The Method of Simulated Quantiles

Yves Dominicy* and David Veredas †

First draft: December 2009 - This version: May 2010

Abstract

We introduce an inference method based on quantiles matching, which is useful for situations where the density function does not have a closed form –but it is simple to simulate– and/or moments do not exist. Functions of theoretical quantiles, which depend on the parameters of the assumed probability law, are matched with sample quantiles, which depend on observations. Since the theoretical quantiles may not be available analytically, the optimization is based on simulations. We illustrate the method with the estimation of α-stable distributions. A thorough Monte Carlo study and an illustration to 22 financial indexes show the usefulness of the method.

Keywords: Quantiles, simulated methods, α-stable distribution, fat tails.

JEL Classification: C32, G14, E44

*Université libre de Bruxelles, Solvay Brussels School of Economics and Management, ECARES. 50, Av Roosevelt CP114, B1050 Brussels, Belgium. Tel: +32(0)26504502; Fax: +3226504475; `yves.dominicy@ulb.ac.be`

†Université libre de Bruxelles, Solvay Brussels School of Economics and Management, ECARES. 50, Av Roosevelt CP114, B1050 Brussels, Belgium. Tel: +32(0)26504218; Fax: +32(0)26504475; `david.veredas@ulb.ac.be`

We are grateful to Rob Engle, Marcelo Fernandes, Marc Hallin, J. Huston McCulloch, and Marc Paolella, as well as the audiences at the conferences in ”recent developments in time series” (Rennes, France), ”statistical inference in multivariate models and time-series models” (Kagoshima, Japan), ”latest developments in heavy-tailed distributions” (Brussels, Belgium), the ”mathematical finance days” (Montreal, Canada), the participants at the Quantitative Financial Econometrics Seminar (Volatility Institute - NYU Stern), the ECORE-KUL seminar, and the doctoral course on ”fat-tailed distributions” in the University of Zurich for insightful remarks. The authors gratefully acknowledge financial support from the Belgian National Bank and the IAP P6/07 contract, from the IAP programme (Belgian Scientific Policy), 'Economic policy and finance in the global economy'. The authors are members of ECORE, the recently created association between CORE and ECARES. Any remaining errors and inaccuracies are ours.

1 Introduction

Estimation of the parameters of an econometric or economic parametric model is a first order concern. In the case that we know the probability law that governs the random variables, Maximum Likelihood (ML henceforth) is the benchmark technique. If we relax the assumption of knowledge of the distribution but we still have knowledge of the moments, the Generalized Method of Moments (GMM henceforth) is the benchmark technique. However, there are economic and econometric models that cannot be easily estimated with ML or GMM: stochastic volatilities, models with stochastic regimes switches, or involving expected utilities to name a few.

To circumvent these estimation difficulties, numerous estimation methods based on simulations have been developed. Gouriéroux and Monfort (1996) and Hajivassiliou and Ruud (1994) introduce Simulated ML (SML), similar to ML except that simulated probabilities are used instead of the exact probabilities. McFadden (1989), Pakes and Polland (1989), and Duffie and Singleton (1989) independently introduced the Method of Simulated Moments (MSM), which is based on matching sample moments and theoretical moments that are generated by simulations. Gouriéroux, Monfort and Renault (1993) propose Indirect Inference (IndInf), a method that is based on estimating indirectly the parameters of the model of interest through an auxiliary model. The Efficient Method of Moments (EMM) of Gallant and Tauchen (1996) is based on the same idea.

In this article we introduce the Method of Simulated Quantiles (MSQ henceforth). MSQ is appropriate when standard estimation methods fail because of lack of closed form expression of the probability density function or/and lack of existence of moments. Indeed, since it is based on quantiles, it is a moment-free method. And since it is based on simulations, we do not need closed form expressions of any function that represents the probability law of the process. Also, due to the robustness of quantiles, MSQ is appropriate when data show unusually large observations that do not follow the same process as the rest of observations.[1] In a nutshell, MSQ is based on a vector of functions of quantiles. These functions can be either computed from data (the sample functions) or from the distribution (the theoretical functions). The estimated parameters are those that minimize a quadratic distance between both. Since the theoretical functions of quantiles may not have a closed form expression, we rely on simulations. Throughout the article, we illustrate the method with the estimation of α-stable distributions.

Example

Let X_t be a random variable distributed with an α-stable distribution that is characterized by four parameters –α, β, σ and μ– and is represented as $X_t \sim S_\alpha(\sigma, \beta, \mu)$. The parameter $\alpha \in (0, 2]$,

[1] The above cited inference methods have been robustified. Genton and Ronchetti (2003), Ortelli and Trojani (2005) and Czellar et al. (2007) propose the robust IndInf, robust EMM and Indirect robust GMM respectively.

often denoted as tail index, measures the thickness of the tails and governs the existence of moments: $E[X_t^p] < \infty$, $\forall p < \alpha$. Asymmetry is captured by $\beta \in [-1, 1]$. For $\beta = 1$ the distribution is completely right-asymmetric and for $\beta = -1$ the distribution is completely left-asymmetric. The dispersion parameter $\sigma \in \mathbb{R}^+$ expands or contracts the distribution, and the location parameter $\mu \in \mathbb{R}$ controls the location of the distribution. As $\alpha \to 2$, β loses importance and it becomes unidentified. For $\alpha = 2$, the distribution becomes Gaussian and β is irrelevant. Other particular cases of the α-stable distribution are the Cauchy ($\alpha = 1$ and $\beta = 0$), and the Lévy distributions ($\alpha = 0.5$ and $\beta = \pm 1$). The α-stable distributions posses the property of stability: Linear combinations of $i.i.d.$ α-stable random variables with the same α are also α-stable distributed. A related property of the α-stable distribution is that it is a domain of attraction –this is often known as the Generalized Central Limit Theorem– which states that the only distribution that arises as limit from sums of $i.i.d.$ random variables (suitably scaled and centered) is the α-stable distribution.

The pdf of the α-stable distribution does not have a closed form. Since it is a complicated integral, even difficult to evaluate numerically, estimation by ML has been often not considered in applied work (though the theoretical properties of the ML estimator exist, Dumouchel, 1973, and the actual estimation has been performed by Nolan, 2001). However, the characteristic function (CF hereafter) has a manageable closed form:

$$E[e^{i\theta X_t}] = \begin{cases} \exp\{-\sigma^\alpha |\theta|^\alpha (1 - i\beta(\text{sign}\theta) \tan \frac{\pi\alpha}{2}) + i\mu\theta\} & \text{if} \quad \alpha \neq 1 \\ \exp\{-\sigma|\theta|(1 + i\beta \frac{2}{\pi}(\text{sign}\theta) \ln|\theta|) + i\mu\theta\} & \text{if} \quad \alpha = 1. \end{cases}$$

All the methods based on the CF match the theoretical and sample counterparts, but in different ways. Since the sample CF is a random variable with complex values, one can think about comparing (i) moments associated to real and imaginary components respectively (Press, 1972, Fielitz and Rozelle, 1981), (ii) minimizing a distance between the sample and the theoretical CF functions (Paulson, Holcomb and Leitch, 1975, Feuerverger and McDunnough, 1981, and Carrasco and Florens, 2002), (iii) performing a regression analysis between the real and imaginary parts of the sample and theoretical CF (Koutrevelis, 1980), or (iv) using the fast Fourier transform to express the likelihood as a function of the CF (Chenyao, Mittnik and Doganoglu, 1999). A problem inherent on these methods is the choice of the grid of frequencies at which to evaluate the CF. While Fielitz and Rozelle (1981) recommend, on the basis of Monte Carlo results, to match only a few frequencies, others, like Feuerverger and McDunnough (1981), recommend to use as many frequencies as possible. However, in the latter case, Carrasco and Florens (2002) have shown that, even asymptotically, matching a continuum of moment conditions introduces a fundamental singularity problem.

An alternative is the use of simulation-based methods. Since random numbers from α-stable distributions can be obtained straightforwardly, simulation-based methods such as IndInf and EMM are appealing, as it has been shown by Garcia, Renault and Veredas (2008) and Lombardi and Calzolari (2008). They both use a skewed-t distribution as auxiliary model. This is a sensible choice, since the skewed-t distribution also has four parameters that measure the same features as the four parameters of the α-stable distribution.

Finally, Fama and Roll (1971) and McCulloch (1986) propose to use functions of quantiles. Four specific functions of quantiles are constructed to capture the same features as those captured by α, β, σ and μ. Since the pdf does not have a closed form, so do the cumulative density function and so do the quantiles. Estimation has to be done either by simulation or by tabulation. They opt for the later. Fama and Roll (1971) and McCulloch (1986) estimate the parameters by calibrating the value of the sample functions of quantiles with tabulated values of the theoretical quantiles. This is a fast way to estimate the parameters, since it avoids optimization, but the theoretical properties remain unclear and the extension to the case of linear combinations of α-stable random variables is not possible (since the tail index has to be the same for all the random variables, estimation has to be done jointly). $\square$

MSQ combines the simulation-based and the quantile-based methods. It is broader than Fama and Roll (1971) and McCulloch (1986) that were designed for the estimation of the α-stable distribution. In fact, MSQ is very general as it does not make any assumptions on the functional forms for the functions of quantiles. A second advantage is that the method is not based on tabulations but on simulations. This allows a larger flexibility and accuracy. Indeed, tabulation requires interpolation if the sample functions of quantiles are not exactly equal to the tabulated theoretical functions of quantiles. Third, we provide an asymptotic theory that shows the consistency, asymptotic normality and the asymptotic variance-covariance matrix of the estimated parameters.

Estimation via quantiles is a natural alternative to moment-based methods and tracks back to Aitchison and Brown (1957). In this book on the log-normal distribution, they estimate a three-parameter log-normal distribution by matching quantiles. Quantiles can also be used to construct functions that measure aspects of the probability distribution. Let q_τ denote the τ-th quantile of X_t for $\tau \in (0,1)$. The median, $q_{0.50}$, is often used as an estimator of the location. The interquartile range, $q_{0.75} - q_{0.25}$ is a natural measure of dispersion. Bowley (1920) proposed the quartile skewness (known as the Bowley coefficient):

$$BC = \frac{(q_{0.75} - q_{0.5}) - (q_{0.5} - q_{0.25})}{(q_{0.75} - q_{0.25})},$$

which was extended by Hinkley (1975):

$$Hink = \frac{(q_\tau - q_{0.5}) - (q_{0.5} - q_{1-\tau})}{(q_\tau - q_{1-\tau})}.$$

The smaller τ, the less sensitive to outliers, but the less information from the tails it uses. These measures of asymmetry are dispersion and location invariant, i.e. $\gamma(aX_t + b) = \gamma(X_t)$, where γ denotes one of the two above measures.

As far as measures for tail thickness are concerned, Crow and Siddiqui (1967) proposed a measure based on the ratio of two interquantile ranges:

$$CS = \frac{(q_{\tau_1} - q_{1-\tau_1})}{(q_{\tau_2} - q_{1-\tau_2})},$$

where $\tau_1 > \tau_2$. Their choices for τ_1 and τ_2 were 0.975 and 0.75 respectively. Alternatively, a measure of tail thickness can be interpreted as a measure of the dispersion of a distribution around $\mu \pm \sigma$. Based on this interpretation, Moors (1988) proposed

$$Mo = \frac{(q_{0.875} - q_{0.625}) + (q_{0.375} - q_{0.125})}{(q_{0.750} - q_{0.250})}.$$

The two terms in the numerator are large if little probability mass is concentrated in the neighborhood of the first and third quartile. Note that CS and Mo are standardized by the interquartile range to guarantee invariance under linear transformations.[2]

The rest of the paper is organized as follows. In Section 2 we first introduce notation followed by MSQ. Each step of the presentation of the method is illustrated with our example. We also show the assumptions and the asymptotic distribution of the estimators. In Section 3 we report the results of a Monte Carlo study based on our example. We consider univariate and multidimensional estimation of α-stable distributions. For the univariate case, our method is compared with McCulloch (1986). In Section 4 we show an illustration to 22 world-wide market indexes, assumed to be distributed according to α-stable distributions. We first estimate the parameters independently. Then we estimate them jointly assuming a common tail index, which is needed for the construction of linear combinations, as we show in the last part of the section. Section 5 concludes. Proofs and other technicalities are relegated to the Appendix.

2 The Method of Simulated Quantiles

We consider a $J \times 1$ strongly stationary random vector that follows a distribution $\mathcal{D}(\boldsymbol{\theta})$, where $\boldsymbol{\theta}$ denotes the vector of unknown parameters that are in an interior point of the compact parameter set $\Theta \subset \mathbb{R}^p$. The J elements of the random vector are independent random variables. The j-th random variable follows the marginal distribution $\mathcal{D}_j(\boldsymbol{\theta}_j)$, where $\boldsymbol{\theta}_j \in \Theta^j \subset \mathbb{R}^{p_j}$. We allow for common parameters between $\{\boldsymbol{\theta}_j\}_{j=1}^M$.

Let $x_{i,j}$ be the i-th realization of the j-th random variable and let $\mathbf{x}_j = (x_{1,j}, ..., x_{i,j}, ..., x_{N,j})^{\mathrm{T}}$ be the vector of N realizations. Denote by $\hat{\boldsymbol{q}}_j = (\hat{q}_{j,\tau_1}, ..., \hat{q}_{j,\tau_s})^{\mathrm{T}} \in \mathbb{R}^s$ and $\hat{\boldsymbol{q}}_j^* = (\hat{q}_{j,\tau_1}, ..., \hat{q}_{j,\tau_b})^{\mathrm{T}} \in \mathbb{R}^b$ two $s \times 1$ and $b \times 1$ vectors of sample quantiles of $\mathbf{x}_j$. That is, $\hat{q}_{j,\tau_k}$ denotes the τ_k-th sample quantile of $\mathbf{x}_j$. Let $\mathbf{h}(\hat{\boldsymbol{q}}_j)$ and $\mathbf{g}(\hat{\boldsymbol{q}}_j^*)$ be two $M \times 1$ vectors of continuous and once differentiable

[2] For more details about quantile based kurtosis measures see Groeneveld (1984), Groeneveld (1998) and Brys (2006).

functions $\mathbb{R}^s \to \mathbb{R}^M$ and $\mathbb{R}^b \to \mathbb{R}^M$. Consider their Hadamard (element-by-element) product $\hat{\boldsymbol{\phi}}_j = \mathbf{h}(\hat{\boldsymbol{q}}_j) \odot \mathbf{g}(\hat{\boldsymbol{q}}_j^*)$.

