Estimation of Stable Distributions by Indirect Inference

René Garcia,† Eric Renault‡ and David Veredas§

First Draft February 2003, this version April 2004

Abstract

This article deals with the estimation of the parameters of an \( \alpha \)-Stable distribution by the indirect inference method with the skewed-t distribution as an auxiliary model. The latter distribution appears as a good candidate for an auxiliary model since it has the same number of parameters as the \( \alpha \)-Stable distribution, with each parameter playing a similar role. In fact, to improve the properties of the estimator in finite sample, we use a variant of the method called Constrained Indirect Inference, recently introduced by Calzolari, Fiorentini and Sentana (2001). In a Monte Carlo study, we show that this method delivers estimators with good properties in finite sample. In particular they are much more efficient than two other prevalent methods based on the characteristic function and the empirical quantiles of the \( \alpha \)-Stable distribution.

Keywords: Stable distribution, Indirect Inference, Constrained Indirect Inference, Skewed-t distribution.

JEL classification: C13, C15, G11.

*We are grateful to François Lamontagne and Marine Carrasco for useful discussions and remarks. Marine Carrasco kindly provided us with the computer code for estimating \( \alpha \)-Stable distributions with continuous GMM. The first two authors gratefully acknowledge financial support from the Institut de Finance Mathématique de Montréal (IFM2), the Fonds québécois de la recherche sur la société et la culture (FQRSC), the Social Sciences and Humanities Research Council of Canada (SSHRC) and the MITACS Network of Centres of Excellence. The first author is grateful to Hydro-Québec and the Bank of Canada for financial support. An important part of this work was carried out while the third author was a postdoctoral fellow at CIRANO. The third author thanks CIRANO, CIREQ and IFM2 for financial support.

†CIRANO and CIREQ, Université de Montréal; email: rene.garcia@umontreal.ca
‡CIRANO and CIREQ, Université de Montréal; email: eric.renault@umontreal.ca
§Econometrics and Finance Group, CentER, Tilburg University and CORE, Université catholique de Louvain; email: d.veredas@uvt.nl.
1 Introduction

The $\alpha$-Stable distribution has been widely used for fitting data in which extreme values are rather frequent. As shown in early work by Mandelbrot (1963) and Fama (1965), it is a good candidate to accommodate heavy-tailed financial series, with the consequence that it produces measures of risk based on the tails of the distribution, such as value at risk, which are more reliable. The $\alpha$-Stable distribution is also able to capture skewness in a series, another feature of financial series. The distribution is also preserved under convolution. This property is appealing when considering portfolios of assets, especially when the skewness and fat tails of returns can be taken into account. These properties motivate its use in the modelling of financial series in particular by Mittnik and Rachev (1993) and Mittnik, Paolella and Rachev (2000).\footnote{Basic references on the $\alpha$-Stable distribution are Feller (1971), Zolotarev (1986) and Samorodnitsky and Taqqu (1994).}

However, estimation of the parameters of the $\alpha$-Stable distribution is challenging. Its density function does not have a closed form. It is defined as an integral without any analytical solution and complex enough as to make its numerical integration difficult. Therefore, although DuMouchel (1973) has shown that the maximum likelihood (ML hereafter) estimator is theoretically consistent, asymptotically normal and efficient, the estimation by maximum likelihood is most of the times unfeasible. Therefore, many alternative methods have been proposed. They are based either on calibration of quantiles or on the characteristic function (CF), which has an analytical expression.

The quantile approach is based on finding the parameters which match the theoretical quantiles with their empirical counterparts. This is the method proposed by Fama and Roll (1971) and enhanced by McCulloch (1986). These estimators are consistent albeit not efficient. Several methods based on the CF are available. In general, they are all based on the comparison of the theoretical CF with its empirical counterpart. The methods differ by the basis of comparison. Since the CF is a complex gaussian random variable, one can think about comparing i) its real and complex theoretical and empirical absolute means, like Press (1972) and Fielitz and Rozelle (1981), or ii) its theoretical and empirical $r$th mean distance, like Paulson, Holcomb and Leitch (1975), Feuerverger and McDunnough (1981) and Carrasco and Florens (2002), or about iii) regressing its real and complex theoretical and empirical square absolute values, like Koutrouvelis (1980) or iv) minimizing its gaussian log-likelihood function, like Feuerverger and McDunnough (1981). All these CF methods, except Carrasco and Florens (2002), have one main drawback: the choice of the frequencies where to evaluate the CF is unclear and somewhat arbitrary. Some authors, like Fielitz and Rozelle (1981), recommend to choose few frequencies based on Monte Carlo studies while others, such as Feuerverger...
and McDunnough (1981), recommend to use as many frequencies as possible based on asymptotic arguments.

Another set of methods is based on the asymptotic tail behaviour of the distribution. The most known is the estimation via the Pareto distribution since the asymptotic tails of an $\alpha$-Stable distribution behave like a Pareto distribution. Its tail index ML estimator is the Hill estimator (Hill, 1975). These methods have also as drawbacks the arbitrary choice of the quantile from where the tail is considered and, more importantly, the fact that only the stability index $\alpha$ is estimated.

In this paper we propose an easy and intuitive way to estimate the parameters of an $\alpha$-Stable distribution. We use indirect inference, a simulation-based estimation method introduced by Gouriéroux, Monfort and Renault (1993) based on Smith (1993). The method is particularly suited to situations where the model of interest is difficult to estimate but relatively easy to simulate. This is the case of the $\alpha$-Stable distribution where several methods (see in particular Chambers, Mallows and Stuck, 1976) are known to simulate $\alpha$-Stable random variables despite the fact that the distribution function is intractable. In indirect inference, the parameters are estimated indirectly through an auxiliary model. The basic idea is to estimate the parameters of the auxiliary model first with the data, second with simulated trajectories of the process of interest based on given values of its parameters. Estimates of the parameters of interest are obtained by minimizing the distance between these two sets of parameters for the auxiliary model.

We use as an auxiliary model the skewed-t distribution introduced independently by Fernández and Steel (1998) and Hansen (1994). It is a Student-t with inverse scale factor in the positive and negative orthants, allowing for asymmetries. The distribution has four parameters which have a one-to-one correspondence with those of the $\alpha$-Stable in terms of interpretation.\footnote{During the course of this project, we were made aware by M. J. Lombardi that Lombardi, Calzolari and Gallo (2003) use the same auxiliary model to estimate a stable distribution. The two projects were conducted independently.} We first show that the ML estimators of the four parameters of the skewed-t distribution are asymptotically normal even when the observations are generated by a $\alpha$-Stable distribution. We also establish that the parameters of the $\alpha$-Stable distribution obtained by indirect inference are asymptotically normal. This is an important feature of the choice of this particular auxiliary model since the use of polynomials as in the conventional method of moments (see Gallant and Tauchen, 1999) can lead to estimators which are asymptotically $\alpha$-Stable.

Asymptotic results are important to guide the choice of the auxiliary model but they are not a guarantee that the estimators will have nice properties in finite samples. Indeed, we show in a Monte Carlo study that as the stability parameter becomes closer to
2, its upper bound, the estimated number of degrees of freedom of the skewed-t diverges towards infinity. This is due to the slow convergence of the estimator towards its asymptotic normal distribution. We therefore constrain the degrees-of-freedom parameter in the skewed-t distribution following an idea of Calzolari, Fiorentini and Sentana (2004), who dubbed the method Constrained indirect inference. This method is an extension of indirect inference procedures which allows for equality and inequality constraints on the auxiliary model parameters.

We show the usefulness of this indirect inference procedure and its constrained version when necessary for estimating $\alpha$-Stable parameters in a thorough Monte Carlo study. We also compare it with the quantile method of McCulloch (1986) and continuous GMM of Carrasco and Florens (2002). We show that it has better properties in terms of both bias and variance.

The structure of the paper is as follows. Section 2 describes the properties of $\alpha$-Stable distributions and presents efficient methods for their estimation, namely maximum likelihood and characteristic function-based methods. In Section 3, we introduce several alternative moment matching-methods, which are based on quantiles, regressions, and quasi-likelihood. Constrained indirect inference is explained in Section 4. It describes indirect inference for the $\alpha$-Stable distribution and the skewed-t distribution chosen for the auxiliary model. It also shows that the estimators are asymptotically normal. Section 5 reports the results of the Monte Carlo study where indirect inference is compared to methods using continuous GMM and empirical quantiles. Section 6 concludes.

2 The $\alpha$-Stable distributions and their efficient estimation methods

In general, the stable distributions do not have closed form expressions for density and distribution functions. On the contrary, since they are divisible, the Levy and Khinchine theorem allows to describe them easily by their characteristic functions. The family of univariate $\alpha$-stable distributions is well-defined as a parametric family of distributions indexed by four real parameters which vary freely in some intervals. However, efficient estimation of these parameters is not a trivial issue since the likelihood function is generally unknown. We remind in this section the interpretation of the four parameters of $\alpha$-stable distributions as well as the strategies available in the literature for their efficient estimation. While this efficient estimation will remain a benchmark when considering in Section 3 alternative moment-matching methods of estimation, the interpretation of the parameters will be crucial to define well-suited moments to match.
2.1 Parameters and properties of $\alpha$-stable distributions

As already mentioned, $\alpha$-stable distributions define a family of probability distributions characterized by four parameters $\alpha$, $\beta$, $\sigma$ and $\mu$, where $\alpha$ is the stability parameter, $\beta$ the skewness parameter, $\sigma$ the scale parameter, and $\mu$ the location parameter.

These parameters define the natural logarithm of the characteristic function as:

$$\ln \psi_\theta(t) = \ln E[\exp(itY)] = i\mu t - \sigma |t|^\alpha [1 - i\beta \text{sign}(t) w(t, \alpha)]$$

where:

$$\theta = (\alpha, \beta, \mu, \sigma) \in \Theta = [0, 2] \times [-1, 1] \times \mathbb{R} \times ]0, +\infty[ $$

is the vector of parameters which vary freely in the indicated intervals, $Y$ is the random variable following the $\alpha$-stable distribution $S(\theta)$ with characteristic function $\psi_\theta(\cdot)$, $\text{sign}(t) = t/|t|$ for $t \neq 0$ (and 0 for $t = 0$), $w(t, \alpha)$ is $\tan(\pi \alpha/2)$ if $\alpha \neq 1$ and $(-2 \pi \ln |t|)$ if $\alpha = 1$.

Note that the four parameters are well identified except the parameter $\beta$ when $\alpha = 2$. This case corresponds to the normal probability distribution. The three parameters $\mu, \sigma$ and $\beta$ are respectively interpreted as location, scale and skewness parameters due to the following property:

$$Y \sim S(\alpha, \beta, \mu, \sigma)$$

$$\Leftrightarrow \frac{Y - \mu}{\sigma} \sim S(\alpha, \beta, 0, 1)$$

$$\Leftrightarrow -(\frac{Y - \mu}{\sigma}) \sim S(\alpha, -\beta, 0, 1).$$

If $\beta$ is positive (resp. negative) the distribution of $Y$ is skewed to the right (resp. to the left) and this affects in particular the tails of the distribution as the property below indicates. For $0 < \alpha < 2$:

$$\lim_{\lambda \to \infty} \lambda^\alpha P[Y > \lambda] = C_\alpha \frac{1 + \beta}{2} \sigma^\alpha$$

$$\lim_{\lambda \to \infty} \lambda^\alpha P[Y < -\lambda] = C_\alpha \frac{1 - \beta}{2} \sigma^\alpha$$

where $C_\alpha = \left[ \int_0^\infty x^{-\alpha} \sin x dx \right]^{-1}$.

Therefore, the stability parameter $\alpha$ characterizes the size of the tails. While for $\alpha = 2$ we have normal probability distributions with finite moments at any order, for
0 < \alpha < 2 we have Pareto tails with infinite moments of order \( p \) for any \( p \geq \alpha \). In particular, \( \alpha \)-stable variables have infinite variance for \( \alpha < 2 \) and the central limit theorem is no longer valid. It is replaced by a stability result, stating that if \( Y_1, Y_2, \cdots, Y_n \) are i.i.d. \( S(\alpha, \beta, \mu, \sigma) \), \( \frac{1}{n^{1/\alpha}} (Y_1 + Y_2 \cdots + Y_n) \) also follows the distribution \( S(\alpha, \beta, \mu, \sigma) \).

