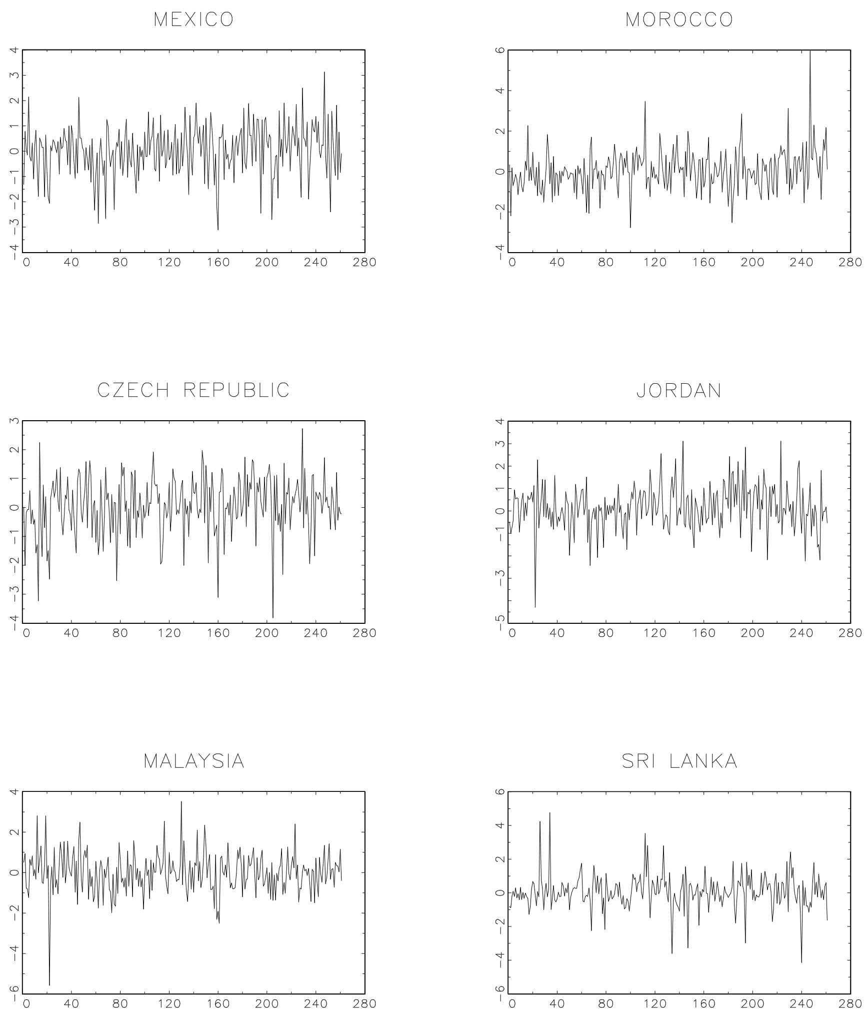

Figure 3: Standardized GARCH(1,1) residuals .

Figure 3 shows the standardized residuals for the same indexes as in the previous figure. The volatility clustering has disappeared, as it is clearly visible for the Czech Republic and Jordan. However, they do not appear to be Gaussian. The kurtosis coefficients range from 3.56 for Mexico to 7.93 for Sri Lanka, meaning that a fat-tailed distribution can be an appropriate choice. Figure 4 shows a heat map of the empirical correlations of the standardized residuals. Clusters by geographical areas are very clear. For instance Eastern European and Latin American countries are very related. Others, like Israel (the second from the upper right) and, surprisingly, China (the fourth from the left bottom), are not correlated at all with any other country.