Likewise, denote by $\boldsymbol{q}_{\boldsymbol{\theta}_j} = (q_{\tau_1,\boldsymbol{\theta}_j}, ..., q_{\tau_s \boldsymbol{\theta}_j})^{\mathrm{T}} \in \mathbb{R}^s$ and $\boldsymbol{q}^*_{\boldsymbol{\theta}_j} = (q^*_{\tau_1,\boldsymbol{\theta}_j}, ..., q^*_{\tau_b,\boldsymbol{\theta}_j})^{\mathrm{T}} \in \mathbb{R}^b$ two $s \times 1$ and $b \times 1$ vectors of theoretical quantiles corresponding to $\mathcal{D}_j(\boldsymbol{\theta}_j)$. That is, $q_{\tau_k,\boldsymbol{\theta}_j}$ denotes the τ_k-th theoretical quantile of $\mathbf{x}_j$. These quantiles may not be available analytically and they may have to be computed through simulation. Let $\mathbf{h}(\boldsymbol{q}_{\boldsymbol{\theta}_j})$ and $\mathbf{g}(\boldsymbol{q}^*_{\boldsymbol{\theta}_j})$ be two $M \times 1$ vectors of continuous and once differentiable functions $\mathbb{R}^s \to \mathbb{R}^M$ and $\mathbb{R}^b \to \mathbb{R}^M$. Consider their Hadamard product $\boldsymbol{\phi}_{\boldsymbol{\theta}_i} = \mathbf{h}(\boldsymbol{q}_{\boldsymbol{\theta}_j}) \odot \mathbf{g}(\boldsymbol{q}^*_{\boldsymbol{\theta}_j})$.

Example (cont.)

For $J = 1$ McCulloch (1986) defines four functions of quantiles that represent the four parameters of the α-stable distribution. Let $\hat{\boldsymbol{q}} = (\hat{q}_{0.95}, \hat{q}_{0.75}, \hat{q}_{0.50}, \hat{q}_{0.25}, \hat{q}_{0.05})^{\mathrm{T}}$ and $\hat{\boldsymbol{q}}^* = (\hat{q}_{0.95}, \hat{q}_{0.75}, \hat{q}_{0.25}, \hat{q}_{0.05})^{\mathrm{T}}$.[3] The functions $\mathbf{h}(\hat{\boldsymbol{q}})$ and $\mathbf{g}(\hat{\boldsymbol{q}}^*)$ are both 4×1:

$$\begin{aligned}
\mathbf{h}(\hat{\boldsymbol{q}}) &= \begin{pmatrix} \hat{q}_{0.95} - \hat{q}_{0.05} \\ (\hat{q}_{0.95} - \hat{q}_{0.50}) + (\hat{q}_{0.05} - \hat{q}_{0.50}) \\ \hat{q}_{0.75} - \hat{q}_{0.25} \\ \hat{q}_{0.50} \end{pmatrix} \\
\mathbf{g}(\hat{\boldsymbol{q}}^*) &= \begin{pmatrix} (\hat{q}_{0.75} - \hat{q}_{0.25})^{-1} \\ (\hat{q}_{0.95} - \hat{q}_{0.05})^{-1} \\ 1 \\ 1 \end{pmatrix}.
\end{aligned}$$

And the vector of functions of quantiles $\hat{\boldsymbol{\phi}}$ is

$$\hat{\boldsymbol{\phi}} = \begin{pmatrix} \frac{\hat{q}_{0.95} - \hat{q}_{0.05}}{\hat{q}_{0.75} - \hat{q}_{0.25}} \\ \frac{(\hat{q}_{0.95} - \hat{q}_{0.50}) + (\hat{q}_{0.05} - \hat{q}_{0.50})}{\hat{q}_{0.95} - \hat{q}_{0.05}} \\ \hat{q}_{0.75} - \hat{q}_{0.25} \\ \hat{q}_{0.50} \end{pmatrix}.$$

The first two elements are essentially special cases of Crow and Siddiqui (1967) and Hinkley (1975) respectively. The last two are the interquantile range and the median.

The two upper elements in $\hat{\boldsymbol{\phi}}$ are dispersion and location invariant, meaning that they are insensitive to μ and σ. This is why McCulloch (1986) standardizes the sample. So if the process truly has unit dispersion, the theoretical and sample interquantile ranges should be the same: $q_{0.75,\boldsymbol{\theta}} - q_{0.25,\boldsymbol{\theta}} = \hat{q}_{0.75} - \hat{q}_{0.25}$. Otherwise, the theoretical interquantile range is re-scaled by σ: $(q_{0.75,\boldsymbol{\theta}} - q_{0.25,\boldsymbol{\theta}})\sigma = \hat{q}_{0.75} - \hat{q}_{0.25}$. Similarly, if the process truly has location zero and unit dispersion, the theoretical and

[3]To simplify notation, in the univariate setting we skip the j-th index.

sample medians should be the same: $q_{0.50,\boldsymbol{\theta}} = \hat{q}_{0.50}$. Otherwise, the theoretical median is re-scaled and re-located by σ and μ respectively: $\mu + \sigma q_{0.50,\boldsymbol{\theta}} = \hat{q}_{0.50}$.[4] Putting all these elements together, the vector of theoretical functions of quantiles equals:

$$\boldsymbol{\phi}_{\boldsymbol{\theta}} = \begin{pmatrix} \frac{q_{0.95,\boldsymbol{\theta}} - q_{0.05,\boldsymbol{\theta}}}{q_{0.75,\boldsymbol{\theta}} - q_{0.25,\boldsymbol{\theta}}} \\ \\ \frac{(q_{0.95,\boldsymbol{\theta}} - q_{0.50,\boldsymbol{\theta}}) + (q_{0.05,\boldsymbol{\theta}} - q_{0.50,\boldsymbol{\theta}})}{q_{0.95,\boldsymbol{\theta}} - q_{0.05,\boldsymbol{\theta}}} \\ \\ (q_{0.75,\boldsymbol{\theta}} - q_{0.25,\boldsymbol{\theta}})\sigma \\ \\ \mu + \sigma q_{0.50,\boldsymbol{\theta}} \end{pmatrix}.$$

□

Let $\hat{\boldsymbol{\Phi}} = (\hat{\boldsymbol{\phi}}_1^{\mathrm{T}}, ..., \hat{\boldsymbol{\phi}}_J^{\mathrm{T}})^{\mathrm{T}}$ and $\boldsymbol{\Phi}_{\boldsymbol{\theta}} = (\boldsymbol{\phi}_{\boldsymbol{\theta}_1}^{\mathrm{T}}, ..., \boldsymbol{\phi}_{\boldsymbol{\theta}_J}^{\mathrm{T}})^{\mathrm{T}}$ be two $JM \times 1$ vectors containing the sample and theoretical functions of quantiles for the J random variables. Identification requires that $JM \geq p$. The principle of MSQ is to find the value of the parameters that best match the sample and theoretical functions of quantiles. This is done by minimizing the quadratic distance between $\hat{\boldsymbol{\Phi}}$ and $\boldsymbol{\Phi}_{\boldsymbol{\theta}}$:

$$\hat{\boldsymbol{\theta}} = argmin_{\boldsymbol{\theta} \in \Theta} \, (\hat{\boldsymbol{\Phi}} - \boldsymbol{\Phi}_{\boldsymbol{\theta}})^{\mathrm{T}} \mathbf{W}_{\boldsymbol{\theta}} (\hat{\boldsymbol{\Phi}} - \boldsymbol{\Phi}_{\boldsymbol{\theta}}), \tag{1}$$

where $\mathbf{W}_{\boldsymbol{\theta}}$ is a $JM \times JM$ symmetric positive definite weighting matrix defining the metric. Three particular cases are nested in (1). The first is when no simulations are needed. If $\boldsymbol{\Phi}_{\boldsymbol{\theta}}$ can be computed explicitly, then (1) can be solved by the standard optimization techniques. An example is the Tukey lambda distribution (Joiner and Rosenblatt, 1971), which is only defined in terms of the quantiles. For the j-th random variable, the τ_k-th theoretical quantile is

$$q_{\tau_k,\theta_j} = \frac{\tau_k^{\theta_j} - (1-\tau_k)^{\theta_j}}{\theta_j},$$

for $\theta_j \neq 0$. The Tukey lambda distribution has its generalized version (Ramberg and Schmeiser, 1974), also only defined in terms of the quantiles:

$$q_{\tau_k,\boldsymbol{\theta}_j} = \theta_{j,1} + \frac{\tau_k^{\theta_{j,3}} - (1-\tau_k)^{\theta_{j,4}}}{\theta_{j,2}},$$

where $\theta_{j,1}$ and $\theta_{j,2}$ are location and dispersion parameters respectively, and $\theta_{j,3}$ and $\theta_{j,4}$ are shape parameters.

[4] McCulloch (1986) notices that $q_{0.50,\boldsymbol{\theta}}$ has a double singularity as α crosses 1 when $\beta \neq 0$. This makes interpolation meaningless between $\alpha = 0.9$ and $\alpha = 1.1$. To circumvent this problem, McCulloch (1986) uses $\sigma q_{0.50,\boldsymbol{\theta}} = \hat{q}_{0.50} - \zeta$ where, see Zolotarev (1957),

$$\zeta = \begin{cases} \mu + \beta\,\sigma \tan\frac{\pi\alpha}{2} & \text{for} \quad \alpha \neq 1 \\ \mu & \text{for} \quad \alpha = 1. \end{cases}$$

The second particular case is when $JM = p$. Then $\mathbf{W}_{\boldsymbol{\theta}}$ is irrelevant and the problem boils down to solving a system of nonlinear equations via simulations. This is the case of the univariate α-stable distribution where there are four parameters and four functions of quantiles.

The third particular case combines the first and second: if $\mathbf{\Phi}_{\boldsymbol{\theta}}$ can be computed explicitly and $JM = p$, the problem reduces to find the $\boldsymbol{\theta}$ such that $\hat{\mathbf{\Phi}} = \mathbf{\Phi}_{\boldsymbol{\theta}}$. This is the case of the Tukey lambda distribution, where there is only one parameter that can estimated with just one function of quantiles (which could be the τ_k-th quantile itself, i.e. matching $\hat{q}_{\tau_k}$ with q_{τ_k,θ_j}), or its generalized version with four parameters and four functions of quantiles.

Example (cont.)

As mentioned above, since $\boldsymbol{\phi}_{\boldsymbol{\theta}}$ and $\hat{\boldsymbol{\phi}}$ are 4×1 vectors, optimization boils down to solve the system $\boldsymbol{\phi}_{\boldsymbol{\theta}} = \hat{\boldsymbol{\phi}}$. Because the functions of quantiles for α and β are location and dispersion invariant, they can be matched independently from those for σ and μ. Let $\boldsymbol{\theta}^{\dagger} = (\alpha, \beta)$,

$$\boldsymbol{\phi}^{\dagger}_{\boldsymbol{\theta}^{\dagger}} = \begin{pmatrix} \frac{q_{0.95,\boldsymbol{\theta}^{\dagger}} - q_{0.05,\boldsymbol{\theta}^{\dagger}}}{q_{0.75,\boldsymbol{\theta}^{\dagger}} - q_{0.25,\boldsymbol{\theta}^{\dagger}}} \\ \\ \frac{(q_{0.95,\boldsymbol{\theta}^{\dagger}} - q_{0.50,\boldsymbol{\theta}^{\dagger}}) + (q_{0.05,\boldsymbol{\theta}^{\dagger}} - q_{0.50,\boldsymbol{\theta}^{\dagger}})}{q_{0.95,\boldsymbol{\theta}^{\dagger}} - q_{0.05,\boldsymbol{\theta}^{\dagger}}} \end{pmatrix},$$

and

$$\hat{\boldsymbol{\phi}}^{\dagger} = \begin{pmatrix} \frac{\hat{q}_{0.95} - \hat{q}_{0.05}}{\hat{q}_{0.75} - \hat{q}_{0.25}} \\ \\ \frac{(\hat{q}_{0.95} - \hat{q}_{0.50}) + (\hat{q}_{0.05} - \hat{q}_{0.50})}{\hat{q}_{0.95} - \hat{q}_{0.05}} \end{pmatrix}.$$

McCulloch (1986) finds the values, through tabulation, of α and β that match $\boldsymbol{\phi}^{\dagger}_{\boldsymbol{\theta}^{\dagger}}$ with $\hat{\boldsymbol{\phi}}^{\dagger}$. The functions of quantiles for α are bounded from below by 2.439 (this is the value that corresponds to $\alpha = 2$). However, due to the randomness inherent in finite samples, the estimated functions of quantiles may take values below 2.439. To avoid this finite sample artifact, McCulloch (1986) imposes the constraint that the function of quantiles for α equals 2.439 if the estimated one is below. Alternative solutions in our simulation-based framework are discussed below.