### 2.2 Maximum likelihood estimation

The maximum likelihood estimation of the parameters \( \theta = (\alpha, \beta, \mu, \sigma) \) of an \( \alpha \)-stable distribution for an \( \theta \) interior point of \( \Theta \):

\[
0 < \alpha < 2, |\beta| < 1, \mu \in \mathbb{R}, \sigma \in ]0, +\infty[\]

raises two specific difficulties. The likelihood function is not known in closed form in general and special numerical procedures are needed to maximize it. For instance, Mitnink, Rachev, Doganoglu and Chenyao (1999) propose to recover the density function from the characteristic function by using fast Fourier transforms. More importantly, the derivation of the asymptotic distribution theory for maximum likelihood estimators in the context of an \( \alpha \)-stable family of probability distributions is not trivial. While the law of large numbers and the central limit theorem are cornerstones of this theory, they are no longer valid for i.i.d. sequences \( Y_1, Y_2, \cdots Y_n \) of variables with stable distribution, \( S(\alpha, \beta, \mu, \sigma) \), \( \alpha < 2 \). Hence, \( \frac{1}{n} \sum_{i=1}^{n} |Y_i|^p \) converges to infinity for \( p \geq \alpha \), \( \frac{1}{n^{1/\alpha}} \sum_{i=1}^{n} Y_i \) is no longer bounded in probability, but \( \frac{1}{n^{1/\alpha}} \sum_{i=1}^{n} Y_i \) is (asymptotically) distributed as \( S(\alpha, \beta, \mu, \sigma) \).

However, the score function remains asymptotically root-\( n \) normal as for more common parametric models and this fact has been used by DuMouchel (1973) to develop the asymptotic theory of maximum likelihood in the context of a family of \( \alpha \)-stable distributions. Basically, DuMouchel (1973) was able to show that the standard tools of maximum likelihood theory (mainly root-\( n \) asymptotic normality and Cramer-Rao bounds) may be applied to estimation of \( \theta = S(\alpha, \beta, \mu, \sigma) \) insofar as the domain of possible values of \( \theta \) is limited in the following way:

\[
\alpha \in ]1, 2[ \quad \text{or} \quad \alpha \in ]\varepsilon, 1[ \quad \text{for some} \ \varepsilon > 0,
\]

and

\[
|\beta| < \min (\alpha, 2 - \alpha).
\]

In particular, totally skewed-stable distribution (\(|\beta| = 1\)) are discarded as well as arbitrarily fat tails (\( \alpha \) arbitrarily close to zero).
Because of the numerical difficulties associated with maximum likelihood estimation, we will not use it in this paper. However, the results of DuMouchel (1973) are important for two reasons. They prove that efficient parametric estimation is a sensible goal, even for the parameters of $\alpha$-stable distributions. Root-$n$ asymptotic normality and standard Cramer-Rao lower bound are reached by MLE. This remark motivates the search for efficiency in the context of characteristic function-based methods of estimation considered below. After all, the characteristic function also characterizes the probability distribution and should convey the same information as the likelihood function for efficient parametric estimation. Moreover, they show that asymptotic normality of M-estimators like MLE or QMLE can be derived by the application of standard central limit theory to well-chosen (pseudo)-score functions rather than to moments of $Y$ which do not exist. This idea is the main motivation of the indirect inference strategy proposed in this paper.

2.3 Characteristic function-based methods

Let $Y_1, Y_2, \ldots, Y_n$ be $n$ observations drawn in the same probability distribution as $Y \sim S(\alpha, \beta, \mu, \sigma)$. Characteristic function techniques are built on fitting the sample characteristic function $\frac{1}{n} \sum_{j=1}^{n} \exp[itY_j]$ to the theoretical one $\psi_\theta(t)$ defined above. Press (1972) proposed several fitting methods: minimum distance, minimum $r^{th}$-mean distance and the method of moments. The problem is that it takes an infinite number of moment conditions, indexed by $t \in \mathbb{R}$, to summarize the informational content of the characteristic function:

$$Eh(t, Y, \theta) = 0, \quad \forall t \in \mathbb{R},$$

where $h(t, Y, \theta) = \exp[itY] - \psi_\theta(t)$.

Feuerverger and McDunnough (1981) and Singleton (2001) choose to work with a finite grid $t_1, t_2, \ldots, t_K$, that is to apply the standard theory of GMM to the set of moment conditions:

$$E[\exp(it_k Y) - \psi_\theta(t_k)] = 0, \quad k = 1, \ldots, K.$$

Note that this amounts to a set of $(2K)$ moment restrictions:

$$E[g_K(\theta, Y)] = 0$$
where the $2K$-dimensional vector $g_K(\theta, Y)$ is formed by stacking both the real parts:

$$\text{Re } h(t_k, Y, \theta) = \cos(t_k Y) - \text{Re } \psi(t_k)$$

and the imaginary parts

$$\text{Im } h(t_k, Y, \theta) = \sin(t_k Y) - \text{Im } \psi(t_k)$$

of

$$h(t_k, Y, \theta), h = 1, \ldots K.$$ 

Efficient GMM is obtained by using as weighting matrix a consistent estimator of the inverse of the long term asymptotic covariance matrix $\Sigma_K$ of $\frac{1}{\sqrt{n}} \sum_{j=1}^{n} \exp(itY_j)$. Note that, at least in the i.i.d case, $\Sigma_K$ admits a simple closed form expression deduced from the identity:

$$E[\exp(itY)\exp(isY)] = \Psi_\theta(t + s).$$

A simple way to back out the coefficients of $\Sigma_K$ is then:

$$\text{cov}[\cos(tY), \cos(sY)] = \frac{1}{2} \left[ \text{Re} \Psi_\theta(t+s) + \text{Re} \Psi_\theta(t-s) \right] - (\text{Re} \Psi_\theta(t)) (\text{Re} \Psi_\theta(s))$$

$$\text{cov}[\cos(tY), \sin(sY)] = \frac{1}{2} \left[ \text{Im} \Psi_\theta(t+s) + \text{Im} \Psi_\theta(t-s) \right] - (\text{Re} \Psi_\theta(t)) (\text{Im} \Psi_\theta(s))$$

$$\text{cov}[\sin(tY), \sin(sY)] = \frac{1}{2} \left[ \text{Re} \Psi_\theta(t+s) + \text{Re} \Psi_\theta(t-s) \right] - (\text{Im} \Psi_\theta(t)) (\text{Im} \Psi_\theta(s))$$

By using the empirical characteristic function $\hat{\phi}(t) = \frac{1}{n} \sum_{j=1}^{n} \exp(itY_j)$ as a consistent estimator of $\Psi_\theta(t)$, consistent estimators of the coefficients of $\Sigma_K$ are then easily deduced.

Then, the asymptotic variance of efficient GMM is given by the standard formula:

$$\Omega_K = \left\{ \left[ E \frac{\partial g_K}{\partial \theta}(\theta, Y) \right] \Sigma_K^{-1} \left[ E \frac{\partial g_K}{\partial \theta'}(\theta, Y) \right] \right\}^{-1}$$

Both Feuerverger and McDunnough (1981) and Singleton (2001) argue that when the grid becomes infinitely fine ($K \to \infty$), the efficient GMM covariance matrix $\Omega_K$ tends to the Cramer-Rao bound for estimation of $\theta$. 

8
However, as first noticed by Carrasco and Florens (2000), this does not provide a way to estimate efficiently $\theta$ based on the characteristic function since the $2K$ moment conditions $E[g_K(\theta, Y)] = 0$ will suffer from a multicollinearity problem when $K$ goes to infinity. In other words, it does not make sense to think about the limit of $\sum_{K}^{-1}$ when $K$ goes to infinity and one must instead think in terms of covariance operator. Since, for any $t$:

$$|h(t, Y, \theta)| \leq 2,$$

the random function $t \rightarrow h(t, Y, \theta)$ defines a stochastic process $h(Y, \theta)$ which is squared integrable for any probability measure $\Pi$ on $\mathbb{R}$:

$$h(Y, \theta) \in L^2(\mathbb{R}, \Pi).$$

The associated covariance operator $\Omega$ is the linear mapping from $L^2(\mathbb{R}, \Pi)$ to $L^2(\mathbb{R}, \Pi)$ such that, for any $f \in L^2(\mathbb{R}, \Pi)$:

$$\Omega = \int w(t, s)f(s)\Pi(ds)$$

where $w(t, s) = E[h(t, Y, \theta)h(s, Y, \theta)]$. Then under standard regularity conditions, $\sqrt{n}h_n(\theta) = \frac{1}{\sqrt{n}} \sum_{j=1}^{n} h(Y_j, \theta)$ converges in $L^2(\mathbb{R}, \Pi)$ towards a Gaussian process $\mathcal{N}[0, \Omega]$. This paves the way for defining optimal GMM for the continuum of moment conditions of interest. More precisely, the operator $\Omega$ is consistently estimated by the operator $\Omega_n$ with kernel:

$$w_n(t, s) = \frac{1}{n} \sum_{j=1}^{n} h(t, Y_j, \tilde{\theta}_n)h(s, Y_j, \tilde{\theta}_n)$$

where $\tilde{\theta}_n$ is a first-step consistent estimator of $\theta$. Note that $\tilde{\theta}_n$ may be for instance obtained from GMM applied with a finite grid of values of $t$.

Intuitively, efficient GMM based on the whole continuum of moment conditions would amount to minimize $\|\Omega^{-1}h_n(\theta)\|$ with $\Omega$ replaced by its consistent estimator $\Omega_n$. However, since $\Omega$ is a compact operator, its eigenvalues converge to zero and the operator $\Omega^{-1}$ is not continuous. Then a regularization scheme is needed. For instance, a Tykhonov regularized inverse of $\Omega_n$ is defined by:

$$\left(\Omega_n^{(\beta_n)}\right)^{-1} = (\beta_n Id + \Omega_n^2)^{-1}\Omega_n$$
where \( Id \) is the identity operator and \( \beta_n > 0 \) is a penalization parameter. Then, Carrasco and Florens (2000) propose to estimate \( \theta \) as:

\[
\hat{\theta}_n = \operatorname*{arg\,min}_{\theta \in \Theta} \left\| \left( \Omega_n^{(\beta_n)} \right)^{-1/2} \hat{h}_n(\theta) \right\|.
\]

They show (see also Carrasco, Chernov, Florens and Ghysels (2001)) that for a sequence \( \beta_n, n \in \mathbb{N} \), of regularization parameters such that \( \beta_n \to \infty \) but \( n\beta_n^{5/2} \to \infty \) when \( n \to \infty \), \( \hat{\theta}_n \) is not only optimal among GMM estimators but reaches the Cramer-Rao efficiency bound. It may then provide a way to reach the efficiency bound more easily than with MLE from a numerical point of view.

Indeed, the computational tractability of the efficient GMM estimator \( \hat{\theta}_n \) is tightly related to the way used to compute the sequence of operators \( \left( \Omega_n^{(\beta_n)} \right)^{-1/2} \). Carrasco, Florens and Renault (2003) provide a survey of estimation methods based on spectral decomposition and regularization. However, for this particular example, Carrasco, Chernov, Florens and Ghysels (2001) give a way to compute the objective function \( \left\| \left( \Omega_n^{(\beta_n)} \right)^{-1/2} \hat{h}_n(\theta) \right\|^2 \) without resorting to any spectral decomposition. They show that:

\[
\left\| \left( \Omega_n^{(\beta_n)} \right)^{-1/2} \hat{h}_n(\theta) \right\|^2 = v(\theta)' \left[ Id_n - C \left[ \beta_n Id_n + C^2 \right] C \right] v(\theta)
\]

where \( C \) is a \( n \times n \) matrix with \((i, j)\) element \( c_{ij} \), \( Id_n \) is the \( n \times n \) identity matrix and \( v(\theta) = (v_1(\theta), \cdots, v_n(\theta))' \) with:

\[
v_i(\theta) = \int h(t, Y_i, \tilde{\theta}_n) \hat{h}_n(t, \theta) \Pi(dt)
\]

and

\[
c_{ij} = \frac{1}{n} \int h(t, Y_i, \tilde{\theta}_n)h(t, Y_j, \tilde{\theta}_n) \Pi(dt).
\]

Note however that the theoretical asymptotic result does not indicate how to select the penalization parameter \( \beta_n \) in practice. A data-driven method may be desirable (see Carrasco and Florens (2000)).