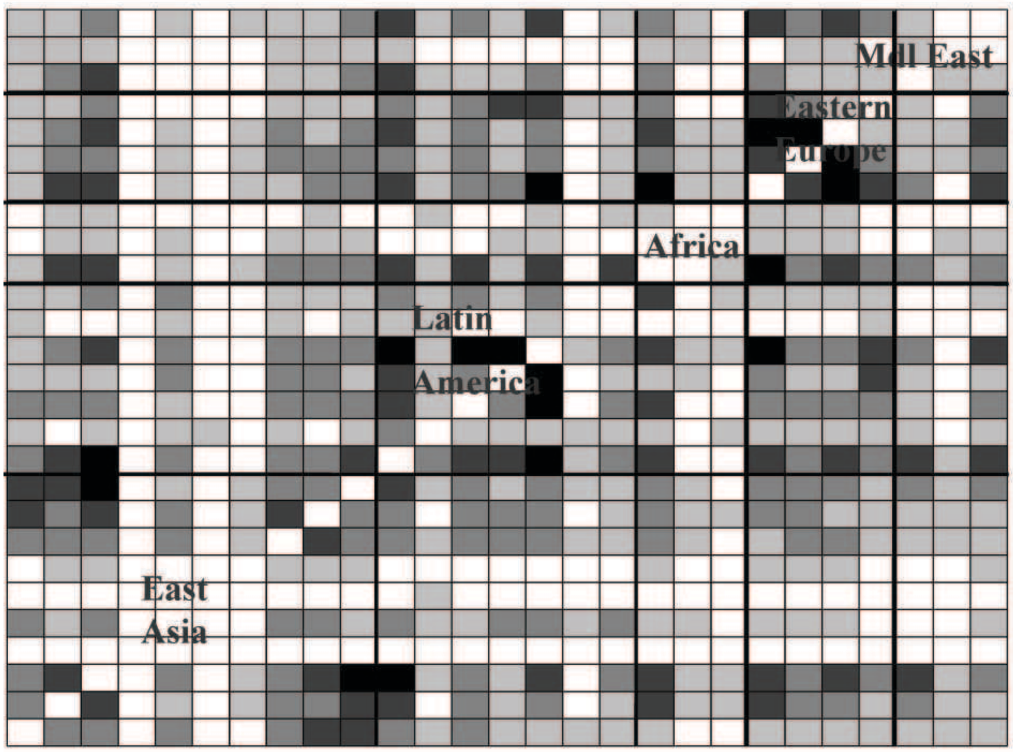

Figure 4: Heat map of the empirical correlations of standardized residuals. The darker (lighter) the higher (lower) the empirical correlations. Black corresponds to correlations around 0.6 and white close to 0. For representation purposes, main diagonal has been replaced by zeros, and hence the white.

Admittedly, this application has a number of drawbacks. First, the tail index α is the same for all countries. We estimated univariate stable distributions for each country and, as expected, the tail indexes are not constant across countries. They vary from 1.5 to 2. Yet this shortcoming is in fact applicable to any multivariate distribution like Student-t, skewed-t or Gaussian (for which the tail index is fixed at 2). Second, data are skewed yet we do not allow for asymmetries. Last we consider constant correlation, which may not be the case for the MSCI indexes. Despite all these shortcomings, this estimation exercise is purely illustrative and an application in a dynamic and asymmetric context is beyond the scope of the paper; though it would be an interesting research avenue, as explained in the conclusions.

The estimated tail index is 1.75, implying thicker tails than in the Gaussian case. The estimated degree of freedom for Student's t is 7.19 , which produces a mismatch in the existence of moments with respect to the stable distribution. Figure 5 shows the estimates for the location vector. They all vary around zero, which makes sense as the standardized residuals have mean equal to zero. As for the correlation matrix, we do not show all the results for reasons of space.[13] Instead, we present the estimated covariations for the six countries considered above,

[13]Detailed results are available upon request.

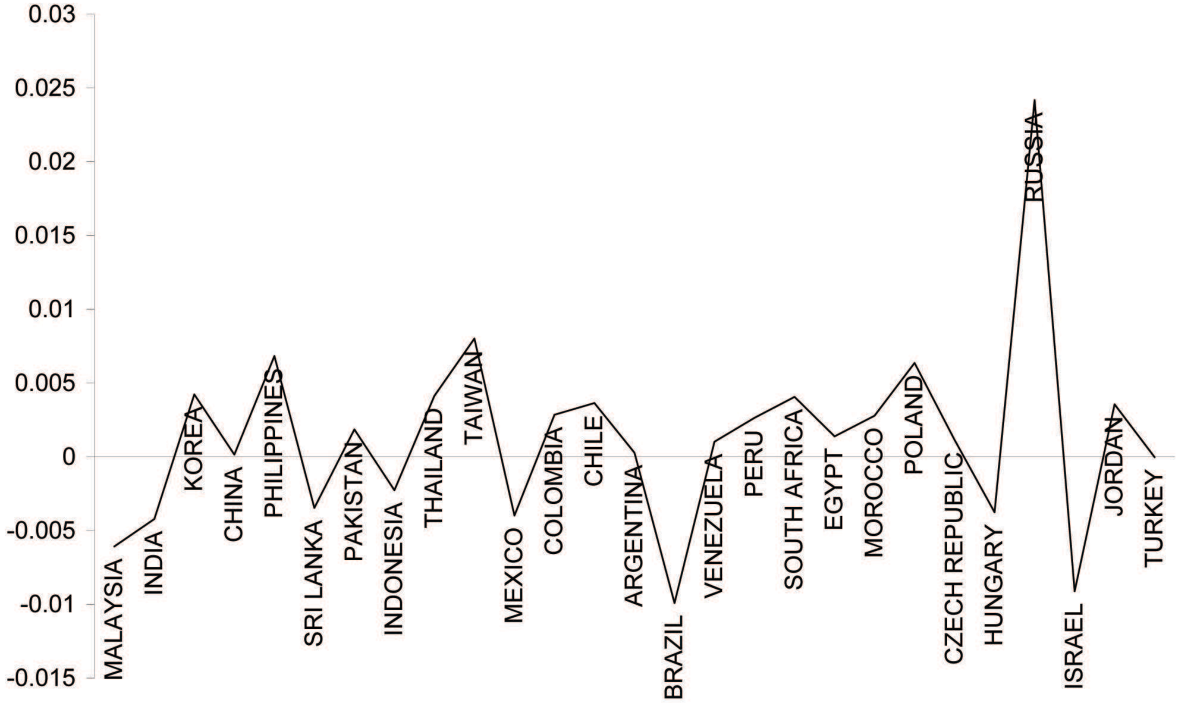

Figure 5: Estimated location parameters

which are the solid lines in Figure 6, while the dotted lines are the empirical correlations. The estimated covariations are very closed to the sample correlations in all cases. We may compare then with the estimated correlations matrices of Student-t, plotted in Figure 7. Clearly the estimated correlation matrix of Student-t distribution has a worst fit than that of the stable distribution. In fact, one can observe that, for Student-t, they are higher than one, which is not sensible.