Once $\hat{\boldsymbol{\theta}}^{\dagger} = (\hat{\alpha}, \hat{\beta})$ are obtained, estimates for σ and μ are straightforwardly obtained. For σ:

$$\hat{\sigma} = \frac{\hat{q}_{0.75} - \hat{q}_{0.25}}{q_{0.75,\hat{\boldsymbol{\theta}}'} - q_{0.25,\hat{\boldsymbol{\theta}}'}}, \tag{2}$$

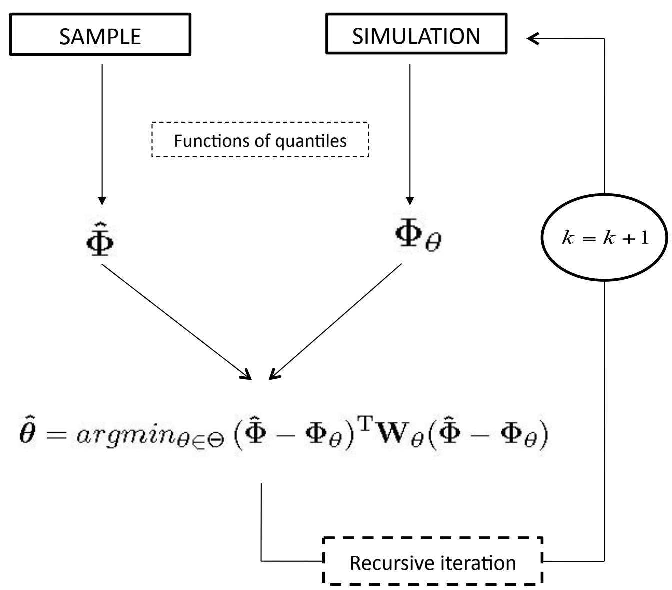
where $q_{0.75,\hat{\boldsymbol{\theta}}^{\dagger}}$ and $q_{0.25\hat{\boldsymbol{\theta}}^{\dagger}}$ are the theoretical 0.75-th and 0.25-th quantiles of a standardized α-stable distribution evaluated at $\hat{\alpha}$ and $\hat{\beta}$. Likewise, for μ:

$$\hat{\mu} = \begin{cases} \hat{q}_{0.50} + \hat{\sigma}\left(\hat{\beta}\tan\frac{\pi\hat{\alpha}}{2} - q_{0.50,\hat{\boldsymbol{\theta}}'}\right) & \text{for} \quad \alpha \neq 1 \\ \hat{q}_{0.50} - \hat{\sigma} q_{0.50,\hat{\boldsymbol{\theta}}'} & \text{for} \quad \alpha = 1. \end{cases} \tag{3}$$

where $q_{0.50,\hat{\boldsymbol{\theta}}^{\dagger}}$ is the theoretical 0.50-th quantile of a standardized α-stable distribution evaluated at $\hat{\alpha}$ and $\hat{\beta}$. $\square$

The optimization (1) works as follows. First, the sample functions of quantiles $\hat{\mathbf{\Phi}}$ are estimated from the observations. Second, given some initial values of the parameters, we simulate from the probability law that generates the process. The simulated sample is used to compute $\mathbf{\Phi}_{\boldsymbol{\theta}}$ (as the sample quantiles of the simulated sample). Third, an iterative process starts to find $\boldsymbol{\theta}$ that minimizes (1). The simulation and calculation of $\mathbf{\Phi}_{\boldsymbol{\theta}}$ are repeated at each iteration of the algorithm (using always the same seed). The iterative process continues until the convergence criterion is achieved. Figure 1 shows a diagrammatic representation of the estimation process.

Figure 1: MSQ iterative process



Several remarks are in order.

First, optimization (1) depends on the weighting matrix $\mathbf{W}_{\boldsymbol{\theta}}$ that in turn depends on the estimated parameters. This is a similar problem to GMM and IndInf. So we proceed similarly: we optimize (1) with $\mathbf{W}_{\boldsymbol{\theta}} = \mathbf{I}$, a $JM \times JM$ identity matrix. The estimated parameters, $\tilde{\boldsymbol{\theta}}_N$, albeit inefficient, are consistent. Then we replace $\boldsymbol{\theta}$ by $\tilde{\boldsymbol{\theta}}_N$ in $\mathbf{W}_{\boldsymbol{\theta}}$, and we optimize again (1). The optimal choice of $\mathbf{W}_{\boldsymbol{\theta}}$ is shown below.

Second, in many situations there are constraints between parameters to be estimated (e.g. equality or proportionality). A first thought to account for them is to optimize (1) subject to the constraints. This leads to a complicated constrained optimization problem that may involve Lagrange multipliers and Kuhn-Tucker conditions. In our simulation-based framework there is no need of such a constrained optimization. The constraints between the parameters can be easily imposed in the simulation step.

Third, notwithstanding the previous appealing feature, there are also constraints in the parameter spaces that the optimization has to handle. For instance, for the α-stable distribution $\alpha \in (0, 2]$, $\beta \in [-1, 1]$ and $\sigma \in \mathbb{R}^+$. An appropriate re-parametrization –a bijective function mapping the real line into a desired interval– may avoid this extra complexity.[5]

Fourth, the choice of the functions of quantiles does not affect the consistency of the estimators but the asymptotic variance (cf. next section). We could just match a small number of empirical and theoretical quantiles avoiding the functions $\mathbf{h}(\cdot)$ and $\mathbf{g}(\cdot)$. But in practice the choice of the functions of quantiles is crucial. As mentioned in the introduction, these functions should be informative about the parameters. A non informative function of quantiles about one of the parameters will make estimation to fail, as explained below in the Monte Carlo study. Alternatively, we could consider a thin grid of quantiles (i.e. a vector of quantiles of a thin grid of τ's) instead of a small number of functions of quantiles. This approach would produce consistent estimators but i) the grid of τ's needs to be thin enough to contain information about all the parameters which entails a significant increase in the complexity of the optimization (both time and computationally wise) and, ii) when the grid becomes sufficiently thin, it would not be possible to estimate efficiently since the quantiles become collinear when the number as τ's goes to infinity (this was pointed out by Carrasco and Florens (2002) in a GMM set up). As a result, the inverse of the weighting function is not continuous and it needs to be stabilized by a regularization parameter. The extension of MSQ to a continuum of quantiles, similarly to the Continuous GMM of Carrasco and Florens (2002), is an interesting extension left for further research.

Example (cont.)

We now consider the case where there are two $i.i.d.$ random variables $X_{i,j} \sim S_\alpha(\sigma_j, \beta_j, \mu_j)$, $j = 1, 2$. Let $y_{i,j}$ be the standardized values of $x_{i,j}$, so that $Y_{i,j} \sim S_\alpha(1, \beta_j, 0)$, $j = 1, 2$. Let $\boldsymbol{\theta}_1^\dagger = (\alpha, \beta_1)^{\mathrm{T}}$, $\boldsymbol{\theta}_2^\dagger = (\alpha, \beta_2)^{\mathrm{T}}$ and $\boldsymbol{\theta}' = (\alpha, \beta_1, \beta_2)^{\mathrm{T}}$.

Since the functions of quantiles for α and β are location and dispersion invariant, the vector of

[5] If a parameter θ is constrained to belong to a specific interval (a, b) then $0 < \frac{\theta - a}{b - a} < 1$, which can be modelled with a logistic function:

$$\frac{\theta - a}{b - a} = \frac{exp(\xi)}{1 + exp(\xi)},$$

where ξ is the transformed parameter ranging from $-\infty$ to ∞. Estimation of ξ is unconstrained and $\hat{\theta}_N$ can be backed from $\hat{\xi}_N$.

sample quantile functions is

$$\hat{\mathbf{\Phi}}^{\dagger} = \begin{pmatrix} \frac{\hat{q}_{1,0.95}-\hat{q}_{1,0.05}}{\hat{q}_{1,0.75}-\hat{q}_{1,0.25}} \\ \frac{\hat{q}_{2,0.95}-\hat{q}_{2,0.05}}{\hat{q}_{2,0.75}-\hat{q}_{2,0.25}} \\ \frac{(\hat{q}_{1,0.95}-\hat{q}_{1,0.50})+(\hat{q}_{1,0.05}-\hat{q}_{1,0.50})}{\hat{q}_{1,0.95}-\hat{q}_{1,0.05}} \\ \frac{(\hat{q}_{2,0.95}-\hat{q}_{2,0.50})+(\hat{q}_{2,0.05}-\hat{q}_{2,0.50})}{\hat{q}_{2,0.95}-\hat{q}_{2,0.05}} \end{pmatrix}.$$

And the vector of theoretical functions of quantiles is

$$\mathbf{\Phi}^{\dagger}_{\boldsymbol{\theta}^{\dagger}} = \begin{pmatrix} \frac{q_{0.95,\boldsymbol{\theta}_1^{\dagger}},-q_{0.05,\boldsymbol{\theta}_1^{\dagger}}}{q_{0.75,\boldsymbol{\theta}_1^{\dagger}}-q_{0.25,\boldsymbol{\theta}_1^{\dagger}}} \\ \frac{q_{0.95,\boldsymbol{\theta}_2^{\dagger}}-q_{0.05,\boldsymbol{\theta}_2^{\dagger}}}{q_{0.75,\boldsymbol{\theta}_2^{\dagger}}-q_{0.25,\boldsymbol{\theta}_2^{\dagger}}} \\ \frac{(q_{0.95,\boldsymbol{\theta}_1^{\dagger}}-q_{0.50,\boldsymbol{\theta}_1^{\dagger}})+(q_{0.05,\boldsymbol{\theta}_1^{\dagger}}-q_{0.50,\boldsymbol{\theta}_1^{\dagger}})}{q_{0.95,\boldsymbol{\theta}_1^{\dagger}}-q_{0.05,\boldsymbol{\theta}_1^{\dagger}}} \\ \frac{(0.95,q_{\boldsymbol{\theta}_2^{\dagger}}-q_{0.50,\boldsymbol{\theta}_2^{\dagger}})+(q_{0.05,\boldsymbol{\theta}_2^{\dagger}}-q_{0.50,\boldsymbol{\theta}_2^{\dagger}})}{q_{0.95,\boldsymbol{\theta}_2^{\dagger}}-q_{0.05,\boldsymbol{\theta}_2^{\dagger}}} \end{pmatrix}.$$

This is a system of 4 equations and 3 parameters. The location and dispersion parameters μ_1, μ_2, σ_1 and σ_2 are estimated similarly to (2) and (3). Several remarks: First, the constraint that the tail indexes for X_1 and X_2 to be the same is imposed in the simulation step and there is no need to use constrained optimization techniques. However, since the tail index and the asymmetry parameters have bounded spaces, we may apply the logistic transformation explained earlier. Second, simulation of the α-stable processes is easily done with the method of Chambers, Mallows and Stuck (1976). Third, in finite samples the problem that the function of quantiles that represents the tail index may take values below 2.439 is carried over for any dimension. A way to solve this finite sample artifact is constraining the function to be larger than its lower bound, as explained earlier. An alternative is to simulate H times and take the average simulated data. If the quantile function is still below the lower bound, then the constraint is imposed. A second alternative is to simulate H times, compute H times the functions of quantiles and compute their average. In any case, this problem is likely to be encountered only when α gets close to 2. It is rarely found for values of α below, say, 1.8. Last, in the case of J random variables there are $3J+1$ parameters to estimate and $4J$ equations. The system $\mathbf{\Phi}^{\dagger} - \mathbf{\Phi}^{\dagger}_{\boldsymbol{\theta}^{\dagger}}$ consists however of $2J$ equations and $J+1$ unknown parameters. For instance, if $J = 100$, the optimization consists of matching 200 equations with respect to 101 parameters. This may seem to be a large system. However, an educated guess of the initial values may alleviate it. For instance, univariate estimates can be used as initial values.[6] The $2J$ dispersion and location parameters are easily estimated regardless of the dimension. $\square$

[6] A reasonable initial value for α is the average of all the univariate estimates.

2.1 Asymptotic Properties

As MSQ is based on quantiles, we recall an old lemma (the proof can be found in Cramér, 1946, p. 369) that shows the asymptotic properties of the sample quantiles.[7]

Lemma 1 *Let $X_1, ..., X_N$ be N i.i.d. draws from a cumultative distribution function F with a continuous density function f. Let $0 < \tau_1 < ... < \tau_s < 1$. Suppose that F has a density function f in the neighborhoods of $q_{\tau_1}, ..., q_{\tau_s}$ and that f is positive and continuous at $q_{\tau_1}, ..., q_{\tau_s}$. Then $\hat{\boldsymbol{q}} = (\hat{q}_{\tau_1}, ..., \hat{q}_{\tau_s})^{\mathrm{T}}$ has the following asymptotic distribution*

$$\sqrt{N}(\hat{\boldsymbol{q}} - \boldsymbol{q}) \rightarrow^{d} \mathcal{N}(\mathbf{0}, \boldsymbol{\eta}),$$

where $\boldsymbol{q} = (q_{\tau_1}, ..., q_{\tau_s})^{\mathrm{T}}$ and where the i, j-th element of $\boldsymbol{\eta}$ is $\eta_{i,j} = \frac{\tau_i \wedge \tau_j - \tau_i \tau_j}{f(F^{-1}(\tau_i)) f(F^{-1}(\tau_j))}$.

The expression $\tau_i \wedge \tau_j$ stands for the minimum of τ_i and τ_j. The calculation of the sparsity function $f(F^{-1}(\tau_i))$ is shown in the Appendix.

We also need assumptions and further notation. At the beginning of Section 2 we introduced assumptions on the distribution of the random vector (which we denote by A1), compactness of $\boldsymbol{\Theta}$ (which we denote by A2) and on the continuity and differentiability of $\mathbf{h}(\hat{\boldsymbol{q}})$, $\mathbf{g}(\hat{\boldsymbol{q}}^{\boldsymbol{*}})$, $\mathbf{h}(\boldsymbol{q_{\theta_j}})$, and $\mathbf{g}(\boldsymbol{q^*_{\theta_j}})$ (which we denote by A3). To attain consistency we also need the standard identifiability and uniqueness conditions:

(A4) There exists a unique $\boldsymbol{\theta_0}$ such that the sample functions of quantiles equal the theoretical ones. That is $\boldsymbol{\theta} = \boldsymbol{\theta_0} \Rightarrow \hat{\boldsymbol{\Phi}} = \boldsymbol{\Phi_{\theta_0}}$.

(A5) $\hat{\boldsymbol{\theta}}$ is the unique minimizer of $(\hat{\boldsymbol{\Phi}} - \boldsymbol{\Phi_\theta})^{\mathrm{T}} \mathbf{W}_{\boldsymbol{\theta}} (\hat{\boldsymbol{\Phi}} - \boldsymbol{\Phi_\theta})$.