### 3 Alternative moment matching methods

Following Gallant and Tauchen (1999), we can use the terminology Conventional Method of Moments (CMM) to refer to all variants of the minimum chi-squared estimator implemented using polynomial moment functions. The problem with stable distributions
is that polynomial functions are not integrable and one must find other moments to match.

While the efficient methods described in section 2 are fairly involved because the moments to match are either defined through the computationally intractable likelihood function or through the whole characteristic function, we consider in this section several alternative moments-based estimation methods which are easier to implement. However, in order to remain as close as possible to efficiency, the moments to match must be well focused on the parameters of interest. We are going to sketch below three categories of methods which look well-suited to provide informative moments to match.

First, as proposed by McCulloch (1986) extending an idea of Fama and Roll (1971), sample counterparts of the cumulative distribution function (or equivalently of some quantiles) are much easier to deal with than the likelihood function, while possibly keeping its informational content. After all, indicator functions of one-sided subsets of the real line define a versatile basis of integrable functions, the expectations of which define the distribution function. The contribution of McCulloch (1986) is to exhibit four specific functions of empirical quantiles which, precisely because each of them heavily depends upon the value of one of the four parameters of interest (irrespective of the value of the three others), will give us a strong fix on the parameters of the stable distribution.

Second, instead of matching some arbitrarily selected values of the characteristic function, it may be much more informative to use well-focused summaries of the empirical characteristic function. Koutrouvelis (1980) has precisely shown that the known functional form of the characteristic function with respect to the parameters \( \alpha, \beta, \mu \) and \( \sigma \) suggests some regression-based summaries of a set of values \( \Psi_\theta(t_k), k = 1 \ldots K \), which are well informative about \( \theta = (\alpha, \beta, \mu, \sigma)' \).

Finally, the more recent indirect inference literature (Smith (1993), Gouriéroux, Monfort and Renault (1993), Gallant and Tauchen (1996)) suggests to look for moments to match through the quasi-likelihood function of an auxiliary model. While Gallant and Tauchen (1999) provide evidence of the superiority of EMM (method of moments implemented with a seminonparametric auxiliary model) over CMM, the situation is quite different here. First, as explained above, CMM is meaningless due to fat tails. Second fat tails may also invalidate the efficiency argument of EMM since there is no more reason to hope that a seminonparametric (SNP) score generator based on Hermite expansions will be able to span the true score function. The class of densities to fit with SNP considered by Coppejans and Gallant (2002) are indeed weighted with the exponential function \( \exp \left( -\frac{x^2}{\tau} \right) \) which ensures finite moments at any order. This feature is at odd with stable distributions. While other weight functions may possibly be imagined to improve the fit with a SNP family, we choose to focus here on a specific parametric
family of distributions which should be well informative about the four parameters of interest. Since, in addition, we will be led later to put forward the crucial importance of imposing some constraints on the auxiliary parameters, we expect that our constrained indirect approach is safer in this context than an SNP approach involving a possibly infinite number of auxiliary parameters.

3.1 Quantile-based methods

Let us denote by $x_p$ the $p$-th population quantile of $S(\alpha, \rho, \mu, \sigma)$:

$$Y \sim S(\alpha, \beta, \mu, \sigma) \implies P[Y < x_p] = p.$$ 

From (2.2), it is clear that for any different values \(p, q, p', q' \in ]0,1[\):

$$\frac{x_p - x_q}{x_{p'} - x_{q'}}$$

is independent of both $\mu$ and $\sigma$.

McCulloch (1986) idea is to define two functions of such ratios, that is two functions $\phi_1(\alpha, \beta)$ and $\phi_2(\alpha, \beta)$, which will allow to back out both $\alpha$ and $\beta$. Moreover, to get accurate estimators, it is intuitively well-suited to define the first auxiliary parameter $\phi_1 = \phi_1(\alpha, \beta)$ as well focused on $\alpha$ (for each $\beta$) which the second auxiliary parameter $\phi_2 = \phi_2(\alpha, \beta)$ is well focused on $\beta$ (for each $\alpha$). The definition of the two auxiliary parameters $(\phi_1, \phi_2)$ must then be tightly related to the interpretation of the two structural parameters $(\alpha, \beta)$ they are supposed to be informative about. Therefore, McCulloch (1986) proposes to define $\phi_1$ as a measure of the relative sizes of the tails and the middle of the distribution:

$$\phi_1(\alpha, \beta) = \frac{x_{0.95} - x_{0.05}}{x_{0.75} - x_{0.25}} \quad (3.1)$$

A larger $\phi_1$ means fatter tails and then a smaller $\alpha$. McCulloch (1986) remarks that, since $\phi_1(\alpha, \beta)$ is a strictly decreasing function of $\alpha$, for each $\beta$, the estimation of $\phi_1$ will give us a strong fix on $\alpha$. The function $\phi_2$ is defined as a measure of the spread between the right part and the left part of the distribution:

$$\phi_2(\alpha, \beta) = \frac{(x_{0.95} - x_{0.5}) - (x_{0.5} - x_{0.05})}{x_{0.95} - x_{0.05}} \quad (3.2)$$

A larger $\phi_2$ means more weight on the right side and thus a larger $\beta$. In other words, since $\phi_2(\alpha, \beta)$ is a strictly increasing function of $\beta$, for each $\alpha$, the estimation of $\phi_2$ will be very informative about $\beta$.
Therefore, the proposed estimation strategy may be to replace population quantiles $x_p$ by their sample counterparts $\hat{x}_p$ and to define estimators $\hat{\alpha}$ and $\hat{\beta}$ of the structural parameters as solutions of:

\[
\begin{align*}
\phi_1(\hat{\alpha}, \hat{\beta}) &= \hat{\phi}_1 \\
\phi_2(\hat{\alpha}, \hat{\beta}) &= \hat{\phi}_2
\end{align*}
\]

(3.3)

where:

\[
\begin{align*}
\hat{\phi}_1 &= \frac{\hat{x}_{0.95} - \hat{x}_{0.05}}{\hat{x}_{0.75} - \hat{x}_{0.25}} \\
\hat{\phi}_2 &= \frac{\hat{x}_{0.95} + \hat{x}_{0.05} - 2\hat{x}_{0.5}}{\hat{x}_{0.95} - \hat{x}_{0.05}}.
\end{align*}
\]

(3.4)

McCulloch (1986) provides some tables of possible values of the functions $\phi_1(\alpha, \beta)$ and $\phi_2(\alpha, \beta)$ for a grid of values of $(\alpha, \beta)$ to solve approximately equations (3.3). The tables contain the values $\hat{\alpha}, \hat{\beta}$ of structural parameters for given estimated values $\hat{\phi}_1, \hat{\phi}_2$ of auxiliary parameters. Of course, as detailed at the end of this subsection, these equations can be solved even more precisely through a simulation-based procedure.

To implement this estimation strategy successfully, McCulloch (1986) adds that the sample quantile must be suitably corrected for continuity. Without such a correction, spurious skewness will appear to be present in finite samples. He also sets the smallest possible value of $\phi_1(\alpha, \beta)$ to 2.439, when $\alpha$ increases to 2, irrespective of the value of $\beta$. Of course, in finite sample, $\frac{\hat{x}_{0.95} - \hat{x}_{0.05}}{\hat{x}_{0.75} - \hat{x}_{0.25}}$ may be less than 2.439, and then be off-scale in corresponding tables. Therefore, the definition of the first auxiliary parameter must be slightly modified to incorporate the relevant constraint (as if it was not guaranteed by the stable distribution):

\[
\phi_1 = \begin{cases} 
\frac{x_{0.95} - x_{0.05}}{x_{0.75} - x_{0.25}} & \text{if it is more than 2.439} \\
2.439 & \text{otherwise.}
\end{cases}
\]

The sample counterpart $\hat{\phi}_1$ is defined accordingly and (3.3) is solved from this definition. Note that, when $\hat{\phi}_1 = 2.439$, $\hat{\alpha} = 2$ and $\hat{\beta}$ is not identified. Since, by (2.2), for any $p, q \in [0, 1]$, $x_p - x_q$ is independent of $\mu$ and proportional to $\sigma$, it is natural to define an estimator of the scale parameter $\sigma$ as:

\[
\hat{\sigma} = \frac{\hat{x}_{0.75} - \hat{x}_{0.25}}{\hat{\phi}_3}
\]

(3.5)
(the choice $p = 0.75$ and $q = 0.25$ is intuitively well informative), where the auxiliary parameter $\phi_3 = \phi_3 (\alpha, \beta)$ is defined by:

$$\phi_3 (\alpha, \beta) = \frac{x_{0.75} - x_{0.25}}{\sigma},$$

a quantity which depends neither on $\mu$ nor $\sigma$ and can be tabulated for a grid of values of $(\alpha, \beta)$. The needed estimation $\hat{\phi}_3$ of the auxiliary parameter $\phi_3$ is deduced from the previous estimation (3.3) of $(\alpha, \beta)$:

$$\hat{\phi}_3 = \phi_3 (\hat{\alpha}, \hat{\beta}). \quad (3.6)$$

Note that $\hat{\phi}_3$ will intuitively inform us best about $\sigma$ if it estimates a coefficient of proportionality between $(x_{0.75} - x_{0.25})$ and $\sigma$ that is almost independent of $(\alpha, \beta)$. The table of values of $\phi_3 (\alpha, \beta)$ provided by McCulloch (1986) at least confirms that it does not depend much on $\beta$. The situation is less favourable concerning $\alpha$.

Finally, to back out the location parameter $\mu$, it is natural to locate it with respect to the median $x_{0.5}$ of the distribution through a standardized spread $\frac{\mu - x_{0.5}}{\sigma}$ which is, by (2.2), a function of $(\alpha, \beta)$ independent of $\mu$ and $\sigma$. Unfortunately, although well defined when $\alpha = 1$, this function goes to $(-\infty)$ (resp. $+\infty$), for all $\beta$, when $\alpha$ goes to 1 by smaller (resp. larger) values. McCulloch (1986) advocates a result of Zolotarev (1954) to claim that a convenient way to erase the discontinuity of this function at $\alpha = 1$ is to modify its definition as

$$\phi_4 (\alpha, \beta) = \frac{\mu - x_{0.5}}{\sigma} + \beta \tan \left( \Pi \frac{\alpha}{2} \right) \quad (3.7)$$

knowing that

$$\phi_4 (1, \beta) = \frac{\mu - x_{0.5}}{\sigma} = \lim_{\alpha \rightarrow 1} \phi_4 (\alpha, \beta) \quad (3.8)$$

From the estimation $\hat{\phi}_4 = \phi_4 (\hat{\alpha}, \hat{\beta})$, we then deduce the estimator of $\mu$ from previously defined estimators of $(\alpha, \beta, \sigma)$:

$$\hat{\mu} = \hat{x}_{0.5} + \hat{\sigma} \left[ \hat{\phi}_4 - \hat{\beta} \tan \left( \Pi \frac{\hat{\alpha}}{2} \right) \right] \quad (3.9)$$

To summarize, the estimators proposed by McCulloch (1986) for the structural parameters $\theta = (\alpha, \beta, \mu, \sigma)'$ can be seen as a particular case of indirect inference estimators as defined by Smith (1993) and Gourieroux, Monfort and Renault (1993).
Indeed, everything starts from a summary of the data sample through the estimator
\( \hat{\Psi} = (\hat{\Psi}_1, \hat{\Psi}_2, \hat{\Psi}_3, \hat{\Psi}_4) \) of a vector \( \Psi = \text{plim} \hat{\Psi} \) of auxiliary parameters. This estimator is defined by:

\[
\begin{align*}
\hat{\Psi}_1 &= \max \left[ \frac{\hat{x}_{0.95} - \hat{x}_{0.05}}{\hat{x}_{0.75} - \hat{x}_{0.25}}, 2.439 \right] \\
\hat{\Psi}_2 &= \frac{\hat{x}_{0.95} - \hat{x}_{0.05}}{\hat{x}_{0.05} - 2\hat{x}_{0.05}} \\
\hat{\Psi}_3 &= f [\hat{x}_{0.05}, \hat{x}_{0.25}, \hat{x}_{0.5}, \hat{x}_{0.75}, \hat{x}_{0.95}] \\
\hat{\Psi}_4 &= g [\hat{x}_{0.05}, \hat{x}_{0.25}, \hat{x}_{0.5}, \hat{x}_{0.75}, \hat{x}_{0.95}]
\end{align*}
\]

where \( f \) and \( g \) are known functions defined by:

\[
\begin{align*}
f [\hat{x}_{0.05}, \hat{x}_{0.25}, \hat{x}_{0.5}, \hat{x}_{0.75}, \hat{x}_{0.95}] &= \frac{\hat{x}_{0.75} - \hat{x}_{0.25}}{\phi_3 (\hat{\alpha}, \hat{\beta})} \\
g [\hat{x}_{0.05}, \hat{x}_{0.25}, \hat{x}_{0.5}, \hat{x}_{0.75}, \hat{x}_{0.95}] &= \hat{x}_{0.5} + (\hat{x}_{0.75} - \hat{x}_{0.25}) \frac{\phi_4 (\hat{\alpha}, \hat{\beta}) - \hat{\beta} \tan \left( \frac{\Pi \hat{\alpha}}{2} \right)}{\phi_3 (\hat{\alpha}, \hat{\beta})}
\end{align*}
\]

where \( \hat{\alpha} \) and \( \hat{\beta} \) are the functions of \( \hat{x}_{0.05}, \hat{x}_{0.25}, \hat{x}_{0.5}, \hat{x}_{0.75} \) and \( \hat{x}_{0.95} \) defined as solutions of (3.3).