6 Conclusions

In this paper we propose indirect estimation methods for multivariate elliptical stable distributions. Theoretical results prove that Student-t distribution is an adequate auxiliary model and standard asymptotics of indirect methods apply; a Monte Carlo study shows that even in small samples the estimator performs reasonably well. Finally, an empirical study further illustrates the proposed method.

Further research can take several directions. First, an obvious generalization is to allow for skewness. The reader may be tempted to extend the indirect methods to asymmetric stable distributions. However, we do not think that this will be a fruitful path as in the asymmetric case, skewness and scatter are merged into the so-called spectral measure, which takes very complicated forms. Furthermore, simulation from an asymmetric multivariate stable distributions turns out to be difficult. This of course hinders the use of indirect estimation methods, which were designed for situations in which simulating from the model of interest is straightforward.[14] An alternative is to use a recent method proposed by Nolan (2005), which

[14]Nonetheless, it is worth noticing that simulation of multivariate stable distributions is possible, cf. Modarres and Nolan (1994).

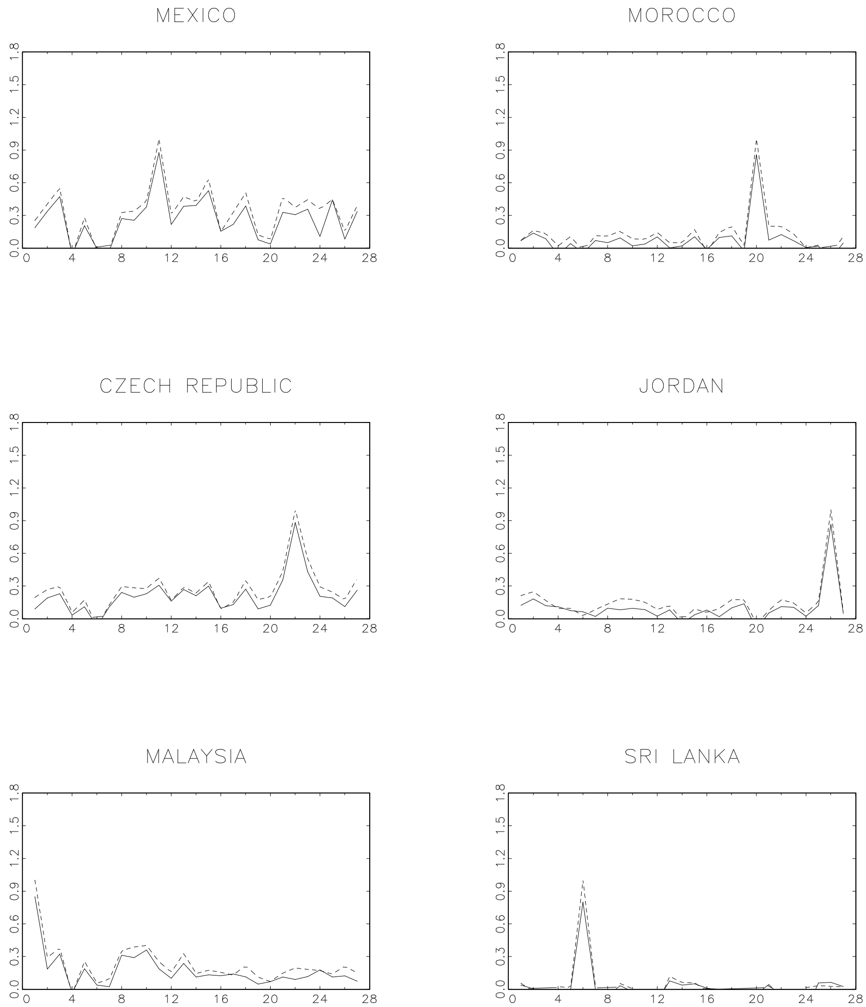

Figure 6: Results for the stable distribution. Estimated covariations (solid line) and empirical correlations (dotted line) of standardized residuals for six countries.

is based on projections parameter functions. Another alternative is the use of generalized elliptical distributions (cf. Frahm, 2004), which share many of the properties of the elliptical distributions.

Another direction for further research is the extension to a time series context. In particular, allow the location vector and the scatter matrix to be time varying. For instance VAR and multivariate GARCH types models are natural choices. On these grounds, one should be careful of the way the VAR and GARCH models are defined, as the inexistance of moments of orders higher that α entails some difficulties. For instance, univariate GARCH models under stable distribution have been analyzed by Mittnik, Paolella and Rachev (2002). This extension seems appropriate, given that most of the economic processes are time dependent. Furthermore, from a theoretical perspective it is feasible, since Gourieroux, Monfort and Renault (1993) and Calzolari, Fiorentini and Sentana (2004) do not assume *i.i.d.* returns in their analysis.