Denote by $\hat{\mathbf{G}}_j$ a $M \times M$ diagonal matrix with diagonal elements $\mathbf{g}(\hat{\boldsymbol{q}}^*_j)$. We gather all these vectors into $\mathbf{g}(\hat{\boldsymbol{q}}^*) = (\mathbf{g}(\hat{\boldsymbol{q}}^*_1)^{\mathrm{T}}, \ldots, \mathbf{g}(\hat{\boldsymbol{q}}^*_J)^{\mathrm{T}})^{\mathrm{T}}$ with the corresponding block diagonal $JM \times JM$ matrix $\hat{\mathbf{G}} = diag(diag(\mathbf{g}(\hat{\boldsymbol{q}}^*_1)), \ldots, diag(\mathbf{g}(\hat{\boldsymbol{q}}^*_J)))$. Similarly, let $\mathbf{G}_{\boldsymbol{\theta}}$ be a $JM \times JM$ diagonal matrix composed of J diagonal blocks, each of size M: $\mathbf{G}_{\boldsymbol{\theta}} = diag(diag(\mathbf{g}(\boldsymbol{q^*_{\theta_1}})), \ldots, diag(\mathbf{g}(\boldsymbol{q^*_{\theta_M}})))$.

(A6) $\hat{\mathbf{G}}$ converges to the non-stochastic matrix $\mathbf{G}_{\boldsymbol{\theta}}$.

Denote by $\tilde{\boldsymbol{\Omega}}_j N^{-1}$ the $M \times M$ sample variance-covariance matrix of $\mathbf{h}(\hat{\boldsymbol{q}}_j)$. We gather all these vectors into $\mathbf{h}(\hat{\boldsymbol{q}}) = (\mathbf{h}(\hat{\boldsymbol{q}}_1)^{\mathrm{T}}, \ldots, \mathbf{h}(\hat{\boldsymbol{q}}_J)^{\mathrm{T}})^{\mathrm{T}}$ with the corresponding block diagonal $JM \times JM$ matrix $\hat{\boldsymbol{\Omega}} = diag(\hat{\boldsymbol{\Omega}}_1, \ldots, \hat{\boldsymbol{\Omega}}_J)$.

(A7) $\lim_{N \to \infty} \hat{\boldsymbol{\Omega}} = \boldsymbol{\Omega_\theta}$.

[7] For a more detailed proof of the consistency see Serfling (1980, chapter 2) and for the asymptotic normality see Koenker (2005, p.71).

Note that $\mathbf{\Omega}_{\boldsymbol{\theta}} = \mathbf{\Omega}(\tau_1, ..., \tau_s; \mathcal{D}(\boldsymbol{\theta}))$. This is the equivalent of the sparsity function $f(F^{-1}(\tau))$ in Lemma 1. We need the following assumptions for the computation of the variance-covariance of $\hat{\boldsymbol{\theta}}_N$:

(A8) The $JM \times JM$ matrix $(\mathbf{G}_{\boldsymbol{\theta}}\mathbf{\Omega}_{\boldsymbol{\theta}}\mathbf{G}_{\boldsymbol{\theta}})$ is is non-singular.

Let

$$\mathbf{D}_{\boldsymbol{\theta}} = \left(\frac{\partial \mathbf{\Phi}_{\boldsymbol{\theta}}^{\mathrm{T}}}{\partial \boldsymbol{\theta}} \mathbf{W} \frac{\partial \mathbf{\Phi}_{\boldsymbol{\theta}}}{\partial \boldsymbol{\theta}^{\mathrm{T}}}\right)^{-1} \frac{\partial \mathbf{\Phi}_{\boldsymbol{\theta}}}{\partial \boldsymbol{\theta}}.$$

be a $p \times JM$ matrix that links the asymptotic properties of $\mathbf{\Phi}_{\hat{\boldsymbol{\theta}}}$ with those of $\hat{\boldsymbol{\theta}}$. The last assumption is

(A9) The $p \times p$ matrix $\left(\frac{\partial \mathbf{\Phi}_{\boldsymbol{\theta}}^{\mathrm{T}}}{\partial \boldsymbol{\theta}} \mathbf{W} \frac{\partial \mathbf{\Phi}_{\boldsymbol{\theta}}}{\partial \boldsymbol{\theta}^{\mathrm{T}}}\right)$ is non-singular.

Equipped with these assumptions, we announce a preliminary Lemma.

Lemma 2 *Given Lemma 1 and under (A1), (A3) and (A7),*

$$\sqrt{N}(\mathbf{h}(\hat{\boldsymbol{q}}) - \mathbf{h}(\boldsymbol{q}_\theta)) \to^d \mathcal{N}(\mathbf{0}, \mathbf{\Omega}_{\boldsymbol{\theta}}),$$

where $\mathbf{h}(\boldsymbol{q}_\theta) = (\mathbf{h}(\boldsymbol{q}_{\theta_1}), \ldots, \mathbf{h}(\boldsymbol{q}_{\theta_J}))^{\mathrm{T}}$ *and* $\mathbf{\Omega}_{\boldsymbol{\theta}} = \frac{\partial \mathbf{h}(\boldsymbol{q}_\theta)^{\mathrm{T}}}{\partial \boldsymbol{q}_\theta} \boldsymbol{\eta} \frac{\partial \mathbf{h}(\boldsymbol{q}_\theta)}{\partial \boldsymbol{q}_\theta^{\mathrm{T}}}$. *Furthermore, under assumption (A6)*

$$\sqrt{N}(\hat{\mathbf{\Phi}} - \mathbf{\Phi}_{\boldsymbol{\theta}}) \to^d \mathcal{N}(\mathbf{0}, \mathbf{G}_{\boldsymbol{\theta}}\mathbf{\Omega}_{\boldsymbol{\theta}}\mathbf{G}_{\boldsymbol{\theta}}).$$

The last step is to map the properties $\hat{\mathbf{\Phi}}$ on those of $\hat{\boldsymbol{\theta}}$:

Theorem *Given Lemma 2 and under (A2), (A4), (A5), (A8), (A9),*

$$\sqrt{N}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0) \to^d \mathcal{N}(\mathbf{0}, \mathbf{D}_{\boldsymbol{\theta}}\mathbf{W}_{\boldsymbol{\theta}}\mathbf{G}_{\boldsymbol{\theta}}\mathbf{\Omega}_{\boldsymbol{\theta}}\mathbf{G}_{\boldsymbol{\theta}}\mathbf{W}_{\boldsymbol{\theta}}^{\mathrm{T}}\mathbf{D}_{\boldsymbol{\theta}}^{\mathrm{T}}).$$

The multilayer sandwich form of the variance-covariance matrix has an intuitive explanation: $\mathbf{\Omega}_{\boldsymbol{\theta}}$ is the variance-covariance matrix of $\mathbf{h}(\hat{\boldsymbol{q}}_N)$. The first layer $\mathbf{G}_{\boldsymbol{\theta}} \cdot \mathbf{G}_{\boldsymbol{\theta}}$ accounts for the product of $\mathbf{h}(\hat{\boldsymbol{q}})$ with $\mathbf{g}(\hat{\boldsymbol{q}}^*)$. The second layer $\mathbf{W}_{\boldsymbol{\theta}} \cdot \mathbf{W}_{\boldsymbol{\theta}}^{\mathrm{T}}$ is the effect of the weighting matrix. The last layer, $\mathbf{D}_{\boldsymbol{\theta}} \cdot \mathbf{D}_{\boldsymbol{\theta}}^{\mathrm{T}}$ captures the effect of the mapping of the properties of $\hat{\mathbf{\Phi}}$ on those of $\hat{\boldsymbol{\theta}}$. Everything is known except $\mathbf{W}_{\boldsymbol{\theta}}$, which has to be chosen optimally, in the sense that it maximizes the information contained in the functions of quantiles.

Corollary The optimal weighting matrix $\mathbf{W}_{\boldsymbol{\theta}}^* = (\mathbf{G}_{\boldsymbol{\theta}}\mathbf{\Omega}_{\boldsymbol{\theta}}\mathbf{G}_{\boldsymbol{\theta}})^{-1}$ so that

$$\sqrt{N}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_{\mathbf{0}}) \to^d \mathcal{N}\left(\mathbf{0}, \left(\frac{\partial \mathbf{\Phi}_{\boldsymbol{\theta}}^{\mathrm{T}}}{\partial \boldsymbol{\theta}} (\mathbf{G}_{\boldsymbol{\theta}}\mathbf{\Omega}_{\boldsymbol{\theta}}\mathbf{G}_{\boldsymbol{\theta}})^{-1} \frac{\partial \mathbf{\Phi}_{\boldsymbol{\theta}}}{\partial \boldsymbol{\theta}^{\mathrm{T}}}\right)^{-1}\right).$$

As usual, and as it has been outlined above, the computation of the optimal weighting matrix $\mathbf{W}_{\boldsymbol{\theta}}^*$ requires a preliminary estimator $\tilde{\boldsymbol{\theta}}$. Note that the calculation of the asymptotic variance-covariance matrix of $\hat{\boldsymbol{\theta}}$ needs estimators of $\mathbf{G}_{\boldsymbol{\theta}}$ and $\mathbf{\Omega}_{\boldsymbol{\theta}}$. The matrix $\mathbf{G}_{\boldsymbol{\theta}}$ depends on the theoretical functions of quantiles, which are obtained by simulations. The matrix $\mathbf{\Omega}_{\boldsymbol{\theta}}$ is a function of the grid of τ's and the sparsity function. Last, the vector of derivatives $\frac{\partial \mathbf{\Phi}_{\boldsymbol{\theta}}}{\partial \boldsymbol{\theta}}$ can be computed numerically.

3 Monte-Carlo Simulations

We carry out a Monte Carlo experiment to determine the finite sample performance of the MSQ estimators. We generate 200 samples of 10000 observations for 3 different scenarios. In the first we consider a univariate α-stable distribution, so we can compare with McCulloch (1986). In the second and third we consider 5 and 10 univariate independent α-stable distributions with the same tail index –and hence their parameters have to be estimated jointly. We will refer to the first as univariate and the two last ones as 5- and 10-dimensional.[8]

For the univariate case, we show results for 12 cases, divided in 4 values of α, namely 1.5, 1.7, 1.9 and 1.95, and 3 values of β: -0.5, 0 and 0.5. Location and dispersion are always 0 and 1. For the 5-dimensional we consider the same values for α but only one combination of the asymmetry parameters $\beta_1 = -0.5$, $\beta_2 = -0.25$, $\beta_3 = 0$, $\beta_4 = 0.25$, and $\beta_5 = 0.5$. Last, for the 10-dimensional case, we also consider the same values for α and one combination for the asymmetry parameters $\beta_1 = -0.9$, $\beta_2 = -0.7$, $\beta_3 = -0.5$, $\beta_4 = -0.3$, $\beta_5 = -0.1$, $\beta_6 = 0.1$, $\beta_7 = 0.3$, $\beta_8 = 0.5$, $\beta_9 = 0.7$ and $\beta_{10} = 0.9$. Results are in Tables 1, 2 and 3. We report the median of the 200 estimates and their RMSE. For Table 1 we also show the results from McCulloch (1986).[9]

Results in the Tables show that MSQ accurately estimates the parameters. Estimates are essentially unbiased and the RMSE are quite small, even for the 10–dimensional case. In the univariate case, the comparison with McCulloch (1986) reveals that MSQ estimates have in general slightly larger RMSE, meaning that the simulation error in MSQ is slightly larger than the interpolation error in McCulloch (1986).

Although not reported here, but available under request, we did the same simulation study under different specifications. We estimated with draws of 1000 and 5000 observations, with values of α below 1, values for β very close to 1 and -1, and different values of σ and μ. Results are on the same lines to the ones shown here. In all cases we choose as initial values those coming from univariate estimations using McCulloch (1986). For α in the multidimensional cases, we take as initial value the average of the univariate estimates. To check the relevance of the initial values, we tried with some very far from the true parameter. Convergence took longer but the estimates were as closed to the true value as those shown in the Tables. Last, we also tried with different functions of quantiles. For instance, for the function of quantiles that represents the tail index, we tried with just the median. Estimation turned out to fail, meaning that the choice of the functions of quantiles are indeed important. A deeper analysis on the choice of the quantile functions is worth investigating. Last, we also studied the finite sample properties of the estimators when

[8]Consciously we are abusing of the terminology by denoting the last two cases as multidimensional, but it is convenient to make the difference with independent univariate estimations.

[9]Round quantiles from data are computed in Matlab associating the i-th largest value with $\tau = i/(1+N)$ and interpolating.