Then, to back out the indirect inference estimator \( \hat{\theta} \) of structural parameters \( \theta = (\alpha, \beta, \mu, \sigma)' \), from the estimator \( \hat{\Psi} \) of auxiliary parameters, one has just to invert the binding function \( \Psi \) defined as:

\[
\Psi (\theta) = \begin{bmatrix}
\phi_1 (\theta_1, \theta_2) \\
\phi_2 (\theta_1, \theta_2) \\
\phi_3 (\theta_1, \theta_2) \\
\phi_4 (\theta_1, \theta_2)
\end{bmatrix}.
\]

(3.10)

It turns out that this binding function is already known from tables of values of \( \phi_i (\alpha, \beta), i = 1, 2, 3, 4 \) provided by McCulloch (1986). However, the resulting strategy is conformable to the general indirect inference strategy of recovering this binding function through simulations in the structural model. Notice that these simulations can be done for a grid of values of \( (\theta_1, \theta_2) = (\alpha, \beta) \), for given \( (\theta_3, \theta_4) = (\mu, \sigma) \) (for instance \( \mu = 0 \) and \( \sigma = 1 \)) since the effect of these location-scale parameters inside the binding function is known in closed form.
The quantile-based estimators proposed by McCulloch (1986) are generally considered to be quite accurate, but not efficient. In order to assess the quality of these estimators, notice that they define a consistent asymptotically normal estimator \( \hat{\theta} \) as a function of the consistent asymptotically normal sample counterpart of a vector of five quantiles:

\[
\gamma = (x_{0.05}, x_{0.25}, x_{0.5}, x_{0.75}, x_{0.95}).
\]

Therefore, a standard indirect inference strategy could also be applied through the overidentified binding function:

\[
\gamma = \Gamma(\theta).
\] (3.11)

In some respect, resorting to the more parsimonious vector \( \Psi \) of auxiliary parameters (instead of \( \gamma \)) is motivated by the fact that, following McCulloch (1986), we already have some intuition about the right way to solve (3.11). Instead of dealing with it through a blind simulated minimum chi-square procedure, we prefer to work with (3.10), where, as explained by McCulloch (1986), each component \( \Psi_i \) of \( \Psi \) is conceived to give us a strong fix on the corresponding component \( \theta_i \) of \( \theta \).

A second important remark is the additional constraint on \( \phi_1 \) (or \( \Psi_1 \)) introduced by McCulloch (1986). Although immaterial asymptotically when the unknown true value \( \alpha^0 \) of \( \alpha \) is supposed to lie in the open interval \( ]0, 2[ \), this constraint may play a role in finite sample. Although needed, as well explained by McCulloch (1986), this constraint introduces an identification problem. When \( \hat{\Psi}_1 \) is stuck on the value 2.439, the sample is finally characterized by a three-dimensional parameter \( (\hat{\Psi}_2, \hat{\Psi}_3, \hat{\Psi}_4) \) which does not allow to identify the four unknown structural parameters. A relevant solution for this problem is to use the constrained indirect Inference (CII) theory, as recently proposed by Calzolari, Fiorentini and Sentana (2004). The idea is to replace the lacking fourth auxiliary parameter by the value of the Kuhn-Tucker multiplier associated with the constraint. This will be our chosen strategy in section 4, with however an alternative set of auxiliary parameters, as put forward in subsection 3.3 below.

### 3.2 Regression-based methods

Instead of extracting information about the parameters from quantiles, one can use other implications from the characteristic function. Koutrouvelis (1980) proposes a regression-type estimation method of the parameters of the stable law. Starting with the usual expression for the characteristic function (2.1), one can deduce a set of equations:

\[
\log \left( -\log |\phi(t)|^2 \right) = \log (2\sigma^\alpha) + \alpha \log |t|
\] (3.12)
and (for $\alpha \neq 1$):

\[ \begin{align*}
\text{Re} \phi(t) &= \exp\left(-|\sigma t|^\alpha\right) \cdot \cos \left[ \mu t - |\sigma t|^\alpha \beta \text{sgn}(t) \tan\left(\frac{\Pi \alpha}{2}\right) \right] \\
\text{Im} \phi(t) &= \exp\left(-|\sigma t|^\alpha\right) \cdot \sin \left[ \mu t - |\sigma t|^\alpha \beta \text{sgn}(t) \tan\left(\frac{\Pi \alpha}{2}\right) \right]
\end{align*} \tag{3.13} \]

for the real and imaginary parts of $\phi(t)$.

The two equations in (3.13) lead to:

\[ \arctan \left( \frac{\text{Im} \phi(t)}{\text{Re} \phi(t)} \right) = \mu t - \beta \sigma^\alpha \tan\left(\frac{\alpha \Pi}{2}\right) \text{sgn}(t) |t|^\alpha \]

Equation (3.12) suggests a regression of $y = \log(-\log|\phi t|^2)$ on $w = \log |t|$:

\[ y = m + \alpha w_k + \varepsilon_k, \quad k = 1, 2, \ldots, K \]

where $(t_k; k = 1, 2, \ldots, \kappa)$ is an appropriate set of real numbers and $m = \log(2\sigma^\alpha)$. This regression model provides estimates of $\sigma$ and $\alpha$. Given these estimates, one can use the regression model:

\[ z_l = \mu u_l - \beta \sigma^\alpha \tan\left(\frac{\Pi \alpha}{2}\right) \text{sgn}(u_l) |u_l|^\alpha + \eta_l \quad l = 1, 2, \ldots, L \]

where $(u_l; l = 1, 2, \ldots, K)$ is an appropriate set of real numbers, to obtain estimates of $\mu$ and $\beta$.

Therefore, this two-step procedure provides estimates of the four parameters of the stable law through two well-chosen functions based on the characteristic function and uses the sample characteristic function to obtain the estimates. According to the simulation results of Koutrouvelis (1980), this regression method is better than other methods based on moments (Press, 1972) or the minimization of a normalized distance between the empirical and the theoretical characteristic functions. Akgiray and Lamoureux (1989) provide a simulation study which compares the regression method to the quantile method of McCulloch (1986). The results indicate that both the fractile method

---

3 The function $z$ is equal to $\arctan\left(\frac{\text{Im} \phi_n(u)}{\text{Re} \phi_n(u)}\right) + \pi k_n(u)$ where $\arctan$ denotes the principal value of the arctan function and the integer $k_n(\mu)$ accounts for possible nonprincipal branches of the arctan function.

4 Koutrouvelis (1980) describes several refinements of the procedure by introducing certain standardizations to the data and by approximately choosing the points $t_k$ and $u_{\ell}$. 

17
and the regression method provide accurate estimates of the characteristic exponent $\alpha$. However, they note that in general the estimates of the skewness parameter $\beta$ are not as good as the estimates of index $\alpha$. The mean squared errors as well as the biases for both methods are relatively large. This is especially true when $\alpha$ is close to 2, as already explained in the previous section. The regression method does not therefore improve significantly over the McCulloch (1986) method.5

### 3.3 Quasi-likelihood-based method

To be better informed about the four parameters of interest $(\alpha, \beta, \mu, \sigma)$, it seems intuitively preferable to go through a quasi-likelihood function which entails similar parameters with similar interpretations. Therefore, we propose in this subsection to focus on the family of skewed-Student distributions as introduced by Fernandez and Steel (1998) (see also Hansen (1994)).

Let us consider the skewed-$t$ quasi-likelihood function:

$$l(y; \nu, \gamma, \omega, \lambda) = \frac{h(\nu)}{\sqrt{\nu}} \frac{1}{\lambda^{\nu + \frac{1}{2}} \Gamma\left(\frac{\nu + 1}{2}\right)} \left\{ 1 + \frac{1}{\nu} \left(\frac{y - \omega}{\lambda}\right)^2 g_\omega(y, \gamma) \right\}^{-\frac{\nu + 1}{2}}$$

where:

$$h(\nu) = \frac{2\Gamma\left(\frac{\nu + 1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)}$$

and

$$g_\omega(y, \gamma) = \begin{cases} \frac{1}{\gamma} & \text{if } y \geq \omega \\ \gamma^2 & \text{if } y < \omega. \end{cases}$$

While the possibly non integer degrees of freedom $\nu$ of a Student distribution capture the thickness of the tails as $\alpha$ does for stable distributions, location $\omega$ and scale parameter $\lambda$ can easily be introduced to match the two parameters $\mu$ and $\sigma$. Finally, the skewed-$t$ extension allows one to accommodate skewness through an additional parameter $\gamma$ which should be well informative about $\beta$.

We then define a vector $\phi(\alpha, \beta, \mu, \sigma)$ of four auxiliary parameters:

$$\begin{align*}
\phi_1(\alpha, \beta, \mu, \sigma) &= \nu \\
\phi_2(\alpha, \beta, \mu, \sigma) &= \gamma \\
\phi_3(\alpha, \beta, \mu, \sigma) &= \omega \\
\phi_4(\alpha, \beta, \mu, \sigma) &= \lambda
\end{align*}$$

5However, they also provide bootstrapping results based on samples drawn from stock-market data and recommend the regression method based on these results.
characterized by:

\[ \phi(\alpha, \beta, \mu, \sigma) = \arg \max_{(\nu, \gamma, \omega, \lambda)} E[\log l(Y; \nu, \gamma, \omega, \lambda)] \]

when \( Y \sim S(\alpha, \beta, \mu, \sigma) \).

In other words, \( \phi(\alpha, \beta, \mu, \sigma) \) defines the pseudo-true value of the skewed-t parameters \((\nu, \gamma, \omega, \lambda)\) when the true probability distribution is the stable one \( S(\alpha, \beta, \mu, \sigma) \). We claim that these auxiliary parameters \( \phi(\theta) \) will be very informative about the corresponding structural parameters \( \theta = (\alpha, \beta, \mu, \sigma) \). The binding function will not only be one-to-one but will remain true to the intuitive associations: \( \nu \leftrightarrow \alpha, \gamma \leftrightarrow \beta, \omega \leftrightarrow \mu, \lambda \leftrightarrow \sigma \). To see this, we prove four results.

**Proposition 3.1:** For any real number \( a \), \( \phi_3(\alpha, \beta, \mu + a, \sigma) = \phi_3(\alpha, \beta, \mu, \sigma) + a \), that is in short: \( \omega(\mu + a) = \omega(\mu) + a \)

Proposition 3.1 confirms that the auxiliary parameter \( \omega = \phi_3 \) should inform us very well on the location parameter \( \mu \).

**Proposition 3.2:** For any \( a > 0 \), \( \phi_4(\alpha, \beta, a\mu, a\sigma) = a\phi_4(\alpha, \beta, \mu, \sigma) \) that is, in short: \( \lambda(a\sigma) = a\lambda(\sigma) \).

Proposition 3.2 confirms that the auxiliary parameter \( \lambda = \phi_4 \) should be very informative about the scale parameter \( \sigma \).