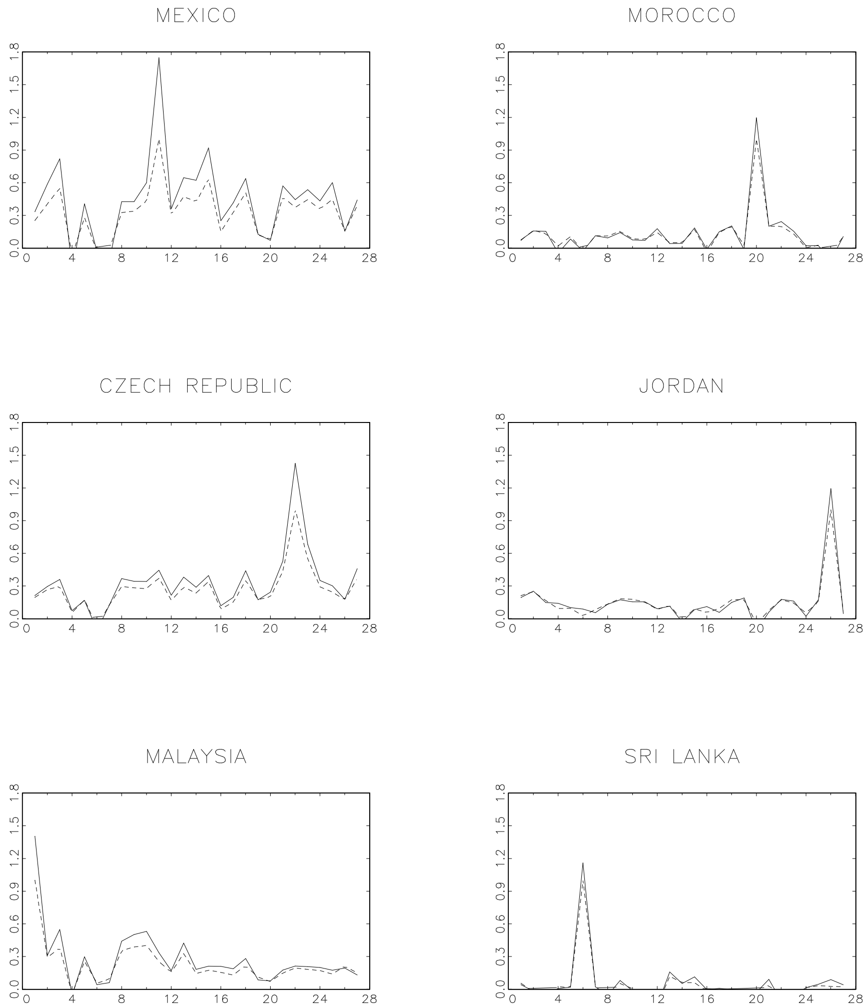

Figure 7: Results for the Student-t distribution. Estimated covariations (solid line) and empirical correlations (dotted line) of standardized residuals for six countries.

A Assumptions

C1. $\mathbf{X}$ is strictly stationary and ergodic.

C2. $\ln \tilde{\ell}(\boldsymbol{\zeta}, \mathbf{x})$ is twice continuously differentiable with respect to $\boldsymbol{\zeta}$.

C3. $E_{\boldsymbol{\theta}}[\ln \tilde{\ell}(\boldsymbol{\zeta}, \mathbf{x})]$ is twice continuously differentiable with respect to $\boldsymbol{\theta}$ and $\boldsymbol{\zeta}$ and has an unique maximum at $\boldsymbol{\theta} = \boldsymbol{\theta}_0$.

C4. $\mathbf{b}(\boldsymbol{\theta})$ is unique for $\boldsymbol{\theta}$.

C5.

$$\frac{\partial^2 \ln \tilde{\ell}(\boldsymbol{\zeta}, \mathbf{x})}{\partial \boldsymbol{\zeta} \partial \boldsymbol{\zeta}'} - \mathcal{H}_0 = o_p(1)$$

$$\sqrt{T} E_{\boldsymbol{\theta}}[\ln \tilde{\ell}(\boldsymbol{\zeta}, \mathbf{x})] \to^d \mathcal{N}(0, \mathcal{I}_0)$$

C6. The asymptotic covariance between the gradients of two units s_1 and s_2 of the simulated sample is constant.

D1. $\mathbf{X}(\hat{\boldsymbol{\xi}})$ is strictly stationary and ergodic.

D2. $\ln \tilde{\ell}(\boldsymbol{\zeta}, \mathbf{x}(\hat{\boldsymbol{\xi}}))$ is twice differentiable with respect to $\boldsymbol{\zeta}$ and $\hat{\boldsymbol{\xi}}$.

D3. $E_{\boldsymbol{\theta}}[\ln \tilde{\ell}(\boldsymbol{\zeta}, \mathbf{x}(\hat{\boldsymbol{\xi}}))]$ is twice continuously differentiable with respect to $\boldsymbol{\theta}$ and $\boldsymbol{\zeta}$ and $\hat{\boldsymbol{\xi}}$, and has an unique maximum at $\boldsymbol{\theta} = \boldsymbol{\theta}_0$.