Table 1: Univariate Monte Carlo study

		α		β		σ		μ	
		median	RMSE	median	RMSE	median	RMSE	median	RMSE
MSQ									
α=1.5	β=-0.5	1.4751	0.0336	-0.5318	0.0463	1.0199	0.0242	-0.0429	0.0484
α=1.5	β=0	1.4987	0.0196	-0.0486	0.0596	1.0385	0.0406	-0.0333	0.0464
α=1.5	β=0.5	1.5216	0.0316	0.4869	0.0377	1.0454	0.0469	-0.0132	0.0355
α=1.7	β=-0.5	1.6975	0.0273	-0.5229	0.0905	1.0317	0.0335	-0.0253	0.0341
α=1.7	β=0	1.7157	0.0445	-0.0582	0.0802	1.0375	0.0391	-0.0133	0.0244
α=1.7	β=0.5	1.7081	0.0652	0.5123	0.0974	1.0462	0.0472	-0.0040	0.0224
α=1.9	β=-0.5	1.9296	0.0460	-0.5079	0.2227	1.0348	0.0359	0.0126	0.0248
α=1.9	β=0	1.9084	0.0508	0.0385	0.1870	1.0319	0.0328	0.0058	0.0185
α=1.9	β=0.5	1.9197	0.0581	0.5259	0.2171	1.0361	0.0366	-0.0042	0.0258
α=1.95	β=-0.5	1.9653	0.0295	-0.4987	0.3564	1.0248	0.0267	0.0099	0.0215
α=1.95	β=0	1.9656	0.0389	0.0065	0.1761	1.0259	0.0276	0.0092	0.0198
α=1.95	β=0.5	1.9588	0.0450	0.4521	0.2365	1.0045	0.0137	0.0144	0.0237
McCulloch									
α=1.5	β=-0.5	1.4977	0.0228	-0.5034	0.0372	0.9970	0.0131	-0.0055	0.0366
α=1.5	β=0	1.4978	0.0197	0.0008	0.0359	0.9986	0.0123	-0.0015	0.0318
α=1.5	β=0.5	1.4968	0.0231	0.5017	0.0367	0.9976	0.0131	0.0036	0.0345
α=1.7	β=-0.5	1.6981	0.0277	-0.5092	0.0835	0.9980	0.0117	-0.0053	0.0247
α=1.7	β=0	1.6964	0.0258	-0.0001	0.0543	0.9992	0.0124	-0.0018	0.0224
α=1.7	β=0.5	1.6970	0.0281	0.5162	0.0780	0.9985	0.0126	0.0014	0.0224
α=1.9	β=-0.5	1.8987	0.0361	-0.5041	0.1911	0.9998	0.0124	0.0048	0.0218
α=1.9	β=0	1.9062	0.0342	0.0006	0.1621	1.0009	0.0118	-0.0004	0.0179
α=1.9	β=0.5	1.9064	0.0345	0.5274	0.1802	1.0008	0.0124	-0.0065	0.0222
α=1.95	β=-0.5	1.9797	0.0347	-0.2449	0.3366	1.0069	0.0136	0.0172	0.0234
α=1.95	β=0	1.9577	0.0341	0.0068	0.1593	1.0064	0.0132	-0.0001	0.0156
α=1.95	β=0.5	1.9543	0.0369	0.2645	0.3075	0.9997	0.0133	-0.0095	0.0207

Monte Carlo estimation results for univariate α-stable distributions. Top panel shows the results for the MSQ and bottom panel for McCulloch (1986). The first two columns show the 12 different parameter configurations. In all cases $\sigma = 1$ and $\mu = 0$. The subsequent couples of columns show the median and root mean square error (RMSE) of 200 draws of 10000 observations each.

quantiles are matched instead of functions of quantiles. We estimated the parameters using a vector of 19 quantiles for $\tau = 0.05, \ldots, 0.95$. The estimated parameters were closed to the true values but convergence took much longer than using the parsimonious 4 functions of quantiles.

4 An Illustration

We illustrate the method with an application to 9 years of daily returns of 22 major world-wide market indexes that represent three geographical areas: America (S&P500, NASDAQ, TSX, Merval, Bovespa and IPC), Europe and Middle East (AEX, ATX, FTSE, DAX, CAC40, SMI, MIB and TA100), and East Asia and Oceania (HgSg, Nikkei, StrTim, SSEC, BSE, KLSE, KOSPI and

Table 2: 5 dimensional Monte Carlo study

		α		β_i		σ_i		μ_i	
		median	RMSE	median	RMSE	median	RMSE	median	RMSE
α=1.5	β_1=-0.5	1.5007	0.0277	-0.5118	0.0068	0.9998	0.0143	0.0115	0.0462
	β_2=-0.25	1.5007	0.0277	-0.2544	0.0138	1.0186	0.0126	-0.0147	0.0279
	β_3=0	1.5007	0.0277	0.0354	0.0355	1.0058	0.0222	0.0236	0.0325
	β_4=0.25	1.5007	0.0277	0.2624	0.0157	1.0032	0.0135	0.0048	0.0306
	β_5=0.5	1.5007	0.0277	0.5088	0.0141	1.0042	0.0144	-0.0044	0.0415
α=1.7	β_1=-0.5	1.7067	0.0319	-0.5256	0.0290	1.0063	0.0131	-0.0108	0.0309
	β_2=-0.25	1.7067	0.0319	-0.2531	0.0065	0.9904	0.0148	-0.0026	0.0198
	β_3=0	1.7067	0.0319	-0.0334	0.0336	0.9841	0.0195	-0.0139	0.0223
	β_4=0.25	1.7067	0.0319	0.2937	0.0449	1.0060	0.0132	0.0511	0.0553
	β_5=0.5	1.7067	0.0319	0.4908	0.0113	0.9923	0.0137	-0.0278	0.0395
α=1.9	β_1=-0.5	1.9161	0.0371	-0.5012	0.0104	1.0021	0.0129	-0.0090	0.0275
	β_2=-0.25	1.9161	0.0371	-0.2484	0.0047	1.0123	0.0176	-0.0417	0.0465
	β_3=0	1.9161	0.0371	-0.0223	0.0224	1.0135	0.0184	0.0198	0.0252
	β_4=0.25	1.9161	0.0371	0.2697	0.0210	1.0015	0.0116	-0.0294	0.0358
	β_5=0.5	1.9161	0.0371	0.5206	0.0235	1.0050	0.0130	-0.0017	0.0237
α=1.95	β_1=-0.5	1.9556	0.0273	-0.5076	0.0132	1.0158	0.0203	0.0234	0.0310
	β_2=-0.25	1.9556	0.0273	-0.2667	0.0178	0.9984	0.0123	-0.0113	0.0213
	β_3=0	1.9556	0.0273	-0.0378	0.0379	0.9997	0.0124	-0.0090	0.0196
	β_4=0.25	1.9556	0.0273	0.2671	0.0187	1.0149	0.0180	0.0020	0.0182
	β_5=0.5	1.9556	0.0273	0.5173	0.0207	1.0021	0.0137	0.0047	0.0221

Monte Carlo estimation results for 5 univariate α-stable distributions that share the same tail index. The first two columns show the 4 different parameter configurations. In all cases locations and scales are set to zero and one. The subsequent couples of columns show the median and root mean square error (RMSE) of 200 draws of 10000 observations each.

AllOrd). Table 4 provides further details on the indexes. The sample consists of 2536 observations and it was downloaded at the website of Yahoo-finance. The initial date is January 4, 2000 and the ending date is September 22, 2009.

Top panel of Figure 2 shows four indexes of four different continents: S&P500, FTSE, Nikkei and All Ordinaries. They all have the stylized facts of financial returns. Visual inspection reveals they are zero-mean processes with strong volatility clustering. This is particularly clear at the end of the sample, the 2008 financial crisis. Top panel of Figure 3 shows the autocorrelogram of the squared returns. The relatively low values and the slow decrease of the autocorrelations are distinguished features of the dynamic pattern in volatility. Returns also show spikes and sudden extreme variations within periods of both high and low volatility. This is a clear indication of thick tails and, possibly, skewness. Indeed, Table 5 shows descriptive statistics for all the indexes. The second column reveals possible negative skewness and the third column evidences a great deal of kurtosis.

Last column of Table 5 shows the estimated tail indexes for each return series. They range from 1.3268 to 1.6511. These values are too low to represent the unconditional tail behavior of the

Table 3: 10 dimensional Monte Carlo study

		α		β_i		σ_i		μ_i	
		median	RMSE	median	RMSE	median	RMSE	median	RMSE
α=1.5	β_1=-0.9	1.5029	0.0224	-0.8983	0.0107	0.9940	0.0155	0.0087	0.0539
	β_2=-0.7	1.5029	0.0224	-0.6888	0.0130	1.0003	0.0137	0.0259	0.0516
	β_3=-0.5	1.5029	0.0224	-0.5218	0.0237	1.0084	0.0167	-0.0253	0.0433
	β_4=-0.3	1.5029	0.0224	-0.2816	0.0184	1.0101	0.0169	0.0314	0.0389
	β_5=-0.1	1.5029	0.0224	-0.1169	0.0171	0.9916	0.0146	0.0131	0.0231
	β_6=0.1	1.5029	0.0224	0.1093	0.0095	0.9929	0.0148	0.0291	0.0357
	β_7=0.3	1.5029	0.0224	0.3002	0.0037	1.0314	0.0355	-0.0038	0.0263
	β_8=0.5	1.5029	0.0224	0.5190	0.0215	1.0036	0.0131	0.0202	0.0402
	β_9=0.7	1.5029	0.0224	0.6851	0.0164	1.0072	0.0165	-0.0082	0.0442
	β_{10}=0.9	1.5029	0.0224	0.8889	0.0139	1.0086	0.0176	0.0055	0.0498
α=1.7	β_1=-0.9	1.6988	0.0228	-0.9159	0.0213	0.9887	0.0177	-0.0212	0.0371
	β_2=-0.7	1.6988	0.0228	-0.7171	0.0208	1.0023	0.0123	-0.0388	0.0480
	β_3=-0.5	1.6988	0.0228	-0.4967	0.0064	0.9960	0.0128	0.0512	0.0557
	β_4=-0.3	1.6988	0.0228	-0.3110	0.0124	0.9960	0.0122	-0.0020	0.0208
	β_5=-0.1	1.6988	0.0228	-0.1066	0.0071	0.9968	0.0128	0.0000	0.0179
	β_6=0.1	1.6988	0.0228	0.1044	0.0048	0.9934	0.0127	0.0088	0.0207
	β_7=0.3	1.6988	0.0228	0.2860	0.0137	0.9926	0.0136	0.0401	0.0458
	β_8=0.5	1.6988	0.0228	0.4982	0.0056	0.9926	0.0143	0.0228	0.0339
	β_9=0.7	1.6988	0.0228	0.7208	0.0244	1.0098	0.0156	0.0161	0.0336
	β_{10}=0.9	1.6988	0.0228	0.9203	0.0263	1.0090	0.0168	0.0498	0.0583
α=1.9	β_1=-0.9	1.9074	0.0242	-0.8905	0.0132	0.9950	0.0128	-0.0149	0.0290
	β_2=-0.7	1.9074	0.0242	-0.7166	0.0223	1.0058	0.0124	-0.0201	0.0311
	β_3=-0.5	1.9074	0.0242	-0.5111	0.0131	1.0049	0.0136	0.0440	0.0487
	β_4=-0.3	1.9074	0.0242	-0.3063	0.0083	0.9985	0.0120	0.0019	0.0183
	β_5=-0.1	1.9074	0.0242	-0.1074	0.0078	0.9969	0.0126	0.0072	0.0196
	β_6=0.1	1.9074	0.0242	0.1109	0.0112	0.9961	0.0112	0.0115	0.0202
	β_7=0.3	1.9074	0.0242	0.3137	0.0148	0.9876	0.0160	0.0448	0.0502
	β_8=0.5	1.9074	0.0242	0.5113	0.0145	0.9889	0.0160	0.0166	0.0269
	β_9=0.7	1.9074	0.0242	0.7103	0.0143	1.0147	0.0197	-0.0020	0.0241
	β_{10}=0.9	1.9074	0.0242	0.8890	0.0122	1.0012	0.0127	0.0336	0.0428
α=1.95	β_1=-0.9	1.9486	0.0168	-0.8814	0.0193	0.9893	0.0166	0.0114	0.0262
	β_2=-0.7	1.9486	0.0168	-0.7225	0.0259	0.9794	0.0234	0.0121	0.0231
	β_3=-0.5	1.9486	0.0168	-0.5213	0.0231	0.9948	0.0129	-0.0076	0.0223
	β_4=-0.3	1.9486	0.0168	-0.3149	0.0165	0.9954	0.0130	0.0218	0.0278
	β_5=-0.1	1.9486	0.0168	-0.1025	0.0030	1.0128	0.0165	0.0253	0.0310
	β_6=0.1	1.9486	0.0168	0.1152	0.0156	1.0044	0.0119	0.0046	0.0174
	β_7=0.3	1.9486	0.0168	0.3280	0.0292	1.0136	0.0182	-0.0102	0.0202
	β_8=0.5	1.9486	0.0168	0.4800	0.0197	0.9926	0.0140	-0.0185	0.0279
	β_9=0.7	1.9486	0.0168	0.7194	0.0220	1.0109	0.0173	0.0136	0.0254
	β_{10}=0.9	1.9486	0.0168	0.8899	0.0162	0.9958	0.0131	-0.0223	0.0317

Monte Carlo estimation results for 10 univariate α-stable distributions that share the same tail index. The first two columns show the 4 different parameter configurations. In all cases locations and scales are set to zero and one. The subsequent couples of columns show the median and root mean square error (RMSE) of 200 draws of 10000 observations each for the different estimated parameters.

Table 4: Market Indexes

Name	Market	Continent
AEX	Amsterdam Stock Exchange	Europe
AllOrd	Australian Stock Market	Oceania
ATX	Wiener Börse	Europe
Bovespa	Sao Paulo Stock Exchange	South America
BSE	Bombay Stock Exchange	Asia
CAC40	Bourse de Paris	Europe
DAX	Deutsche Börse	Europe
FTSE	London Stock Exchange	Europe
HgSg	Hong Kong stock market	Asia
IPC	Mexican Stock Exchange	South America
KLSE	Bursa Malaysia	Asia
KOSPI	Korea Stock Exchange	Asia
Merval	Buenos Aires Stock Exchange	South America
MIB	Borsa Italiana	Europe
NASDAQ	NASDAQ Stock Market	North America
Nikkei	Tokyo Stock Exchange	Asia
SSEC	Shanghai Stock Exchange	Asia
SMI	SIX Swiss Exchange	Europe
S&P500	NYSE	North America
StrTim	Singapore Exchange	Asia
TA100	Tel Aviv Stock Exchange	Europe
TSX	Toronto Stock Exchange	North America

HgSg, StrTim, and AllOrd stand for Hang Seng, Straits Times and All Ordinaries respectively.

observations and they are due to the volatility clustering. Several articles have investigated the relationship between α-stable processes and processes with volatility clustering. De Vries (1991) has shown that under certain conditions on the parameters of a GARCH-like process, the α-stable and GARCH processes are observationally equivalent from the viewpoint of the unconditional distribution. Ghose and Kroner (1995) establish that many of the properties of α-stable models are shared by GARCH models. In particular, both models share the facts that the unconditional distribution has fat tails and that the tail index is invariant under addition. However, they reveal distinctive properties, namely the clustering in volatility, that is not present in α-stable distributions, and the distributions of the extreme values, captured by the tail indexes.