**Proposition 3.3:**

\[ \phi_2(\alpha, -\beta, \mu, \sigma) = [\phi_2(\alpha, \beta, \mu, \sigma)]^{-1} \]

that is, in short:

\[ \gamma(-\beta) = [\gamma(\beta)]^{-1} \]

Proposition 3.3 confirms that the auxiliary parameter \( \gamma = \phi_3 \) should capture the skewness parameter \( \beta \).
Proposition 3.4: \( \nu = \phi_1 (\alpha, \beta, \mu, \sigma) \) is determined as solution of:

\[
\frac{h' (\nu)}{h (\nu)} = \frac{1}{2} E \left\{ \log \left[ 1 + \frac{1}{\nu} \left( \frac{Y - \omega}{\lambda} \right)^2 g_\omega (Y, \gamma) \right] \right\}
\]

where \( Y \sim S (\alpha, \beta, \mu, \sigma) \) and \( (\nu, \gamma, \omega, \lambda) = \phi (\alpha, \beta, \mu, \sigma) \)

In particular:

\[
\begin{align*}
\phi_1 (\alpha, \beta, \mu + a, \sigma) &= \phi_1 (\alpha, \beta, \mu, \sigma) \quad \text{for all } a, \\
\phi_1 (\alpha, \beta, a\mu, a\sigma) &= \phi_1 (\alpha, \beta, \mu, \sigma) \quad \text{for all } a > 0, \\
\phi_1 (\alpha, -\beta, \mu, \sigma) &= \phi_1 (\alpha, \beta, \mu, \sigma)
\end{align*}
\]

In short:

\[
\begin{align*}
\nu (\mu + a) &= \nu (\mu) \\
\nu (a\sigma) &= \nu (\sigma) \\
\nu (-\beta) &= \nu (\beta)
\end{align*}
\]

Proposition 3.4 confirms that the auxiliary parameter \( \nu = \phi_1 \) should correspond to the tail parameter \( \alpha \). In particular it is not modified by symmetry, location and scale changes.

The nice correspondence between the two set of parameters suggests that pseudo-maximum likelihood estimators \( \hat{\nu}, \hat{\gamma}, \hat{\omega}, \hat{\lambda} \) of the skewed-\( t \) parameters should be very informative about the structural parameters \( (\alpha, \beta, \mu, \sigma) \). Indirect inference estimators \( \hat{\alpha}, \hat{\beta}, \hat{\mu}, \hat{\sigma} \) of the latter could be simply computed as solutions of the following equations

\[
\begin{align*}
\hat{\nu} &= \phi_1 (\hat{\alpha}, \hat{\beta}, \hat{\mu}, \hat{\sigma}) \\
\hat{\gamma} &= \phi_2 (\hat{\alpha}, \hat{\beta}, \hat{\mu}, \hat{\sigma}) \\
\hat{\omega} &= \phi_3 (\hat{\alpha}, \hat{\beta}, \hat{\mu}, \hat{\sigma}) \\
\hat{\lambda} &= \phi_4 (\hat{\alpha}, \hat{\beta}, \hat{\mu}, \hat{\sigma})
\end{align*}
\]

Of course, the binding functions \( \phi_1, \phi_2, \phi_3, \) and \( \phi_4 \) are not known in closed form and must be recovered though simulations. However, in finite sample, and most likely for larger values of the true unknown \( \alpha \), \( \hat{\nu} \) may be off the theoretical range corresponding to stable distributions exactly as \( \phi_1 \) could be off the range in the previously described quantile-based approach. To see this, note that since \( \lim_{\nu \to +\infty} h' (\nu) = 0 \), \( \nu \to +\infty \) is always a solution of the equation in proposition 3.4 which defines the pseudo-true value.
of \( \nu \). Intuitively, for \( \alpha \) close to 2, we may expect that observed data will give the spurious feeling that variance is finite, which would imply a normal distribution corresponding to \( \nu = +\infty \) in a Student framework. This is why we will constrain the auxiliary parameter \( \phi_1 \) by imposing on it an upper bound as McCulloch (1986) did for his auxiliary parameter which provided information about \( \alpha \). We choose to impose \( \nu \leq 2 \) that is to redefine \( \phi_1 \) as \( \phi_1 = \min(\nu, 2) \).

However, as already mentioned for the quantile-based example, such a constraint about one auxiliary parameter may cause some lack of identification when it is stuck on its limit value. Constrained indirect inference provides the right methodology to deal with this problem.

## 4 Constrained indirect estimators

Following Calzolari, Fiorentini and Sentana (2004), we consider the Lagrangian function:

\[
Q_n(\Psi) = \ln(\phi) + \delta (2 - \nu)
\]

where:

\[
\Psi = (\phi', \delta')' \quad \text{and} \quad \phi = (\nu, \gamma, \omega, \lambda)
\]

is the vector of auxiliary parameters corresponding to the skewed-\( t \) quasi-likelihood function:

\[
\ln(\phi) = n \ln \left[ \frac{h(\nu)}{\sqrt{\nu}} \cdot \frac{1}{\lambda (\gamma + \frac{1}{\gamma})} \right] - \frac{\nu + 1}{2} \sum_{i=1}^{n} \ln \left[ 1 + \frac{1}{\nu} \left( \frac{Y_i - \omega}{\lambda} \right)^2 g_\omega(Y_i, \lambda) \right]
\]

The parameter \( \delta \geq 0 \) is the Kuhn-Tucker multiplier associated with the constraint \( \nu \leq 2 \). In order to accommodate jointly the standard indirect inference estimator and the constrained indirect inference estimator, we consider here the two cases, first with \( \delta \) identical to zero (no constraint on \( \nu \)) and second with \( \delta \geq 0 \) (inequality constraint on \( \nu \)).
The estimator \( \hat{\Psi} \) of the pseudo-true value of \( \Psi \) is then defined by the first-order conditions:

\[
\begin{align*}
\frac{\partial \ln}{\partial (\gamma, \omega, \lambda)} (\hat{\phi}) &= 0 \\
\frac{\partial \ln}{\partial \nu} (\hat{\phi}) &= \hat{\delta}
\end{align*}
\]  

(4.3)

jointly with the complementary slackness restriction:

\[
\hat{\delta} (2 - \hat{\nu}) = 0
\]

(4.4)

Let us denote \( Y^h_i (\theta), h = 1 - H, i = 1, \ldots, n \), the components of \( H \) simulated paths of an \( \alpha \)-stable process for a given value \( \theta = (\alpha, \beta, \mu, \sigma) \) of structural parameters. For simplicity, we maintain here the assumption of i.i.d. observations but everything could be easily extended to \( m \)-dependent stable sequences as considered in Deo (2000).

The simulated path \( (Y^h_i (\theta))_{1 \leq i \leq n} \) defines a simulated criterion function

\[
Q^h_n (\Psi | \theta) = l^h_n (\phi | \theta) + \delta (2 - \nu)
\]

(4.5)

where \( l^h_n (\phi | \theta) \) is computed as in (4.2) but with simulated data \( (Y^h_i (\theta))_{1 \leq i \leq n} \) instead of observed ones \( (Y_i)_{1 \leq i \leq n} \). Corresponding simulated estimators \( \hat{\Psi}^h (\theta) \) are defined by the system of equations:

\[
\begin{align*}
\frac{\partial l^h_n}{\partial (\gamma, \omega, \lambda)} [\hat{\phi}^h (\theta) | \theta] &= 0 \\
\frac{\partial l^h_n}{\partial \nu} [\hat{\phi}^h (\theta) | \theta] &= \hat{\delta}^h (\theta) \\
\hat{\delta}^h (\theta) \cdot (2 - \hat{\nu}^h (\theta)) &= 0
\end{align*}
\]

(4.6)

Let us then consider the average estimator over the \( H \) simulated paths:

\[
\hat{\Psi}_H (\theta) = \frac{1}{H} \sum_{h=1}^{H} \hat{\Psi}^h (\theta)
\]

The main idea of indirect inference is to choose the estimator \( \hat{\theta} \) of structural parameters \( \theta \) in order to match \( \hat{\Psi}_H (\theta) \) against \( \hat{\Psi} \). For standard (unconstrained) indirect inference, we are here in a just identified setting, so that \( \hat{\theta}^u \) is just defined as solution of the system of four equations:

\[
\hat{\phi}^u = \phi_H (\hat{\theta}^u)
\]

(4.7)
The superscripts, \( u \) for unconstrained, are just a reminder that the corresponding estimators have been computed by choosing a zero Kuhn-Tucker multiplier: \( \hat{\delta} \) and \( \hat{\delta}^h(\theta) \) are fixed to zero, for \( h = 1, \ldots, H \). Note that, from Gourieroux, Monfort and Renault (1993), we know that in this just identified setting, the indirect inference estimator \( \hat{\theta}^u \) numerically coincides with the score matching estimator as put forward by Gallant and Tauchen (1996).

By contrast, in order to perform constrained indirect inference, we are faced with a seemingly overidentfied problem since both \( \tilde{\Psi}_H(\theta) \) and \( \hat{\Psi} \) entail five free parameters while the unknown \( \theta \) is of dimension four. However, we know that this overidentification feature is just a finite sample problem since (see Calzolari, Sentana and Fiorentini (2004), Proposition 1), the asymptotic distributions of \( \hat{\Psi} \) and \( \tilde{\Psi}_H(\theta) \) are singular. Therefore, the overidentified finite sample matching problem can be solved by minimizing an arbitrary distance:

\[
\hat{\theta}^c = \arg\min_\theta \left( \tilde{\Psi}_H(\theta) - \hat{\Psi} \right)' W \left( \tilde{\Psi}_H(\theta) - \hat{\Psi} \right)
\]  

(4.8)

In terms of asymptotic probability distribution of \( \hat{\theta}^c \), the choice of the positive definite weighting matrix \( W \) is immaterial. Note however that when \( \hat{\nu} \) is stuck at its limit value 2, the information content of \( \hat{\Psi} \) about the structural parameters \( \theta \) will go through the Kuhn-Tucker multiplier \( \hat{\delta} = \frac{\partial \ln}{\partial \hat{\nu}} \hat{\phi} \). Therefore, constrained indirect inference will not suffer from the weak identification problem about \( \beta \) that is currently encountered with competing estimation methods when the true unknown value of \( \alpha \) is close to 2.

While Calzolari, Fiorentini and Sentana (2004) only derive the asymptotic probability distribution of the constrained indirect inference estimator for an infinite number \( H \) of simulated paths, we do apply it here with finite \( H \). We know, as a general principle of simulated method of moments, that the only asymptotic consequence of this is to multiply the asymptotic variance matrix of \( \hat{\theta} \) by a factor \( 1 + \frac{1}{H} \).

Generally speaking, standard theory of indirect inference (see Gourieroux, Monfort and Renault (1993)) can be applied insofar as first the information content of auxiliary parameters is sufficient to identify the structural parameters and second the estimator \( \tilde{\Psi} \) of auxiliary parameters is root-n asymptotically normal. We show in the appendix that the latter property is fulfilled by \( \tilde{\Psi} \) solution of (4.3). This makes the important difference between our approach and a conventional method of moments. Moreover, while we have chosen to work with just-identified moment conditions, it would be easy to introduce some degree of overidentification, for instance by adding McCulloch (1986) quantile-based auxiliary parameters to our skewed-\( t \) auxiliary parameters.

Then, standard theory of overidentification test in the indirect inference setting will provide \( \chi^2 \)-based goodness of fit tests for stable observations. As mentioned by Deo...
(2000), formal but user-friendly statistical tests for the goodness of fit for stable distributions are not widely available so far. While Deo (2000) proposes some $\chi^2$-tests based on the characteristic function, the indirect inference approach provides an alternative unified framework.

5 A Monte Carlo study

In this section we carry out an extensive Monte Carlo experiment to determine if the good asymptotic properties of the indirect inference estimators with a skewed-t auxiliary model are maintained in a finite sample context. As we have seen in the previous section the asymptotic distribution of $\hat{\theta}_n$ is determined by the asymptotic distribution of $\hat{\Psi}$. Therefore, it is worthwhile to examine the sample distribution of the parameter estimates for the auxiliary model in an experimental setting where we simulate data from a $\alpha$-Stable distribution with different values of the parameters.