D4. $\mathbf{b}(\boldsymbol{\theta}, \hat{\boldsymbol{\xi}})$ is unique for $\boldsymbol{\theta}$.

D5.

$$\frac{\partial^2 \ln \tilde{\ell}(\boldsymbol{\zeta}, \mathbf{x}(\hat{\boldsymbol{\xi}}))}{\partial \boldsymbol{\zeta} \partial \boldsymbol{\zeta}'} - \mathcal{H}_0 = o_p(1)$$
$$\sqrt{T} E_{\boldsymbol{\theta}}[\ln \tilde{\ell}(\boldsymbol{\zeta}, \mathbf{x}(\hat{\boldsymbol{\xi}}))] \to^d \mathcal{N}(0, \mathcal{I}_0 - \boldsymbol{\Xi})$$

where

$$\begin{aligned}
\boldsymbol{\Xi} &= \mathbf{B}\mathcal{J}_{\boldsymbol{\xi}}^{-1}\mathcal{I}_{\boldsymbol{\xi}}\mathcal{J}_{\boldsymbol{\xi}}^{-1}\mathbf{B}' + \mathbf{G}\mathbf{B}' + \mathbf{B}\mathbf{G}' \\
\mathbf{B} &= \lim_{T\to\infty} E\left(\frac{\partial^2 \ln \tilde{\ell}(\boldsymbol{\zeta}, \mathbf{x}(\hat{\boldsymbol{\xi}}))}{\partial \boldsymbol{\zeta} \partial \hat{\boldsymbol{\xi}}'}\right), \\
\mathbf{G} &= E\left(\frac{\partial \ln \tilde{\ell}(\boldsymbol{\zeta}, \mathbf{x}(\hat{\boldsymbol{\xi}}))}{\hat{\boldsymbol{\zeta}}}\hat{\boldsymbol{\xi}}\right)
\end{aligned}$$

D6. The asymptotic covariance between the gradients of two units s_1 and s_2 of the simulated sample is constant.

References

[1] Azzalini, A. and Capitanio, A. (2003), "Distributions generated by permutation of symmetry with emphasys on a multivariate skew t distribution", *Journal of the Royal Statistical Society B*, 64, 367-389.

[2] Bauwens, L., Laurent, S. and Rombouts, J. (2006), "Multivariate GARCH: A survey", *Journal of Applied Econometrics*, 21/1, 79-109.

[3] Bauwens, L. and Laurent, S. (2005), "A New Class of Multivariate Skew Densities with Application to GARCH models", *Journal of Business and Economic Statistics*, 23/3, 346-354.

[4] Bradley, B.O. and Taqqu, M.S. (2001), "Financial Risk and Heavy Tails", in Heavy tailed distributions in finance. Rachev S.T. ed.: Noth Holland.

[5] Calzolari, G., Fiorentini, G. and Sentana, E. (2004), "Constrained Indirect Estimation", *Review of Economic Studies* 71(4), 945-973.

[6] Cambais, S., Huang, S. and Simons, G. (1981) "On the theory of elliptically contoured distributions", *Journal of Multivariate Analysis* 11, 368-385.

[7] Chambers, J.M., Mallows, C.L. and Stuck, B.W. (1976), "A Method for Simulating Stable Random Variables", *Journal of the American Statistical Association*, 71, 340-344. Corrections 82 (1987): 704, 83 (1988): 581.

[8] Cheng, B.N., Rachev, S.T. (1995) “Multivariate stable future prices”, *Mathematical Finance* 5, 133-153.

[9] Demarta, S. and McNeil, A.J. (2005), ”The t copula and related copulas”, *International Statistical Review* 73: pp. 111–129.

[10] Fama, E. (1965a), “Behavior of Stock Market Prices”, *Journal of Business,* 38, 34-105.

[11] Fama, E. (1965b), “Portfolio Analysis in a Stable Paretian Market”, *Management Science,* 11, 404-419.

[12] Fang K.T., Kotz, S. and Ng, K.W. (1990), *Symmetric multivariate and related distributions*. New York: Chapman and Hall.

[13] Feldheim, E. (1937), *Etude de la stabilite des lois de probabilite*. PhD Thesis, These de la Faculte des Sciences de Paris.

[14] Frahm, G. (2004), *Generalized Elliptical Distributions*. PhD thesis, University of Cologne.

[15] Frahm, G. (2006), ”On the Extremal Dependence Coeficient of Multivariate Distributions”. *Statistics and Probability* Letters 76: pp. 1470–1481.

[16] Frahm, G., Junker, M., and Szimayer, A. (2003), ”Elliptical copulas: applicability and limitations”, *Statistics and Probability Letters* 63: pp. 275–286.

[17] Gallant, A.R. and Tauchen, G. (1996), “Which moments to match?”, *Econometric Theory*, 12, 657-681.

[18] Garcia, R., Renault, E. and Veredas, D. (2006), “Estimation of Stable Distributions by Indirect Inference”, mimeo.