To safeguard against conditional volatility (and possible mean reversion), we adjust the returns with a VAR(2)-CCC model (i.e. Constant Conditional Correlation with GARCH(1,1) models in the conditional volatilities) such that the remaining heteroskedasticity is not due to dynamic conditional volatility. Bottom panel of Figure 2 shows the standardized adjusted returns. The volatility clustering has disappeared. This is also verified in the bottom panel of Figure 3. The autocorrelations of the squared adjusted returns do not show significant autocorrelation. However, they do not appear to be Gaussian. Table 6 shows the skewness and kurtosis. Once volatility clustering

Table 5: Descriptive Statistics: Returns

	s.d.	skewness	kurtosis	$\hat{\alpha}_T$
AEX	0.0163	-0.0420	8.9851	1.4196
AllOrd	0.0101	-0.7292	11.277	1.4986
ATX	0.0146	-0.3767	12.731	1.4699
Bovespa	0.0198	-0.0695	6.7367	1.6511
BSE	0.0176	-0.2021	9.1432	1.4564
CAC40	0.0157	0.0342	8.1883	1.5249
DAX	0.0166	0.0756	7.3018	1.5474
FTSE	0.0132	-0.1035	9.4666	1.4790
HgSg	0.0167	-0.0301	11.196	1.3872
IPC	0.0152	0.0927	7.1575	1.5438
KLSE	0.0099	-0.8620	13.315	1.4774
KOSPI	0.0184	-0.5223	7.8422	1.4789
Merval	0.0220	-0.0476	8.1843	1.4355
MIB	0.0128	-0.1738	10.241	1.4147
NASDAQ	0.0191	0.1223	7.3096	1.4513
Nikkei	0.0160	-0.3126	9.8454	1.5638
SSEC	0.0167	-0.0808	7.5209	1.4053
SMI	0.0130	0.0743	9.2335	1.5288
S&P500	0.0138	-0.0970	11.081	1.4457
StrTim	0.0133	-0.4119	8.7269	1.5244
TA100	0.0142	-0.3606	8.4907	1.3268
TSX	0.0129	-0.7181	12.135	1.4810

Descriptive statistics of the 22 index returns. This table shows the standard deviation -denoted by s.d.-, skewness, kurtosis, and the tail index, denoted by $\hat{\alpha}_T$, estimated independently for each return index.

is removed, kurtosis diminishes but still far from Gaussianity, and a high degree of negative skewness is unveiled. The last two columns display the Cramer-von Mises and the Anderson-Darling empirical distribution tests statistics. Since adjusted returns are standardized, the null hypothesis is a $\mathcal{N}(0,1)$ distribution. Out of the 44 tests, only in seven we are not able to reject the null at 5% (and the largest of the p-values of those seven tests is 13.2%). Hence, we conclude that, after controlling for volatility clustering and mean reversion, market indexes are not Gaussian.

This illustration may be criticized on the grounds that the tail index is the same for all return series. We estimate univariate α-stable distributions. Left part of Table 7 shows the estimation results. The tail indexes are higher, ranging from 1.54 to 1.93, than those in Table 5, confirming the fact that the tail indexes were capturing the volatility clustering, a genuine conditional effect. Figure 4 shows them along with their 5% confidence bands (dashed lines) and the average (horizontal straight dotted line). The constraint of a single tail index does not seem unreasonable. On the other hand, the estimation results dovetail with the descriptive statistics of Table 6. Most of the estimated β_i are negative and, even, sometimes they show a large degree of

Table 6: Descriptive Statistics: Filtered returns

	Skewness	Kurtosis	Cr-von Mis	And-Drlng
AEX	-0.3162	6.5239	0.8293 (0.0062)	4.7114 (0.0040)
AllOrd	-0.1374	3.8951	0.3225 (0.1169)	1.9807 (0.0941)
ATX	-0.1254	4.6071	0.6052 (0.0217)	3.8482 (0.0103)
Bovespa	-0.0651	4.1318	0.9714 (0.0029)	5.4471 (0.0018)
BSE	-0.1887	5.6807	1.0967 (0.0015)	7.1688 (0.0003)
CAC40	-0.0779	8.4249	1.2966 (0.0005)	8.2988 (0.0000)
DAX	0.0888	4.7314	1.0886 (0.0015)	6.6839 (0.0005)
FTSE	-0.1100	4.2751	0.6052 (0.0217)	4.0859 (0.0079)
HgSg	0.0061	3.5935	0.3043 (0.1314)	1.8773 (0.1074)
IPC	0.0213	4.9774	1.1449 (0.0011)	6.8585 (0.0004)
KLSE	-0.7711	9.6183	2.7532 (0.0000)	14.645 (0.0000)
KOSPI	-0.1954	3.9019	0.4467 (0.0545)	3.0087 (0.0271)
Merval	0.0131	5.3476	2.3002 (0.0000)	13.907 (0.0000)
MIB	-0.5260	7.4854	1.2268 (0.0000)	8.1812 (0.0000)
NASDAQ	-0.5143	6.4510	1.4427 (0.0002)	7.9401 (0.0003)
Nikkei	0.0004	4.1094	0.3034 (0.1321)	1.9187 (0.1018)
SSEC	-0.0377	5.9127	2.9071 (0.0000)	16.154 (0.0000)
SMI	-0.1385	4.1140	0.8095 (0.0069)	4.6632 (0.0042)
S&P500	-0.0249	4.1113	0.6167 (0.0204)	3.6129 (0.0135)
StrTim	-0.3062	6.1301	1.5045 (0.0002)	8.3576 (0.0000)
TA100	-0.3204	5.7394	2.9165 (0.0030)	16.292 (0.0003)
TSX	-0.5540	4.7482	1.4292 (0.0003)	10.675 (0.0000)

Descriptive statistics of the 22 filtered returns. This table shows the skewness, kurtosis, the Cramer-von Mises (Cr-von Mis) and the Anderson-Darling (And-Drlng) empirical distribution tests statistics. The null hypothesis is that the filtered returns follow a Gaussian distribution with zero mean and unit variance. Numbers in parenthesis are the p-values.

skewness. These however, are only reliable when α is not too close to 2. The location parameters are basically statistically not different from zero, and the dispersion parameters are smaller than 0.70, which is the value that it would correspond in the Gaussian case with variance equal to one.[10]

Right part of Table 7 shows the multidimensional estimates. The estimated tail index is 1.7483, very similar to the average tail index from the univariate estimations, and implying thicker tails than in the Gaussian case. The other estimated parameters are very similar to those of the univariate estimations, which makes sense, and significantly different from zero, except the location parameters that are very small and most of the times not significant.

From the previous sections it has been highlighted that the advantage of imposing a common tail index is that the distribution remains stable across aggregation. Let $\tilde{\mathbf{r}}_i = (\tilde{r}_{1,i}, \ldots, \tilde{r}_{22,i})$ be the adjusted returns. Each one is distributed as $S_{\hat{\alpha}}(\hat{\sigma}_j, \hat{\beta}_j, \hat{\mu}_j)$. Let $\check{r}_i = \sum_{j=1}^{22} \omega_j \tilde{r}_{i,j}$ with known

[10]Since $S_2(\sigma, 0, \mu) = \mathcal{N}(\mu, 2\sigma^2)$, the standardization implies that σ should be close to $1/\sqrt{2}$.

weights such that $\sum_{j=1}^{22}\omega_j = 1$. Then, since $\hat{\alpha} > 1$, $\check{r}_i \sim S_{\hat{\alpha}}(\check{\sigma}, \check{\beta}, \check{\mu})$ where[11]

$$
\begin{aligned}
\check{\mu} &= \sum_{j=1}^{22}\omega_j\hat{\mu}_j, \\
\check{\sigma} &= \left(\sum_{j=1}^{22}|\omega_j|^{\hat{\alpha}}\hat{\sigma}_j^{\hat{\alpha}}\right)^{1/\hat{\alpha}} \quad \text{and} \\
\check{\beta} &= \frac{\sum_{j=1}^{22}\hat{\beta}_j|\omega_j|^{\hat{\alpha}}\hat{\sigma}_j^{\hat{\alpha}}sign(\omega_j)}{\sum_{j=1}^{22}|\omega_j|^{\hat{\alpha}}\hat{\sigma}_j^{\hat{\alpha}}}.
\end{aligned}
$$

As weights we choose the percentage of market capitalization of the market representing each index relative to the sum of all market capitalizations (shown in the last column of Table 7).[12] The largest weight, by far, corresponds to S&P500, followed by Nikkei, NASDAQ and FTSE. The last row of Table 7 shows the parameters of the α-stable distribution for $\check{r}_i$. As expected, it shows negative skewness. However, and interestingly enough, the dispersion parameter σ is much smaller than the market ones. Aggregation reduces substantially the variations. This is expected since the dispersion of a linear combination of random variables is smaller than the linear combination of the dispersions.[13]

We now construct a world portfolio on the observed returns and compute its conditional distribution, which includes its conditional location, volatility and asymmetry. Denote by $\mathbf{m}_i$ and $\mathbf{H}_i$ the conditional location (a 22×1 vector) and volatility (a 22×22 matrix) used in the adjustment prior to estimation. The j,dth element of $\mathbf{H}_i$ is denoted by $h_{j,d,i}$. Then

$$\mathbf{r}_i = \mathbf{H}_i^{1/2}(\tilde{\mathbf{r}}_i + \mathbf{m}_i)$$

is the vector of returns at time i. Given that $\tilde{\mathbf{r}}_i$ is a vector with independent stable distributed random variables, $\mathbf{H}_i^{1/2}$ can be understood as a time varying mixing matrix.[14]

Given that $\alpha \neq 1$, the returns $\mathbf{r}_i$ are, conditional on past information, multivariate stable distributed $S_{\hat{\alpha}}(\hat{\boldsymbol{\Gamma}}_i, \hat{\boldsymbol{\delta}}_i)$ where $\hat{\boldsymbol{\delta}}_i$ is a 22×1 location vector with jth element

$$\hat{\delta}_{j,i} = \sum_{d=1}^{22} h_{j,d,i}^{1/2} m_{d,i},$$

[11] This result comes from Properties 1.2.1 and 1.2.3 of Samorodnitsky and Taqqu (2004)

[12] The source is the 2008 World Federation of Exchange report.

[13] This is a well known effect in empirical finance: the volatility of a market index is lower than the volatility of their constituents.

[14] For the sake of simplicity the location parameters of $\tilde{\mathbf{r}}_i$ are set to zero since, with a confidence level of 10% the contrary only happens in one case (out of 22).

and discrete spectral measure

$$\begin{aligned}\hat{\mathbf{\Gamma}}_i &= \sum_{j=1}^{22} \frac{1+\hat{\beta}_j}{2}\hat{\sigma}_j^{\hat{\alpha}} \left(\sum_{d=1}^{22} h_{j,d,i}\right)^{\hat{\alpha}/2} \delta\left(\frac{h_{j,1,i}^{1/2}}{\left(\sum_{d=1}^{22} h_{j,d,i}\right)^{1/2}}, \ldots, \frac{h_{j,22,i}^{1/2}}{\left(\sum_{d=1}^{22} h_{j,d,i}\right)^{1/2}}\right) \\ &+ \frac{1-\hat{\beta}_j}{2}\hat{\sigma}_j^{\hat{\alpha}} \left(\sum_{d=1}^{22} h_{j,d,i}\right)^{\hat{\alpha}/2} \delta\left(\frac{-h_{j,1,i}^{1/2}}{\left(\sum_{d=1}^{22} h_{j,d,i}\right)^{1/2}}, \ldots, \frac{-h_{j,22,i}^{1/2}}{\left(\sum_{d=1}^{22} h_{j,d,i}\right)^{1/2}}\right),\end{aligned}$$

where the Dirac measure $\delta(\mathbf{s}_j)$ assigns unit mass to the point $\mathbf{s}_j$ that is inside the 22-dimensional unit sphere. The spectral measure is therefore concentrated in 44 points (and hence its discrete nature). The way the spectral measure is written above is standard but not very intuitive. Another way to understand it is as a vector of weights

$$\begin{aligned}\hat{\mathbf{\Gamma}}_i = &\left(\frac{1+\hat{\beta}_1}{2}\hat{\sigma}_1^{\hat{\alpha}} \left(\sum_{d=1}^{22} h_{1,d,i}\right)^{\hat{\alpha}/2}, \ldots, \frac{1+\hat{\beta}_{22}}{2}\hat{\sigma}_{22}^{\hat{\alpha}} \left(\sum_{d=1}^{22} h_{22,d,i}\right)^{\hat{\alpha}/2},\right. \\ &\left.\frac{1-\hat{\beta}_1}{2}\hat{\sigma}_1^{\hat{\alpha}} \left(\sum_{d=1}^{22} h_{1,d,i}\right)^{\hat{\alpha}/2}, \ldots, \frac{1-\hat{\beta}_{22}}{2}\hat{\sigma}_{22}^{\hat{\alpha}} \left(\sum_{d=1}^{22} h_{22,d,i}\right)^{\hat{\alpha}/2}\right)\end{aligned}$$

corresponding to the points $\mathbf{s}_1, \ldots, \mathbf{s}_{44}$.