We carry out a simulation where we generate 500 samples of 1000 observations for 9 different values of $\alpha$, namely 0.3, 0.7, 1, 1.3, 1.5, 1.7, 1.9, 1.95 and 1.99. We keep the other parameters $\mu$, $\sigma$ and $\beta$ fixed and set them equal to 0, 0.5 and 0 respectively. The simulation experiment is divided in two parts. First we estimate the unconstrained skewed-t distribution, i.e. solving

$$\hat{\phi}^u = \arg \max_{\phi} \ln (Y; \phi),$$

where $\ln (Y; \phi) = \ln l(y; \nu, \gamma, \lambda, \omega)$ and $l(y; \nu, \gamma, \lambda, \omega)$ is as defined as in (4.2). Results are reported in Table 1. Since the true values of $\phi$ are unknown we are only able to draw conclusions comparing the first two moments across different DGPs and looking how far the skewness and kurtosis are from 0 and 3 respectively. For all the values of $\alpha$ the statistics for the location, scale and skewness ($\hat{\omega}$, $\hat{\lambda}$ and $\hat{\gamma}$) are constant in mean and standard deviation and their skewness and kurtosis are close to the gaussian case. Therefore when $\alpha$ approaches its upper bound $\hat{\omega}$, $\hat{\lambda}$ and $\hat{\gamma}$ are still normally distributed even for a sample of 1000 observations.

However $\hat{\nu}$ is ill-behaved, for a finite sample, when $\alpha$ approaches to 2. The mean increases as it can be expected since when $\alpha \to 2$, $\nu \to \infty$; i.e. $\hat{\nu}$ is attracted by $\infty$. The variance increases and the skewness and kurtosis are very far from 0 and 3. Figure 2 represents the kernel densities for the different values of $\alpha$ considered in Table

---

6 In Appendix B, we explain how to simulate from an $\alpha$-Stable distribution
7 We also carried out the same experiment for different values of $\beta$ and obtained similar results. These results can be provided upon request addressed to the corresponding author.
1. For a sample of 1000 observations when $\alpha \geq 1.5$ the estimator $\hat{\nu}$ exhibits serious departures from normality. Notice that for $\alpha$ between 1.3 and 1.95 the densities for $\hat{\nu}$ are more and more peaked and the right tail is fatter, increasing the kurtosis coefficient. This is characteristic of leptokurtic densities. However for $\alpha = 1.99$ the kurtosis is below 3 and the density is much less peaked, but the standard deviation is much larger. This is characteristic of platykurtic densities. Both lepto and platykurtic densities are undesirable. Nonetheless in this context a platykurtic density is a more severe problem as, given the standard deviation, the probability of estimating a $\nu$ very far from its median is certainly large.

As just remarked, the above densities are for a sample size of 1000. We investigate the role of the sample size in Table 2. As $\alpha$ tends to 2 we need a larger sample size to reach normality. For example for $\alpha = 1.9$ and $N = 1000$, $\hat{\nu}$ is far from it (see also Figure 2) and we need a sample size of 10000 for approaching the gaussian distribution. It is even worse when $\alpha$ is closer to 2. In these cases even for $N = 10000$, $\hat{\nu}$ is not normally distributed. This confirms the theoretical results of Appendix A. The rate of convergence of $\hat{\nu}$ is very slow and becomes slower as $\alpha$ tends to 2.

We also show in Figure 3 the link between the degrees of freedom of the auxiliary model and the stability index $\alpha$. We plot the estimated $\hat{\nu}$ as a function of $\alpha$ for a given set of values for $\beta$ (indicated in the legend to the figure). One can conclude from the figure that the relation between $\hat{\nu}$ and $\alpha$ appears to be exponential, confirming that as $\alpha \to 2$ we get closer to a gaussian distribution within the $\alpha$-Stable family and therefore $\hat{\nu} \to \infty$. Moreover, since $\hat{\nu}$ is not always below two, it means that the true process may be an infinite variance process, yet the skewed-t distribution has a finite second moment. Finally, $\beta$ does not have a significant impact on $\hat{\nu}$ as all curves are very close to each other.

To assess the implications of these shortcomings of the auxiliary model on the estimation of the parameter vector $\theta_\alpha$, we generate 500 samples of 1000 observations for different values of $\theta_\alpha$ and estimate for each sample $\hat{\theta}_\alpha$ by indirect inference. The values of $\mu$ and $\sigma$ are set to 0 and 0.5, while $\alpha$ takes the values 1.5 and 1.9 and $\beta$ the values 0 and 0.75. The results, shown in Table 3, indicate that for $\alpha = 1.9$ the skewness and kurtosis for $\mu$ and $\sigma$ tend to explode when $\beta = 0.75$. The distribution of the estimates is much closer to a normal when $\alpha = 1.5$. We conclude from this simulation exercise that the auxiliary model works for most cases but fails when $\alpha$ comes close to 2.

Therefore we propose to use a constrained version of the skewed-t distribution. Since the estimate $\hat{\nu}$ is ill-behaved in finite sample when $\alpha \to 2$ because it is attracted by $\infty$, we impose an upper bound on $\nu$, i.e. $\nu < \nu^c$. We choose $\nu^c = 2$ but in fact even a
larger bound could fulfill the purpose.\footnote{The constrain $\nu < 2$ is sufficient but not necessary, since what matters is to prevent $\hat{\nu}$ from approaching $\infty$. A bound of 10 for example will be as effective although the variance of the estimated distribution will be larger.} We therefore have an additional constraint in the maximization problem which becomes:

$$\hat{\Psi}^c = \arg\max_\psi \ln (Y; \phi) + (\nu - 2) \delta,$$  \hspace{1cm} (5.2)

plus the slackness restriction $(\hat{\nu} - 2) \hat{\delta} = 0$ and the inequality restrictions $\nu \leq 2$.

If $\hat{\nu}^c$, the estimate of $\nu$ under (5.2), is smaller than two the multiplier $\hat{\delta}$ is zero, because of the slackness condition. On the contrary, if $\hat{\nu}^c$ reaches its upper bound, the multiplier is not zero. In the former case implies, the multiplier does not play a role in the estimation and constrained indirect inference is nothing else than indirect inference. In the latter case, it is $\hat{\nu}^c$ who does not play a role and the multiplier is the parameter at work with constrained indirect inference. To verify that the estimated multiplier is well-behaved in finite sample, we draw in Figure 4 the densities of $\hat{\delta}$ when $\alpha$ is getting closer to 2. It can be seen that, contrary to $\hat{\nu}$ (see Figure 2), the distributions are much closer to a normal. The behavior according to the sample size is reported in Table 4.

To assess the performance of the constrained indirect inference method, we will conduct a thorough Monte Carlo study for a number of combinations of values for $\alpha$ and $\beta$, while setting the location and scale parameters, $\mu$ and $\sigma$, to 0 and 0.5 respectively.\footnote{We also carried out the simulation for $\sigma = 1.5$. For space considerations, we do not report the results since the conclusions are basically the same than for $\sigma = 0.5$. Results are available upon request.} We choose four different positive values, namely $\{0.0, 0.25, 0.5, 0.75\}$, for the skewness parameter $\beta$. Finally the stability index $\alpha$ takes the values $\{0.7, 1.1, 1.7, 1.9\}$. We generate 500 samples of 1000 observations for the resulting values of $\theta$. In the indirect inference methods, both constrained and unconstrained, one can choose the number of times ($h$) the simulation is repeated for each estimation in order to reduce the bias. We chose $h = \{1, 2, 5\}$ and report the results in Tables 5, 6 and 7 respectively.\footnote{In Appendix C, we comment on some numerical aspects related to the implementation of the Monte Carlo procedure.} For the weighting matrix $W$, we choose a mixture of 4x4 identity matrices for the cases where $\hat{\nu}$ is equal or greater than 2 or below 2 (that is whether the multiplier is the parameter at work or not). With the 500 estimated parameters we compute some basic statistics: mean, standard deviation, skewness, kurtosis, minimum and maximum, and evaluate if the densities of the estimated parameters depart from normality.

We also compare the constrained indirect inference method to two other methods which have been described in Section 2 and are serious contenders for the estimation of the $\alpha$-Stable distribution. The first one is the continuous GMM method of Carrasco
and Florens (2002) based on the characteristic function with penalized term $\beta_n$ equal to $10^{-6}$. The second one is the empirical quantile method of McCulloch (1986). Results for these two methods are presented in Tables 8 and 9 respectively.

As a general assessment, one can say that the constrained indirect inference method delivers consistent estimators which are close to being distributed normally for all values chosen for $\theta$. The skewness for all parameters is close to 0 and the kurtosis close to 3. Thanks to the constraint imposed on $\nu$ in the auxiliary model, the estimator behaves well even when $\alpha$ approaches 2. However, when $\alpha$ is equal to 1.1 and is therefore close to 1, the parameter $\mu$ is badly estimated since theoretically it becomes infinite. This is a feature which is shared by all estimation methods of the $\alpha$-Stable distribution and confirmed by the two other methods we examined. Increasing the number of simulated draws, $h$, improves the properties of the estimators but not significantly. As expected, the mean of the replications is closer to the true value for $h = 5$ and the standard deviation is smaller.

Constrained indirect inference compares well with the two other methods. First, with respect to continuous GMM, it appears that it estimates much better the parameter $\sigma$. For values of $\alpha$ less than 1, continuous GMM overestimates the value of $\sigma$, while it underestimates for values of $\alpha$ greater than 1. The bias for $\alpha = 1.9$ is quite severe, since the mean of the 500 replications is 0.27 for a true value of 0.5. The bias of indirect inference is also smaller than continuous GMM for the problematic case of $\alpha = 1.1$. Other parameter estimates like $\beta$ suffer when $\alpha$ gets close to 2. This is never the case for indirect inference. The constrained indirect inference method is also more efficient than continuous GMM. The reduction is standard deviation is often by a factor of 2, but in certain cases, say $\alpha = 1.9$ and $\beta = 0.75$, the standard deviation can be almost four times smaller.

The empirical quantile method does not seem to suffer from any systematic bias except for $\beta$ at $\alpha = 1.9$. Its main weakness appears to be its lack of efficiency. Standard deviations are quite larger than in the case of indirect inference and continuous GMM.

This Monte Carlo study shows unequivocally that indirect inference, with its constrained version when necessary, is a reliable method to estimate the parameters of the $\alpha$-Stable distribution.

---

11$10^{-6}$ is the same value that Carrasco and Florens (2000) chose for their Monte Carlo study with the $\alpha$-Stable distribution.
12We use a GAUSS procedure written by J. Huston McCulloch and available in his web page http://www.econ.ohio-state.edu/jhm/jhm.html
13As noted above, continuous GMM has been performed with a fixed ad hoc penalized term. An endogenous choice of $\beta_n$ could improve results, namely suppress the bias for $\sigma$. However, in the Monte Carlo study performed by Carrasco and Florens (2000) for the $\alpha$-Stable distribution, the estimated $\sigma$ does not change significantly when $\beta_n$ is selected in an ad hoc way or endogenously.
\(\alpha\)-Stable distribution and certainly improve in terms of efficiency over classical quantile methods and even over the more recently proposed continuous GMM method.

6 Conclusion

The stable distribution is very useful to model processes with heavy-tailed and skewed distributions which are often encountered in financial series. However, its estimation raises several challenges that we addressed in this paper. Since the density function of a stable distribution does not have a closed form but a stable series is relatively easy to simulate, we proposed an indirect inference estimation method which is ideally suited to such characteristics. In a Monte Carlo study, we showed that the method performed well for almost all values of the parameters and much better than competing methods currently used in terms of efficiency. To improve the properties of the estimator in finite samples when the value of the stability parameter approaches two, we use a variant of the indirect inference method called constrained indirect inference. This new method for estimating stable distributions may prove very useful for financial series exhibiting skewness and kurtosis such as hedge funds returns series or operational loss data. Computation of values at risk based on this indirect inference method will also deliver more reliable estimates.
Appendix A: Proofs

Proof of Propositions 3.1 to 3.4:

Pseudo-true values are defined as maximizing the expectation of the log quasi-likelihood function, that is:

\[ h [(\nu, \gamma, \omega, \lambda); (\alpha, \beta, \mu, \sigma)] = E \left[ \ln l (Y; \nu, \gamma, \omega, \lambda) \right] \]

where \( Y \sim S (\alpha, \beta, \mu, \sigma) \).