[19] Ghose, D. and Kroner, K.F. (1995). “The relationship between GARCH and symmetric stable processes: Finding the source of fat tails in financial data”, *Journal of Empirical Finance*, vol 2, 225-251.

[20] Gouriéroux, C. Monfort, A. and Renault, E. (1993). “Indirect Inference”, *Journal of Applied Econometrics*, vol 8, 85-118.

[21] Ibragimov, I.A. and Linnik Y.V. (1971), *Independent and Stationary Secuences of Random Variables*, The Netherlands: Wolters-Noordhoff Publishing.

[22] Kelker D. (1970), “Distribution theory of elliptical distributions and a location-scale parameter generalization”, *Sankhya* A, 32, 419-430.

[23] Lombardi, M. J., Calzolari, G. (2004), “Indirect Inference for α-Stable Distributions and Processes ”, *Econometrics Journal*, in press.

[24] Lombardi, M. J., Veredas, D. (2007), Indirect Estimation of Elliptical Stable Distributions, CORE DP 2007/18.

[25] Lombardi, M. J., Veredas, D. (2008), Two-stage Indirect Estimation, manuscript.

[26] Mandelbrot, B. (1963), “The Variation of Certain Speculative Prices”, *Journal of Business*, 36, 394-419.

[27] McCulloch, J.H. (2003), “The Risk Neutral Measure and Option Pricing under log-Stable Uncertainty”, mimeo.

[28] Meng, B., van Dijk, H.K. (1995), “The EM algorithm”, *Journal of the Royal Statistical Society B*, 59, 511-567.

[29] Mittnik S., Paolella Marc, Rachev S. (2002), “Stationarity of Stable Power-GARCH Processes”, *Journal of Econometrics* 106.

[30] Modarres, R. and Nolan, J.P. (1994), "A method for simulating stable random vectors", *Computational Statistics* 9, 11-19.

[31] Nolan, J.P. (2005), "Multivariate stable densities and distribution functions general and elliptical case", mimeo.

[32] Nolan, J.P., Panorska, A.K. and McCulloch, J.H. (2001) "Estimation of stable spectral measures", *Mathametical and Computer Modelling* 34, 1113-1122.

[33] Ortobelli, S., Huber, I., Schwartz, E. (2002), "Portfolio Selection with Stable Distributed Returns", *Mathematical Methods of Operations Research*, 55, 265-300.

[34] Ortobelli, S. and Rachev, S.T. (2005), "Risk Management and dynamic portfolio selection with stable paretian", mimeo.

[35] Patton, A. (2004), "On the out-of-sample importance of skewness and asymmetric dependence for asset allocation", *Journal of Financial Econometrics*, 2(1) 130-168.

[36] Pivato, M. and Seco, L. (2003) "Estimating the spectral measure of a multivariate stable distribution via spherical harmonic analysis", *Journal of Multivariate Analysis* 87, 219-240.

[37] Rachev, S.T. and Xin, H. (1993), "Test for association of random variables in the domain of attraction of multivariate stable law", *Probability Mathematical Statistics*, 14, 125-141.

[38] Samorodnitsky, G. and Taqqu, M.S. (1994), *Stable Non-Gaussian Random Processes*. New York: Chapman and Hall.

[39] Smith, A.A. (1993), "Estimating Time Series Models Using Simulated Vector Autoregressions", *Journal of Applied Econometrics*, 8, 563-584.

Table 1: Simulation Results with $k = 2$.

		$S = 1$			$S = 5$		
α	True	1.7	1.9	1.95	1.7	1.9	1.95
	Mean	1.7267	1.8648	1.9304	1.7719	1.9196	1.9492
	Median	1.7211	1.8810	1.9439	1.7843	1.9307	1.9561
	RMSE	0.0753	0.0644	0.0810	0.1395	0.0517	0.0330
	MAE	0.0604	0.0329	0.0192	0.0630	0.0443	0.0217
	Asymp Coverage	0.7853	0.4749	0.4427	0.0402	0.0091	0.0103
	MC Coverage	0.9294	0.9498	0.9402	0.7488	0.9848	0.9349
σ_{11}	True	0.5	0.5	0.5	0.5	0.5	0.5
	Mean	0.8609	0.5866	0.5214	0.7556	0.5854	0.5851
	Median	0.5001	0.4997	0.5012	0.7737	0.5976	0.6051
	RMSE	6.8028	1.4322	0.1161	0.1304	0.2105	0.1893
	MAE	0.4481	0.1525	0.0518	0.0872	0.1784	0.1419
σ_{22}	True	2	2	2	2	2	2
	Mean	1.8488	2.0666	2.0505	2.0252	1.8069	1.9858
	Median	1.8744	2.0064	2.0046	2.0296	1.8200	2.0203
	RMSE	0.1773	0.2764	0.1672	0.4239	0.5826	0.5688
	MAE	0.1400	0.1020	0.0706	0.2411	0.5091	0.4189
σ_{12}	True	0.9	0.9	0.9	0.9	0.9	0.9
	Mean	0.6997	0.9061	0.9187	0.7575	0.7482	0.8450
	Median	0.8063	0.9027	0.9009	0.7805	0.7433	0.8538
	RMSE	1.9783	0.3055	0.1051	0.5037	0.2293	0.2373
	MAE	0.1695	0.0763	0.0462	0.1246	0.1992	0.1741
μ_1	True	0	0	0 0	0	0	
	Mean	-0.0114	0.0202	0.0445	0.1689	0.0669	0.0525
	Median	0.0107	0.0089	0.0050	0.0542	0.0558	0.0210
	RMSE	0.6816	0.3008	0.1182	1.5248	0.1292	0.1341
	MAE	0.0904	0.0698	0.0615	0.1791	0.1018	0.0930
μ_2	True	0	0	0	0	0	0
	Mean	0.0830	0.0706	0.0811	0.0179	0.1146	0.0778
	Median	0.0066	0.0025	0.0002	0.0763	0.0819	0.0106
	RMSE	0.7047	0.2172	0.2559	1.0416	0.2743	0.2877
	MAE	0.1472	0.1126	0.1231	0.2070	0.2132	0.1934