Let $\hat{\mathbf{\Gamma}}_{d,i}$ be the dth weight of $\hat{\mathbf{\Gamma}}_i$ and let $\boldsymbol{\omega} = (\omega_1, \ldots, \omega_{22})'$ denote a vector of the weights used previously. Then, the world portfolio $r_i^W = \boldsymbol{\omega}^{\mathrm{T}}\mathbf{r}_i$ follows, conditional on past information, a univariate stable distribution with tail index $\hat{\alpha}$, and asymmetry, volatility and location

$$\begin{aligned}\beta_i^W &= \frac{\sum_{d=1}^{44} |\boldsymbol{\omega}^{\mathrm{T}}\mathbf{s}_d|^{\hat{\alpha}} sign(\boldsymbol{\omega}^{\mathrm{T}}\mathbf{s}_d)\hat{\Gamma}_{d,i}}{\sum_{d=i}^{2J} |\boldsymbol{\omega}^{\mathrm{T}}\mathbf{s}_d|^{\hat{\alpha}}\hat{\Gamma}_{d,i}}, \\ \sigma_i^W &= \left(\sum_{d=1}^{44} |\boldsymbol{\omega}^{\mathrm{T}}\mathbf{s}_d|^{\hat{\alpha}}\hat{\Gamma}_{d,i}\right)^{1/\hat{\alpha}} \quad \text{and} \\ \mu_i^W &= \boldsymbol{\omega}^{\mathrm{T}}\hat{\boldsymbol{\delta}}_i.\end{aligned}$$

The asymmetry of the world portfolio is time-varying. Let $vol_i^W = \sigma_i^W \sqrt[\hat{\alpha}]{252} \times 100$ denote the annualized volatility in percentage. Figure 5 shows the returns and the estimated conditional location (top left plot), the annualized volatility (top right plot), and the conditional asymmetry (bottom plot). The conditional location of the portfolio returns is basically zero, reflecting the well known fact of unpredictability of returns. The shape of the annualized volatility is as expected, with the increases in volatility due to the burst of the internet bubble, the accounting scandals in the US in 2002, the calm period 2004-2007, the subsequent surge in volatility at the spring of the subprime crisis, and the sky rocketing volatility with the bankrupt of Lehman Brothers. Last, the conditional asymmetry is always negative and shows a very heterogenous pattern, ranging from -0.13 to -0.09. Interestingly enough, there is an inverse relation between volatility and asymmetry: periods of large volatility are tied with the periods of the largest negative asymmetry. This reflects the well known asymmetry reaction of volatility to news: bad news -i.e. negative returns- entail larger volatility than good news.

5 Conclusion

We have presented the Method of Simulated Quantiles, a new estimation method based on quantiles. It is useful for situations where the density function does not have a closed form, but it is simple to simulate, or/and moments do not exist. The basic principle is the construction of functions of theoretical quantiles, which depend on the parameters of the assumed probability law, that are matched with sample quantiles, which depend on data. The theoretical quantiles may not have a closed form expression, and hence we rely on simulation. Theoretically the method is appealing for its simplicity and its finite sample and asymptotic properties.

All along the article, the method is illustrated with the estimation of α-stable distributions. For the univariate case, MSQ can be seen as an extension of McCulloch (1986) quantile method. However, MSQ also handles multidimensional cases as, for instance, the joint estimation of N univariate α-stable distributions with the constraint of a common tail index. This is a situation that may arise when we want to construct portfolios. The portfolio is α-stable distributed only if all its constituents share the same tail index. A thorough univariate, 5- and 10-dimensional Monte Carlo study shows the goodness of the method. In all cases, it provides unbiased estimates with small RMSE. The article is completed with an application to times series of returns of 22 world wide market indexes and, using results on multivariate stable distributions, the construction of a world portfolio with a stable distribution and conditional location, volatility and asymmetry.

Appendix

The Appendix is divided in two parts. The first shows the proofs of Lemma 2 and of the Theorem. The second shows how to estimate the sparsity function.

Proofs

Proof of Lemma 2 Since $\mathbf{h}(\hat{\boldsymbol{q}})$ is a continuous and differentiable transformation of $\hat{\boldsymbol{q}}$, the Delta method applies. And since $\hat{\mathbf{G}}$, by assumption (A6), converges to the non-singular diagonal matrix $\mathbf{G}_{\boldsymbol{\theta}}$, in the limit $\hat{\boldsymbol{\Phi}}$ is just a scale transformation of $\mathbf{h}(\hat{\boldsymbol{q}})$.

Proof of the Theorem Recall the minimization problem

$$\hat{\boldsymbol{\theta}} = argmin_{\boldsymbol{\theta}}\, (\hat{\boldsymbol{\Phi}} - \boldsymbol{\Phi}_{\boldsymbol{\theta}})^{\mathrm{T}}\, \mathbf{W}_{\boldsymbol{\theta}} (\hat{\boldsymbol{\Phi}} - \boldsymbol{\Phi}_{\boldsymbol{\theta}}).$$

The first order condition is

$$\frac{\partial \mathbf{\Phi}_{\hat{\boldsymbol{\theta}}}}{\partial \boldsymbol{\theta}} \mathbf{W}_{\boldsymbol{\theta}} (\hat{\mathbf{\Phi}} - \mathbf{\Phi}_{\hat{\boldsymbol{\theta}}}) = 0.$$

A first order Taylor expansion around $\boldsymbol{\theta}_0$ in $\mathbf{\Phi}_{\hat{\boldsymbol{\theta}}}$

$$\mathbf{\Phi}_{\hat{\boldsymbol{\theta}}} = \mathbf{\Phi}_{\boldsymbol{\theta}_\mathbf{0}} + \frac{\partial \mathbf{\Phi}_{\boldsymbol{\theta}_\mathbf{0}}}{\partial \boldsymbol{\theta}} (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_\mathbf{0}).$$

Introduced in the first order condition and after reordering the elements and multiplying by $\sqrt{N}$ one gets

$$\sqrt{N}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_\mathbf{0}) \approx \left[\frac{\partial \mathbf{\Phi}_{\hat{\boldsymbol{\theta}}}}{\partial \boldsymbol{\theta}}^{\mathrm{T}} \mathbf{W}_{\boldsymbol{\theta}} \frac{\partial \mathbf{\Phi}_{\boldsymbol{\theta}_\mathbf{0}}}{\partial \boldsymbol{\theta}^{\mathrm{T}}} \right]^{-1} \frac{\partial \mathbf{\Phi}_{\hat{\boldsymbol{\theta}}}}{\partial \boldsymbol{\theta}} \mathbf{W}_{\boldsymbol{\theta}} \sqrt{N} (\hat{\mathbf{\Phi}} - \mathbf{\Phi}_{\boldsymbol{\theta}_\mathbf{0}}).$$

Since, by Lemma 2, $\hat{\mathbf{\Phi}}$ is a consistent estimator, $\hat{\boldsymbol{\theta}}$ is also consistent. Further, from the same Lemma the asymptotic normality of $\sqrt{N}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_\mathbf{0})$ is obtained.

The sparsity function

The essential feature of the covariance matrix is to estimate the inverse of the sparsity function $s(\tau) = \frac{1}{f(F^{-1}(\tau))}$ where $\tau \in [0, 1]$ and $f(F^{-1}(\tau))$ is the sparsity function (as well called the quantile density; Parzen, 1979). Siddiqui (1960) suggests to use the following approximation:

$$\hat{s}(\tau) = \frac{N}{2d} \left[x_{([N\tau]+d+1)} - x_{([N\tau]-d+1)} \right],$$

where $x_{(i)}$ denotes the ith order statistic from the one sample $\{x_1, ..., x_N\}$ model, $[x]$ is the greatest integer contained in x and d is a bandwidth. We use the one suggested by Bofinger (1975):

$$d = N^{1/5} \left(\frac{4.5\phi^4(\Phi^{-1}(\tau))}{(2(\Phi^{-1}(\tau))^2 + 1)^2} \right)^{1/5},$$

where ϕ is the normal probability density function and Φ^{-1} is the normal quantile.

References

[1] J. Aitchison and J. Brown (1957). *The Lognormal Distribution.* Cambridge University Press, Cambridge.

[2] T. Bolerslev and J.M. Wooldrige (1992), Quasi-maximum likelihood estimation and inference in dynamic models with time-varying covariances, *Econometrics Reviews* 11, 143-172.

[3] A.L. Bowley (1920). *Elements of statistics.* Charles Scribner's Sons, New York

[4] G. Brys and M. Hubert and A. Struyf (2006). Robust measures of tail weight, *Computational Statistics and Data Analysis*, 50, 733-759

[5] M. Carrasco and J.P. Florens (2002), Efficient GMM Estimation Using the Empirical Characteristic Function, GREMAQ-University of Toulouse Working Paper

[6] J.M. Chambers, C.L. Mallows and B.W. Stuck (1976). A method for simulating stable random variables, *J. Amer. Statist. Assoc.*, 71, 340-344

[7] D. Chenyao, S. Mittnik, and T. Doganoglu (1999). Computing the probability density function of the stable paretian distribution, *Mathematical and Computer Modelling* 29, 235-240.

[8] H. Cramér (1946). *Mathematical methods of statistics.* Princeton N.J.: Princeton University Press.

[9] E.L. Crow and M.M. Siddiqui (1967). Robust estimation of location, *J. Amer. Statist. Assoc.*, 62, 353-389.

[10] V. Czellar, G.A. Karolyi, and E. Ronchetti (2007), Indirect robust estimation of the short-term interest rate process. *Journal of Empirical Finance* 14, 546-563.

[11] C.G. De Vries (1991), On the relation between GARCH and stable processes, *Journal of Econometrics*, 48, 313-324

[12] D. Duffie and K.J. Singleton (1993). Simulated Moments Estimation of Markov Models of Asset Prices, *Econometrica*, 61, 929-952

[13] W.H. DuMouchel (1973), On the Asymptotic Normality of the Maximum-Likelihood Estimate when Sampling from a Stable Distribution, *Annals of Statistics*, 1, 948-957

[14] W.H. Dumouchel (1975). Stable distributions in statistical inference: 2. information from stably distributed samples, *J. Amer. Statist. Assoc.*, 70, 386-393

[15] E.F. Fama and R. Roll (1968). Some properties of symmetric stable distributions, *J. Amer. Statist. Assoc.*, 63, 817-836

[16] E.F. Fama and R. Roll (1971). Parameter estimates for symmetric stable distributions, *J. Amer. Statist. Assoc.*, 66, 331-338

[17] A. Feuerverger and McDunnough (1981), On the efficiency of Empirical Characteristic Function Procedures, *Journal of the Royal Statistical Society*, Series B Statistical Methodology, 43, 20-27

[18] B.D. Fielitz and J. Rozelle (1981), Method-of-Moments Estimators of Stable Distribution Parameters, *Applied Mathematics and Computation*, 8, 303-320

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

[20] R. Garcia, E. Renault and D. Veredas (2010). Estimation of Stable Distributions by Indirect Inference. Forthcoming in *Journal of Econometrics*.

[21] M.G. Genton and E. Ronchetti (2003), Robust indirect inference, *Journal of the American Statistical Association* 98, 67-76.

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

[23] C. Gouriéroux and A. Monfort (1996). *Simulation-based econometric methods*. Oxford University Press

[24] C. Gouriéroux, A. Monfort and E. Renault (1993). Indirect Inference. *Journal of Applied Econometrics*, 8, Supplement: Special issue on econometric inference using simulation techniques, 85-118

[25] R.A. Groeneveld (1984). Measuring skewness and kurtosis, *The Statistician*, 391-399

[26] R.A. Groeneveld (1998). A Class of Quantile Measures for Kurtosis, *The American Statistician*, 52

[27] V. Hajivassiliou and P. Ruud (1994). Classical estimation methods for LDV models using simulation , *Handbook of Econometrics*, 4, 2383-2441

[28] B.L. Joiner and J.R. Rosenblatt (1971). Some properties of the range in samples from Tukey's symmetric lambda distributions. *J. Amer. Statist. Assoc.*, 66, 394-399.

[29] R. Koenker (2005). *Quantile Regression*. Cambridge University Press

[30] I.A. Koutrouvelis (1980) Regression-type Estimation of the Parameters of Stable Laws, *J. Amer. Statist. Assoc.*, 75, 918-928

[31] M.J. Lombardi, G. Calzolari and G.M. Gallo (2004), Indirect Inference for Stable Distributions, Working paper 2004/07, Universita degli Studi di Firenze.

[32] J.H. McCulloch, (1986). Simple consistent estimators of stable distribution parameters. *Commun. Statist. Simula.*, 15(4), 1109-1136

[33] D. McFadden (1989). A Method of Simulated Moments for Estimation of Discrete Response Models without Numerical Integration, *Econometrica*, 57, 995-1026

[34] J.J.A. Moors (1988). A quantile alternative for kurtosis, *The Statistician*, 25-32

[35] J.P. Nolan (2001). Maximum likelihood estimation of stable parameters, *Lévy Processes: Theory and Applications*, 379-400.

[36] C. Ortelli and F. Trojani (2005), Robust efficient method of moments. *Journal of Econometrics* 128, 69-97.

[37] A. Pakes and D. Pollad (1989). The Asymptotics of Simulation Estimators, *Econometrica*, 57, 1027-1058

[38] A.S Paulson, E.W. Holcomb and R.A. Leitch (1975). The Estimation of the Parameters of the Stable Laws, *Biometrika*, 62, 163-170

[39] J.S. Press (1972). Estimation of Univariate and Multivariate Stable Distributions, *J. Amer. Statist. Assoc.*, 67 (340), 842-846

[40] J.S. Ramberg and B.W. Schmeiser (1974). An approximative method for generating asymmetric random variables. *Commun. of the ACM*, 17(2), 78-82.

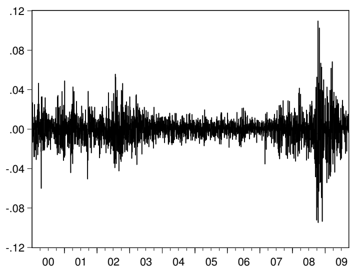
[41] G. Samorodnitsky and M.S. Taqqu (1994). *Stable non-gaussian random processes: Stochastic models with infinite variance.* Chapman & Hall/CRC, Stochastic Modeling

[42] R.J. Serfling (1980). *Approximation theorems of mathematical statistics.* John Wiley & Sons

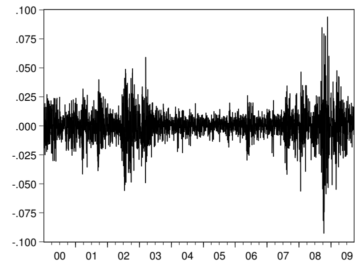
[43] *Annual report and statistics 2008*, World Federation of Exchanges

[44] V.M. Zolotarev (1957). Mellin-Stieltjes transforms in probability theory, *Theory of Probability and its Application*, 2, 433-460

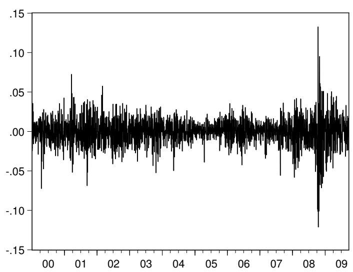
Figure 2: Returns (top panel) and adjusted returns (bottom panel).