We first show that:

\[
\begin{align*}
&h [(\nu, \gamma, \omega + a, \lambda); (\alpha, \beta, \mu + a, \sigma)] = h [(\nu, \gamma, \omega, \lambda); (\alpha, \beta, \mu, \sigma)] \quad \text{for all } a, \\
&h [(\nu, \gamma, a\omega, a\lambda); (\alpha, \beta, a\mu, a\sigma)] = h [(\nu, \gamma, \omega, \lambda); (\alpha, \beta, \mu, \sigma)] \quad \text{for all } a > 0, \\
&h [(\nu, 1/\gamma, 2\mu - w, \lambda); (\alpha, -\beta, \mu, \sigma)] = h [(\nu, \gamma, \omega, \lambda); (\alpha, \beta, \mu, \sigma)]
\end{align*}
\]  

(A.1)

To see this, it is sufficient to notice that:

First, \( Y + a \sim S (\alpha, \beta, \mu + a, \sigma) \) and for all \( y \):

\[ l (y + a; \nu, \gamma, \omega + a, \lambda) = l (y; \nu, \gamma, \omega, \lambda). \]

Second, \( aY \sim S (\alpha, \beta, a\mu, a\sigma), a > 0 \), and for all \( y \):

\[ l (ay; \nu, \gamma, a\omega, \lambda) = l (y; \nu, \gamma, \omega, \lambda). \]

Third, \( (2\mu - Y) \sim S (\alpha, -\beta, \mu, \sigma) \) and for all \( y \):

\[ l (2\mu - y; \nu, 1/\gamma, 2\mu - w, \lambda) = l (y; \nu, \gamma, \omega, \lambda). \]

This achieves to prove (A.1).

Suppose for the moment that we can also prove that, for maximization of the quasi-likelihood function, the first order condition with respect to \( \nu \) is tantamount to the equation stated by proposition 3.4:

\[
\frac{h' (\nu)}{h (\nu)} = \frac{1}{2} E \left\{ \log \left[ 1 + \frac{1}{\nu} \left( \frac{Y - \omega}{\lambda} \right)^2 g_w (Y, \gamma) \right] \right\}
\]

(A.2)
Then, by joint application of (A.1) and (A.2), we see clearly that the changes $Y \rightarrow Y + a$, $Y \rightarrow aY$ and $Y \rightarrow 2\mu - Y$ will have the effects on pseudo-true value that are stated by proposition 3.1 to 3.4. The proof of these propositions will then be completed by the proof of (A.2). To get it, let us write the first order conditions of quasi-likelihood maximization with respect to $\lambda$ and $\nu$. The partial derivatives of the expected Log-quasi-likelihood function with respect to $\lambda$ and $\nu$ are:

$$\frac{\partial h}{\partial \lambda} = -\frac{1}{\lambda} - \frac{\nu + 1}{2\nu} \left( -\frac{2}{\lambda^3} \right) E \left[ \frac{(Y - \omega)^2 g_w(Y, \gamma)}{1 + \frac{1}{\nu} \left( \frac{Y - \omega}{\lambda} \right)^2 g_w(Y, \gamma)} \right]$$

and

$$\frac{\partial h}{\partial \lambda} = \frac{h'(\nu)}{h(\nu)} - \frac{1}{2\nu} - \frac{1}{2} E \left[ \ln \left[ 1 + \frac{1}{\nu} \left( \frac{Y - \omega}{\lambda} \right)^2 g_w(Y, \gamma) \right] \right]$$

$$- \frac{\nu + 1}{2} \left( -\frac{1}{\nu^2} \right) E \left[ \frac{(Y - \omega)^2 g_w(Y, \gamma)}{1 + \frac{1}{\nu} \left( \frac{Y - \omega}{\lambda} \right)^2 g_w(Y, \gamma)} \right].$$

This leads to the following first order conditions, for $\lambda$ and $\nu$ respectively:

$$(\nu + 1) E \left[ \frac{\frac{1}{\nu} \left( \frac{Y - \omega}{\lambda} \right)^2 g_w(Y, \gamma)}{1 + \frac{1}{\nu} \left( \frac{Y - \omega}{\lambda} \right)^2 g_w(Y, \gamma)} \right] = 1 \quad (A.3)$$

and:

$$\frac{h'(\nu)}{h(\nu)} - \frac{1}{2\nu} - \frac{1}{2} E \left[ \ln \left[ 1 + \frac{1}{\nu} \left( \frac{Y - \omega}{\lambda} \right)^2 g_w(Y, \gamma) \right] \right]$$

$$+ \frac{\nu + 1}{2} E \left[ \frac{\frac{1}{\nu} \left( \frac{Y - \omega}{\lambda} \right)^2 g_w(Y, \gamma)}{1 + \frac{1}{\nu} \left( \frac{Y - \omega}{\lambda} \right)^2 g_w(Y, \gamma)} \right] = 0 \quad (A.4)$$

By plugging (A.3) into (A.4), we get the announced first-order conditions (A.2) to characterize the pseudo true value of $\nu$.

Q.E.D

Proof of asymptotic normality

To be completed
Appendix B: Simulating a $\alpha$-Stable distribution

The use of indirect inference or its constrained version necessitates to simulate an $\alpha$-Stable process. For simulating we adopt the method proposed by Chambers, Mallows and Stuck (1976) which is fast and easy to implement.\(^{14}\)

Let $z$ and $y$ two independent random variables, $z$ being uniformly distributed on $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ and $y$ exponentially distributed with mean $1$. When $\alpha \neq 1$,

$$X = \zeta + \frac{\sin \alpha z - \zeta \cos \alpha z}{(\cos z)^{1/\alpha}} \left(\frac{\cos(1-\alpha)z - \zeta \sin(1-\alpha)z}{y}\right)^{\frac{(1-\alpha)}{\alpha}} \sim S_\alpha(\beta, 1, 0), \quad (B.1)$$

where $\zeta = -\beta \tan \frac{\pi \alpha}{2}$. When $\alpha = 1$,

$$X = \frac{2}{\pi} \left[\left(\frac{\pi}{2} + \beta z\right) \tan z - \beta \ln \left(\frac{\pi y \cos z}{\frac{\pi}{2} + \beta z}\right)\right] \sim S_\alpha(\beta, 1, 0). \quad (B.2)$$

To generate $z$ and $y$, we draw two independent uniform$(0, 1)$ random variables $U_1$ and $U_2$ and set $z = \pi (U_1 - \frac{1}{2})$ and $y = -\ln U_2$. Notice that this procedure simulates a process $X$ from a $S_\alpha(\beta, 1, 0)$ distribution. Generating a $S_\alpha(\beta, \sigma, \mu)$ from $S_\alpha(\beta, 1, 0)$ is straightforward using

\[\sigma X + \mu \sim S_\alpha(\beta, \sigma, \mu) \quad \text{if} \quad \alpha \neq 1,\]
\[\sigma X + \frac{2}{\pi} \beta \sigma \ln \sigma + \mu \sim S_\alpha(\beta, \sigma, \mu) \quad \text{if} \quad \alpha = 1.\] \quad (B.3)


Appendix C: Numerical aspects

For the Monte Carlo experiment, we could start the algorithm at the true values of the parameters. However, we also want to propose a practical approach to estimating the parameters of a $\alpha$-Stable distribution. Therefore, we first obtain starting values of the parameters by using an empirical quantile method. In order to reduce the variance of this estimator we bootstrap it, that is we use as initial value of the parameters the mean of the estimates obtained from resampling the series taken as the observed data a certain number of times (10 in this case).

\(^{14}\)The GAUSS procedure for simulating from the $\alpha$-Stable process has been written by J. Huston McCulloch and available in his web page http://www.econ.ohio-state.edu/jhm/jhm.html.
Second, two of the four parameters in $\theta_a$ are constrained. To avoid using a constrained optimization algorithm, we reparametrize the initial parameters. In general if a parameter $\vartheta$ is constrained to belong to a specific interval: $a < \vartheta < b$. Then $0 < \frac{\vartheta - a}{b - a} < 1$ which can be modelled with a logistic function:

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

(C.1)

This means that we can estimate $\xi$, which varies between $-\infty$ and $+\infty$, and then recover $\vartheta$. We apply this transformation to $\alpha$ ($0 < \alpha \leq 2$) and $\beta$ ($-1 \leq \beta \leq 1$). The new parameter set is then $\theta''_a = (\xi_\alpha, \xi_\beta, \sigma, \mu) \in \Theta''_a \subset \mathbb{R}^4$ where $\Theta''_a = [\xi_\alpha, \xi_\beta, \mu \in \mathbb{R}, \sigma \geq 0].$
References


Asymptotic tail behaviours for the $\alpha$-Stable, the Skewed-t and the Gaussian distributions. Parameters are: for the $\alpha$-Stable $\alpha = 1.5$, $\beta = 0$ and $\sigma = 1/\sqrt{2}$. For the Skewed-t $\nu$ is statistically two. Its value is the resulting of an estimation given a simulated sample from an $\alpha$-Stable distribution with the above given parameters.

Figure 1: Asymptotic tail behaviour
Reading in rows. From top to bottom and from left to right the true $\alpha$ are 0.3, 0.7, 1, 1.3, 1.5, 1.7, 1.9, 1.95 and 1.99.

Figure 2: Kernel Densities for $\hat{v}$
Relation between $\alpha$ and $\hat{\nu}$ given that $\beta$ falls in some range of 0.10, in other words each line is
$\hat{\nu} = f(\alpha|\beta \in [\beta_i, \beta_i + 0.10])$ for $\beta_i = \{0, 0.1, 0.2, 0.3, 0.4\}$.

Figure 3: $\hat{\nu}$ sensibility to $\alpha$ and $\beta$
Reading in rows. From top to bottom and from left to right the true $\alpha$ are 1.7, 1.9, 1.95 and 1.99.

Figure 4: Kernel Densities for $\hat{\delta}$
<table>
<thead>
<tr>
<th>$\alpha$</th>
<th>Mean</th>
<th>Sd</th>
<th>Skw</th>
<th>Kur</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.3</td>
<td>0.20941</td>
<td>0.0008</td>
<td>0.0311</td>
<td>0.6414</td>
<td>-0.0645</td>
<td>3.8771</td>
</tr>
<tr>
<td>0.7</td>
<td>0.60229</td>
<td>1.0010</td>
<td>0.0212</td>
<td>0.0378</td>
<td>0.18906</td>
<td>-0.1619</td>
</tr>
<tr>
<td>1</td>
<td>1.00265</td>
<td>0.99967</td>
<td>0.05578</td>
<td>0.03697</td>
<td>0.24542</td>
<td>-0.0262</td>
</tr>
<tr>
<td>1.3</td>
<td>1.85331</td>
<td>1.00515</td>
<td>0.15291</td>
<td>0.04581</td>
<td>0.27424</td>
<td>0.20723</td>
</tr>
<tr>
<td>1.5</td>
<td>2.18338</td>
<td>1.00277</td>
<td>0.20163</td>
<td>0.04515</td>
<td>0.58705</td>
<td>-0.0812</td>
</tr>
<tr>
<td>1.7</td>
<td>3.34125</td>
<td>1.00680</td>
<td>0.46471</td>
<td>0.04708</td>
<td>1.38592</td>
<td>0.14397</td>
</tr>
<tr>
<td>1.9</td>
<td>7.55703</td>
<td>1.00248</td>
<td>3.32274</td>
<td>0.05245</td>
<td>3.42853</td>
<td>0.18672</td>
</tr>
<tr>
<td>1.95</td>
<td>26.9539</td>
<td>1.00651</td>
<td>59.1613</td>
<td>0.05175</td>
<td>4.62974</td>
<td>-0.0390</td>
</tr>
<tr>
<td>1.99</td>
<td>146.564</td>
<td>1.00369</td>
<td>142.759</td>
<td>0.0559</td>
<td>0.49277</td>
<td>-0.0631</td>
</tr>
</tbody>
</table>

The top part of the Table are fits for $\nu$ and $\gamma$, for each statistic the first column is for $\nu$ and the second for $\gamma$. Bottom part are for $\lambda$, first column, and $\tilde{\omega}$, second column. $\mu$, $\beta$ and $\sigma$ are fixed to 0, 0 and 0.5. Sd, Skw, Kur, Min and Max stand for standard deviation, skewness, kurtosis, minimum and maximum.
Table 2: Sensitivity of the finite sample estimates of $\nu$ to changes in the sample size and the stability index $\alpha$