Table 2: Simulation Results with $k = 5$.

		$S = 1$			$S = 5$		
α	True	1.7	1.9	1.95	1.7	1.9	1.95
	Mean	1.6840	1.8775	1.9412	1.6892	1.8840	1.9346
	Median	1.6917	1.8913	1.9438	1.6958	1.8928	1.9465
	RMSE	0.0930	0.1381	0.0254	0.0850	0.0961	0.1141
	MAE	0.0306	0.0270	0.0073	0.0276	0.0214	0.0162
	Asymp Coverage	0.4512	0.2543	0.2231	0.3541	0.2032	0.1011
	MC Coverage	0.9728	0.9249	0.9782	0.9687	0.9394	0.9574
σ_{11}	True	0.25	0.25	0.25	0.25	0.25	0.25
	Mean	0.5893	0.4852	0.2521	0.3081	0.5438	0.4590
	Median	0.2515	0.2517	0.2505	0.2505	0.2518	0.2502
	RMSE	5.2882	2.6483	0.0240	0.5655	3.1190	3.7549
	MAE	0.3446	0.2406	0.0091	0.0667	0.2984	0.2157
σ_{22}	True	0.5	0.5	0.5	0.5	0.5	0.5
	Mean	0.5749	0.6587	0.6096	0.5596	0.6767	0.5589
	Median	0.5127	0.5033	0.5054	0.5101	0.5028	0.5072
	RMSE	0.7932	2.6387	1.9218	0.5816	2.4458	0.6869
	MAE	0.0906	0.1911	0.1154	0.0764	0.2042	0.0670
σ_{33}	True	1	1	1	1	1	1
	Mean	1.0420	1.2141	1.1910	1.1107	1.3601	1.0627
	Median	1.0060	1.0033	1.0058	1.0012	1.0018	1.0033
	RMSE	0.4469	2.5589	2.6972	1.2229	3.2985	0.7168
	MAE	0.0784	0.2417	0.2031	0.1512	0.3918	0.0809
σ_{44}	True	2	2	2	2	2	2
	Mean	2.0471	2.0403	2.2014	2.2007	2.2349	2.0966
	Median	2.0112	2.0004	2.0014	2.0153	1.9995	2.0105
	RMSE	0.6089	0.8294	2.7130	1.8715	2.9724	1.0452
	MAE	0.0880	0.1085	0.2129	0.2245	0.3061	0.1017
σ_{55}	True	4	4	4	4	4	4
	Mean	4.1734	4.1306	4.1878	4.1257	4.2664	4.1699
	Median	4.0002	4.0009	4.0018	3.9994	4.0001	4.0003
	RMSE	1.8511	1.4261	2.6971	1.9573	3.1187	2.0141
	MAE	0.2078	0.1888	0.1971	0.1963	0.3070	0.1787
μ_1	True	0	0	0	0	0	0
	Mean	0.0010	-0.0681	0.0280	0.0100	-0.0351	-0.0298
	Median	0.0008	0.0036	0.0021	0.0018	0.0035	0.0015
	RMSE	0.4878	1.0461	0.2734	0.6206	0.4344	0.3987
	MAE	0.0775	0.1201	0.0333	0.0989	0.0656	0.0451
μ_2	True	0	0	0	0	0	0
	Mean	0.0264	0.0162	0.0091	0.0025	0.0191	0.0484
	Median	0.0027	0.0042	0.0008	0.0011	0.0029	0.0000
	RMSE	0.4172	0.5426	0.2756	0.5014	0.4220	0.9474
	MAE	0.0768	0.1048	0.0444	0.0805	0.0778	0.0812
μ_3	True	0	0	0	0	0	0
	Mean	-0.0015	0.0069	-0.0089	-0.0145	-0.0183	-0.0312
	Median	0.0005	0.0014	-0.0006	-0.0031	0.0007	-0.0012
	RMSE	0.4053	0.7578	0.2808	0.4953	0.5562	0.2764
	MAE	0.0844	0.1274	0.0454	0.0895	0.0919	0.0502
μ_4	True	0	0	0	0	0	0
	Mean	0.0660	0.0286	0.0124	-0.0228	0.0382	0.0155
	Median	0.0007	-0.0012	-0.0004	0.0069	-0.0012	0.0074
	RMSE	0.6611	1.3595	0.2742	0.7870	0.5531	0.1050
	MAE	0.0939	0.1626	0.0389	0.0747	0.0787	0.0353
μ_5	True	0	0	0	0	0	0
	Mean	-0.0201	-0.0911	-0.0125	0.0289	0.0209	-0.0198
	Median	-0.0004	0.0011	0.0004	-0.0052	0.0010	-0.0048
	RMSE	0.6233	1.1853	0.2730	0.5308	0.5586	0.0836
	MAE	0.0821	0.1371	0.0339	0.0485	0.0680	0.0265

Table 3: Simulation Results with $k = 5$, continued.