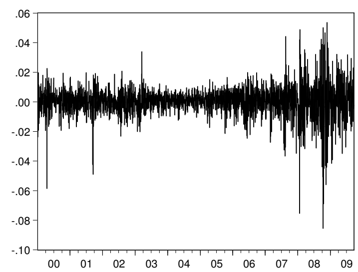
(a) S&P500



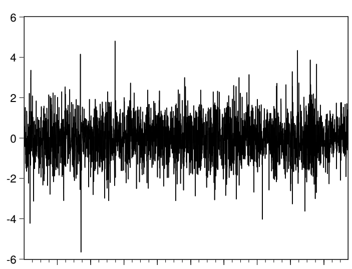
(b) FTSE



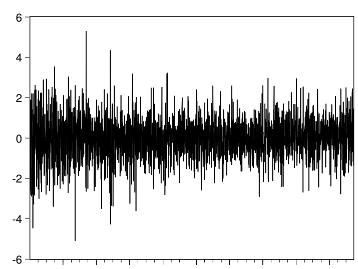
(c) Nikkei



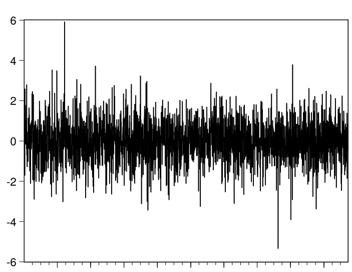
(d) All Ordinaries



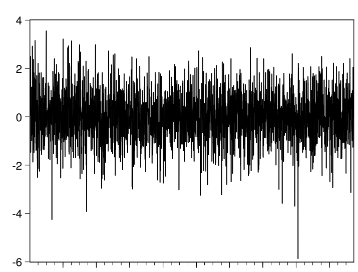
(e) S&P500



(f) FTSE

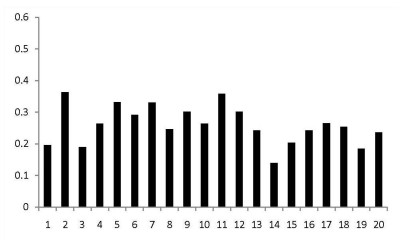


(g) Nikkei

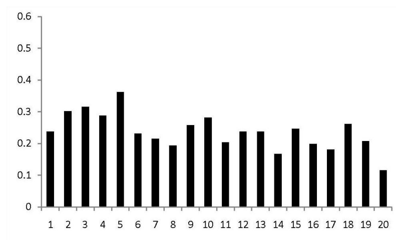


(h) All Ordinaries

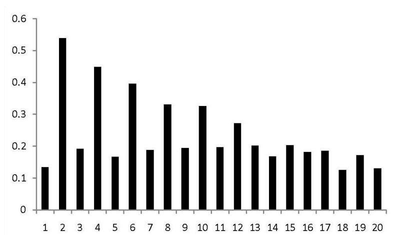
Figure 3: Autocorrelograms of squared returns (top panel) and squared adjusted returns (bottom panel).



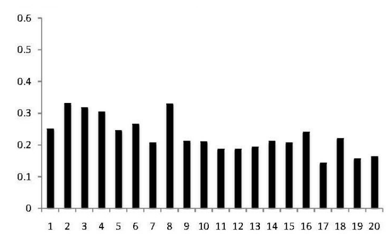
(a) S&P500



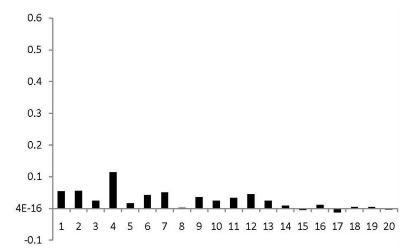
(b) FTSE



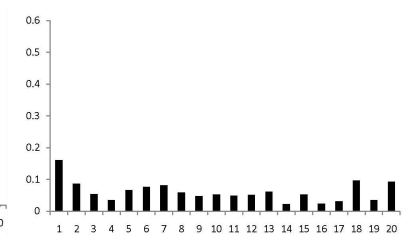
(c) Nikkei



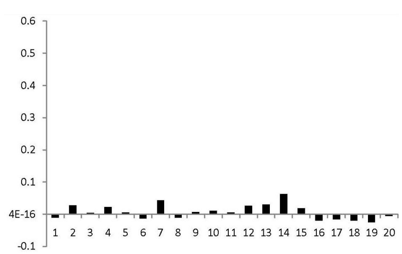
(d) All Ordinaries



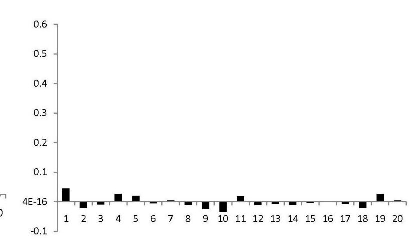
(e) S&P500



(f) FTSE



(g) Nikkei



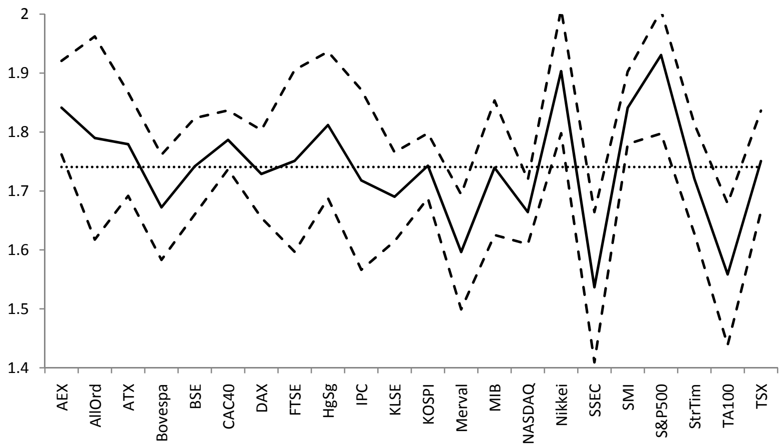
(h) All Ordinaries

Table 7: Estimated parameters

	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\sigma}$	$\hat{\mu}$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\sigma}$	$\hat{\mu}$	Weights (%)
AEX	1.8413	-0.3161	0.6395	0.0005	1.7483	-0.3173	0.6363	-0.0171	3.7131
	(0.0405)	(0.0048)	(0.0044)	(0.0210)	(0.0176)	(0.0068)	(0.0150)	(0.0212)	
AllOrd	1.7898	0.1333	0.6517	-0.0108	1.7483	0.1338	0.6538	-0.0069	2.2493
	(0.0879)	(0.0169)	(0.0119)	(0.0279)	(0.0176)	(0.0142)	(0.0126)	(0.0194)	
ATX	1.7793	-0.1005	0.6374	-0.0315	1.7483	-0.1001	0.6322	-0.0310	0.2509
	(0.0447)	(0.0180)	(0.0104)	(0.0110)	(0.0176)	(0.0107)	(0.0138)	(0.0215)	
Bovespa	1.6722	-0.0431	0.6130	-0.0029	1.7483	-0.0433	0.6138	0.0005	1.9472
	(0.0455)	(0.0157)	(0.0190)	(0.0175)	(0.0176)	(0.0143)	(0.0142)	(0.0210)	
BSE	1.7419	-0.3266	0.6038	-0.0459	1.7483	-0.3278	0.6042	-0.0423	2.1287
	(0.0417)	(0.0164)	(0.0298)	(0.0158)	(0.0176)	(0.0079)	(0.0168)	(0.0226)	
CAC40	1.7867	0.0243	0.6073	0.0137	1.7483	0.0246	0.6066	0.0141	3.7131
	(0.0255)	(0.0079)	(0.0198)	(0.0120)	(0.0176)	(0.0083)	(0.0152)	(0.0255)	
DAX	1.7289	-0.0075	0.6152	-0.0127	1.7483	-0.0075	0.6153	-0.0108	3.6528
	(0.0379)	(0.0061)	(0.0083)	(0.0229)	(0.0176)	(0.0065)	(0.0130)	(0.0212)	
FTSE	1.7510	-0.1611	0.6261	-0.0083	1.7483	-0.1615	0.6240	-0.0130	6.1445
	(0.0787)	(0.0063)	(0.0090)	(0.0103)	(0.0176)	(0.0088)	(0.0157)	(0.0269)	
HgSg	1.8119	0.0944	0.6481	-0.0034	1.7483	0.0948	0.6478	0.0015	4.3704
	(0.0634)	(0.0196)	(0.0044)	(0.0337)	(0.0176)	(0.0096)	(0.0145)	(0.0200)	
IPC	1.7179	-0.1576	0.6053	-0.0305	1.7483	-0.1582	0.6059	-0.0285	0.7698
	(0.0779)	(0.0099)	(0.0155)	(0.0193)	(0.0176)	(0.0113)	(0.0138)	(0.0209)	
Klse	1.6902	0.1771	0.5259	0.0086	1.7483	0.1778	0.5280	0.0001	0.6219
	(0.0385)	(0.0105)	(0.0079)	(0.0072)	(0.0176)	(0.0168)	(0.0140)	(0.0194)	
Kospi	1.7428	-0.0572	0.6425	-0.0136	1.7483	-0.0573	0.6429	-0.0158	1.5485
	(0.0281)	(0.0145)	(0.0229)	(0.0170)	(0.0176)	(0.0083)	(0.0150)	(0.0210)	
Merval	1.5965	0.1045	0.5632	-0.0029	1.7483	0.1049	0.5646	-0.0170	0.1311
	(0.0497)	(0.0166)	(0.0062)	(0.0177)	(0.0176)	(0.0054)	(0.0120)	(0.0202)	
MIB	1.7394	-0.2369	0.6084	-0.0340	1.7483	-0.2384	0.6094	-0.0308	1.7172
	(0.0582)	(0.0050)	(0.0173)	(0.0180)	(0.0176)	(0.0119)	(0.0153)	(0.0208)	
NASDAQ	1.6642	-0.0928	0.5898	-0.0052	1.7483	-0.0932	0.5912	0.0044	7.8818
	(0.0278)	(0.0041)	(0.0151)	(0.0214)	(0.0176)	(0.0121)	(0.0154)	(0.0226)	
Nikkei	1.9032	-0.0737	0.6730	-0.0181	1.7483	-0.0740	0.6688	-0.0308	10.248
	(0.0537)	(0.0150)	(0.0154)	(0.0310)	(0.0176)	(0.0083)	(0.0155)	(0.0212)	
SSEC	1.5367	-0.1198	0.5445	-0.0492	1.7483	-0.1202	0.5457	-0.0242	4.6881
	(0.0650)	(0.0092)	(0.0162)	(0.0117)	(0.0176)	(0.0129)	(0.0138)	(0.0214)	
SMI	1.8411	-0.3072	0.6540	-0.0471	1.7483	-0.3084	0.6501	-0.0669	2.8197
	(0.0313)	(0.0160)	(0.0148)	(0.0126)	(0.0176)	(0.0068)	(0.0148)	(0.0231)	
S&P500	1.9305	-0.2329	0.6785	-0.0440	1.7483	-0.2340	0.6728	-0.0745	30.289
	(0.0679)	(0.0143)	(0.0224)	(0.0249)	(0.0176)	(0.0092)	(0.0147)	(0.0209)	
StrTim	1.7208	0.0530	0.6132	-0.0467	1.7483	0.0531	0.6130	-0.0492	0.8715
	(0.0475)	(0.0107)	(0.0117)	(0.0296)	(0.0176)	(0.0117)	(0.0136)	(0.0215)	
TA100	1.5585	-0.0380	0.5499	-0.0376	1.7483	-0.0381	0.5540	-0.0302	0.4433
	(0.0612)	(0.0105)	(0.0191)	(0.0198)	(0.0176)	(0.0086)	(0.0156)	(0.0224)	
TSX	1.7507	-0.3867	0.6214	-0.0849	1.7483	-0.3885	0.6222	-0.0860	3.3991
	(0.0436)	(0.0298)	(0.0192)	(0.0281)	(0.0176)	(0.0035)	(0.0147)	(0.0247)	
World					1.7483	-0.2366	0.2667	-0.0351	
					(0.0176)	(0.0036)	(0.0014)	(0.0033)	

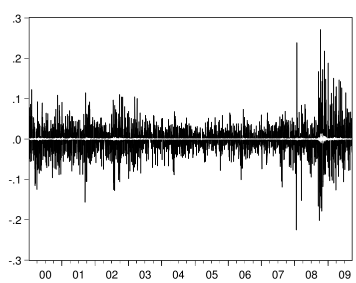
This table shows for each of the 22 financial indexes, the four estimated parameters of an α-stable distribution in the univariate and multidimensional cases. In parenthesis the asymptotic standard errors. The last column shows the weights (in percentage) of each market index on the world index.

Figure 4: Estimated α's

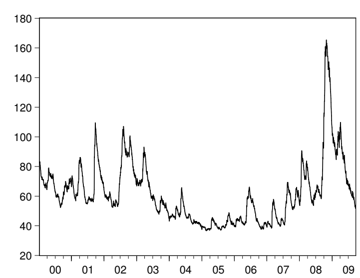


Dashed thin lines are the 95% confidence bands and the straight dotted line is the average α.

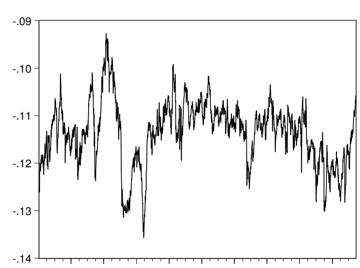
Figure 5: World portfolio conditional location, annualized volatility and asymmetry



(a) Returns and μ_i^W



(b) vol_i^W



(c) β_i^W