		$S = 1$			$S = 5$		
α	True	1.7	1.9	1.95	1.7	1.9	1.95
σ_{12}	True	0.25	0.25	0.25	0.25	0.25	0.25
	Mean	0.8001	0.3143	0.1996	0.3270	0.2646	0.2247
	Median	0.2484	0.2492	0.2501	0.2522	0.2499	0.2512
	RMSE	10.0136	0.6661	0.8108	1.3074	0.8857	1.1987
	MAE	0.6000	0.1001	0.0648	0.1634	0.1296	0.0993
σ_{13}	True	0.4	0.4	0.4	0.4	0.4	0.4
	Mean	0.6825	0.3861	0.4636	0.4344	0.4650	0.5358
	Median	0.3993	0.4009	0.4030	0.3995	0.3997	0.4018
	RMSE	4.8737	0.8970	0.8141	0.6182	1.2563	2.2305
	MAE	0.3462	0.1187	0.0694	0.0944	0.1679	0.1553
σ_{14}	True	0	0	0	0	0	0
	Mean	-0.8002	-0.0028	-0.0011	0.0041	0.0213	-0.0293
	Median	-0.0037	-0.0027	-0.0008	-0.0003	-0.0018	-0.0019
	RMSE	0.8829	1.0609	0.8141	0.5449	1.0631	1.1364
	MAE	0.0905	0.1250	0.0677	0.0809	0.1359	0.1111
σ_{15}	True	0	0	0	0	0	0
	Mean	0.0992	0.0344	-0.0651	0.0009	-0.0244	-0.0245
	Median	-0.0069	-0.0070	-0.0038	0.0016	-0.0058	0.0036
	RMSE	1.5185	0.8649	0.8143	0.3018	1.1487	1.7979
	MAE	0.1523	0.1207	0.0689	0.0850	0.1547	0.1649
σ_{23}	True	0.4	0.4	0.4	0.4	0.4	0.4
	Mean	0.0656	0.3768	0.4673	0.3916	0.3545	0.5555
	Median	0.3891	0.4004	0.4041	0.4026	0.4014	0.4061
	RMSE	5.6644	1.1492	0.8196	1.3310	1.2731	1.9799
	MAE	0.3831	0.1717	0.0744	0.1783	0.1728	0.1598
σ_{24}	True	0	0	0	0	0	0
	Mean	0.0277	0.0171	-0.0467	0.2356	0.0139	0.0022
	Median	-0.0074	-0.0021	-0.0006	0.0026	-0.0019	-0.0002
	RMSE	0.9995	0.6997	0.8201	4.5781	1.1107	0.2963
	MAE	0.1190	0.0950	0.0735	0.2979	0.1392	0.0390
σ_{25}	True	0	0	0	0	0	0
	Mean	-0.1053	-0.0261	-0.0148	0.2958	-0.1193	0.0045
	Median	-0.0198	-0.0141	-0.0053	0.0029	-0.0142	0.0019
	RMSE	1.5437	0.5355	0.8185	6.5364	1.3771	0.2672
	MAE	0.1596	0.0885	0.0730	0.4116	0.1546	0.0359
σ_{34}	True	0	0	0	0	0	0
	Mean	-0.0285	-0.1261	-0.0012	-0.1392	0.0254	-0.0259
	Median	-0.0009	-0.0016	-0.0005	-0.0011	-0.0007	0.0009
	RMSE	0.5776	1.2267	0.8137	2.7696	1.3337	0.4995
	MAE	0.0958	0.1443	0.0705	0.1989	0.1661	0.0480
σ_{35}	True	0	0	0	0	0	0
	Mean	-0.0171	0.0034	0.0438	-0.1806	-0.1686	-0.0533
	Median	-0.0037	-0.0024	0.0001	-0.0026	-0.0011	-0.0029
	RMSE	0.6957	1.0714	0.8185	3.4287	2.6061	0.7032
	MAE	0.1103	0.1679	0.0762	0.2532	0.3127	0.0679
σ_{45}	True	2.55	2.55	2.55	2.55	2.55	2.55
	Mean	2.5206	2.3702	2.5084	2.6777	2.6605	2.6521
	Median	2.5407	2.5481	2.5497	2.5517	2.5509	2.5570
	RMSE	1.4091	3.7083	0.8346	1.9161	3.1640	1.0382
	MAE	0.1827	0.2960	0.0763	0.1953	0.3505	0.1059