<table>
<thead>
<tr>
<th>$\alpha$</th>
<th>N=100</th>
<th>200</th>
<th>500</th>
<th>1000</th>
<th>5000</th>
<th>10000</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Sd</td>
<td>0.2298</td>
<td>0.1353</td>
<td>0.0828</td>
<td>0.0669</td>
<td>0.0243</td>
</tr>
<tr>
<td></td>
<td>Skw</td>
<td>0.9070</td>
<td>0.6197</td>
<td>0.3841</td>
<td>0.3316</td>
<td>0.0549</td>
</tr>
<tr>
<td>1.3</td>
<td>Sd</td>
<td>0.5489</td>
<td>0.2970</td>
<td>0.1739</td>
<td>0.1171</td>
<td>0.0503</td>
</tr>
<tr>
<td></td>
<td>Skw</td>
<td>2.0973</td>
<td>0.9792</td>
<td>0.6197</td>
<td>0.3542</td>
<td>0.3323</td>
</tr>
<tr>
<td></td>
<td>Kur</td>
<td>11.536</td>
<td>4.8421</td>
<td>3.3634</td>
<td>3.5739</td>
<td>3.0352</td>
</tr>
<tr>
<td>1.5</td>
<td>Sd</td>
<td>21.281</td>
<td>0.5676</td>
<td>0.3381</td>
<td>0.2004</td>
<td>0.0879</td>
</tr>
<tr>
<td></td>
<td>Skw</td>
<td>21.895</td>
<td>2.2189</td>
<td>1.2433</td>
<td>0.8046</td>
<td>0.3184</td>
</tr>
<tr>
<td></td>
<td>Kur</td>
<td>496.06</td>
<td>12.361</td>
<td>7.8171</td>
<td>4.9488</td>
<td>3.8585</td>
</tr>
<tr>
<td>1.7</td>
<td>Sd</td>
<td>62.268</td>
<td>18.792</td>
<td>0.7770</td>
<td>0.4968</td>
<td>0.1931</td>
</tr>
<tr>
<td></td>
<td>Skw</td>
<td>4.7446</td>
<td>19.820</td>
<td>1.3933</td>
<td>0.9682</td>
<td>0.4147</td>
</tr>
<tr>
<td></td>
<td>Kur</td>
<td>23.851</td>
<td>442.04</td>
<td>6.2104</td>
<td>4.5532</td>
<td>3.2662</td>
</tr>
<tr>
<td>1.9</td>
<td>Sd</td>
<td>145.39</td>
<td>165.47</td>
<td>38.823</td>
<td>15.467</td>
<td>0.8323</td>
</tr>
<tr>
<td></td>
<td>Skw</td>
<td>0.8854</td>
<td>2.2574</td>
<td>7.1886</td>
<td>3.0826</td>
<td>0.9205</td>
</tr>
<tr>
<td></td>
<td>Kur</td>
<td>1.8296</td>
<td>6.2549</td>
<td>56.530</td>
<td>16.532</td>
<td>4.4781</td>
</tr>
<tr>
<td>1.95</td>
<td>Sd</td>
<td>158.77</td>
<td>145.75</td>
<td>102.03</td>
<td>58.269</td>
<td>2.5094</td>
</tr>
<tr>
<td></td>
<td>Skw</td>
<td>0.2557</td>
<td>0.9314</td>
<td>2.4688</td>
<td>5.5896</td>
<td>2.7023</td>
</tr>
<tr>
<td></td>
<td>Kur</td>
<td>1.1004</td>
<td>1.5366</td>
<td>7.3776</td>
<td>34.003</td>
<td>19.298</td>
</tr>
<tr>
<td>1.99</td>
<td>Sd</td>
<td>156.38</td>
<td>148.11</td>
<td>150.54</td>
<td>143.28</td>
<td>86.330</td>
</tr>
<tr>
<td></td>
<td>Skw</td>
<td>-0.396</td>
<td>-0.045</td>
<td>0.4966</td>
<td>0.4791</td>
<td>2.3712</td>
</tr>
<tr>
<td></td>
<td>Kur</td>
<td>1.1185</td>
<td>1.2754</td>
<td>1.0784</td>
<td>1.3418</td>
<td>7.4779</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>$\beta$</td>
<td>Mean</td>
<td>Sd</td>
<td>Skw</td>
<td>Kur</td>
<td>Min</td>
</tr>
<tr>
<td>---------</td>
<td>--------</td>
<td>------</td>
<td>-----</td>
<td>-----</td>
<td>------</td>
<td>------</td>
</tr>
<tr>
<td>$\alpha = 1.5$</td>
<td>$\beta = 0$</td>
<td>1.47593</td>
<td>0.01169</td>
<td>0.04659</td>
<td>0.69317</td>
<td>0.03458</td>
</tr>
<tr>
<td>$\beta = 0.75$</td>
<td>1.49288</td>
<td>0.70918</td>
<td>0.05816</td>
<td>0.11778</td>
<td>0.26090</td>
<td>0.57982</td>
</tr>
<tr>
<td>$\alpha = 1.9$</td>
<td>$\beta = 0$</td>
<td>1.82605</td>
<td>0.16182</td>
<td>0.79555</td>
<td>0.21406</td>
<td>-3.1243</td>
</tr>
<tr>
<td>$\beta = 0.75$</td>
<td>1.85128</td>
<td>0.73687</td>
<td>0.08693</td>
<td>0.21192</td>
<td>-0.1666</td>
<td>-1.1151</td>
</tr>
</tbody>
</table>

The top part of the Table are results for $\tilde{\alpha}$ and $\tilde{\beta}$, for each statistic the first column is for $\tilde{\alpha}$ and the second for $\tilde{\beta}$. Bottom part are for $\tilde{\sigma}$ and $\tilde{\mu}$. $\mu$ and $\sigma$ are fixed to 0 and 0.5. Sd, Skw, Kur, Min and Max stand for standard deviation, skewness, kurtosis, minimum and maximum.
Table 4: Sensitivity of the finite sample estimates of \( \rho \) to changes in the sample size and the stability index \( \alpha \)

<table>
<thead>
<tr>
<th>( \alpha )</th>
<th>N=100</th>
<th>200</th>
<th>500</th>
<th>1000</th>
<th>5000</th>
<th>10000</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.7</td>
<td>Sd</td>
<td>0.3967</td>
<td>0.1851</td>
<td>0.0070</td>
<td>0.0045</td>
<td>0.0022</td>
</tr>
<tr>
<td></td>
<td>Skw</td>
<td>4.2254</td>
<td>9.7938</td>
<td>-0.238</td>
<td>-0.180</td>
<td>-0.053</td>
</tr>
<tr>
<td></td>
<td>Kur</td>
<td>18.964</td>
<td>97.537</td>
<td>3.4076</td>
<td>2.7903</td>
<td>2.9022</td>
</tr>
<tr>
<td>1.9</td>
<td>Sd</td>
<td>0.0094</td>
<td>0.0068</td>
<td>0.0044</td>
<td>0.0031</td>
<td>0.0015</td>
</tr>
<tr>
<td></td>
<td>Skw</td>
<td>-0.714</td>
<td>-0.436</td>
<td>-0.297</td>
<td>-0.149</td>
<td>-0.230</td>
</tr>
<tr>
<td></td>
<td>Kur</td>
<td>3.9675</td>
<td>3.4135</td>
<td>2.9608</td>
<td>2.8653</td>
<td>3.3603</td>
</tr>
<tr>
<td>1.95</td>
<td>Sd</td>
<td>0.0084</td>
<td>0.0058</td>
<td>0.0042</td>
<td>0.0028</td>
<td>0.0012</td>
</tr>
<tr>
<td></td>
<td>Skw</td>
<td>-0.585</td>
<td>-0.131</td>
<td>-0.277</td>
<td>-0.252</td>
<td>-0.123</td>
</tr>
<tr>
<td></td>
<td>Kur</td>
<td>3.7321</td>
<td>2.9502</td>
<td>3.1520</td>
<td>3.4036</td>
<td>3.1047</td>
</tr>
<tr>
<td>1.99</td>
<td>Sd</td>
<td>0.0067</td>
<td>0.0053</td>
<td>0.0034</td>
<td>0.0023</td>
<td>0.0010</td>
</tr>
<tr>
<td></td>
<td>Skw</td>
<td>-0.233</td>
<td>-0.406</td>
<td>-0.125</td>
<td>0.0145</td>
<td>-0.007</td>
</tr>
<tr>
<td></td>
<td>Kur</td>
<td>2.9394</td>
<td>3.2607</td>
<td>2.7799</td>
<td>3.2751</td>
<td>3.4280</td>
</tr>
<tr>
<td>( \alpha )</td>
<td>( \beta )</td>
<td>Mean</td>
<td>Sd</td>
<td>Skw</td>
<td>Kur</td>
<td>Min</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>0.7</td>
<td>0</td>
<td>0.69914</td>
<td>-0.0017</td>
<td>0.02525</td>
<td>0.04299</td>
<td>0.33602</td>
</tr>
<tr>
<td>0.25</td>
<td>0.69497</td>
<td>0.24104</td>
<td>0.02538</td>
<td>0.04065</td>
<td>0.28667</td>
<td>0.1459</td>
</tr>
<tr>
<td>0.5</td>
<td>0.69434</td>
<td>0.49150</td>
<td>0.02610</td>
<td>0.04185</td>
<td>0.07106</td>
<td>-0.0683</td>
</tr>
<tr>
<td>0.75</td>
<td>0.69014</td>
<td>0.73893</td>
<td>0.02485</td>
<td>0.03611</td>
<td>0.13553</td>
<td>-0.0905</td>
</tr>
<tr>
<td>1.1</td>
<td>0</td>
<td>1.09009</td>
<td>-0.0017</td>
<td>0.04792</td>
<td>0.05536</td>
<td>0.07621</td>
</tr>
<tr>
<td>0.25</td>
<td>1.08352</td>
<td>0.24102</td>
<td>0.04470</td>
<td>0.05768</td>
<td>0.06737</td>
<td>0.09699</td>
</tr>
<tr>
<td>0.5</td>
<td>1.08703</td>
<td>0.49088</td>
<td>0.04577</td>
<td>0.05858</td>
<td>0.13993</td>
<td>0.01076</td>
</tr>
<tr>
<td>0.75</td>
<td>1.10633</td>
<td>0.73934</td>
<td>0.04975</td>
<td>0.04467</td>
<td>0.01838</td>
<td>0.0679</td>
</tr>
<tr>
<td>1.5</td>
<td>0</td>
<td>1.50170</td>
<td>-0.0021</td>
<td>0.05033</td>
<td>0.09211</td>
<td>0.04064</td>
</tr>
<tr>
<td>0.25</td>
<td>1.50130</td>
<td>0.25655</td>
<td>0.04557</td>
<td>0.10691</td>
<td>-0.0118</td>
<td>0.09777</td>
</tr>
<tr>
<td>0.5</td>
<td>1.50768</td>
<td>0.54749</td>
<td>0.04110</td>
<td>0.09308</td>
<td>-0.0500</td>
<td>0.06233</td>
</tr>
<tr>
<td>0.75</td>
<td>1.53243</td>
<td>0.79642</td>
<td>0.04467</td>
<td>0.08517</td>
<td>0.0574</td>
<td>-0.0475</td>
</tr>
<tr>
<td>1.9</td>
<td>0</td>
<td>1.87921</td>
<td>0.04829</td>
<td>0.03315</td>
<td>0.33591</td>
<td>-0.0274</td>
</tr>
<tr>
<td>0.25</td>
<td>1.92028</td>
<td>0.24223</td>
<td>0.01856</td>
<td>0.21779</td>
<td>-0.2136</td>
<td>0.25077</td>
</tr>
<tr>
<td>0.5</td>
<td>1.96126</td>
<td>0.54333</td>
<td>0.03037</td>
<td>0.36026</td>
<td>-0.1035</td>
<td>-0.5706</td>
</tr>
<tr>
<td>0.75</td>
<td>1.80619</td>
<td>0.74200</td>
<td>0.04439</td>
<td>0.13274</td>
<td>0.37799</td>
<td>-0.4841</td>
</tr>
</tbody>
</table>

The top part of the Table are results for \( \bar{\alpha} \) and \( \bar{\beta} \). The bottom part are for \( \bar{\sigma} \) and \( \bar{\mu} \) for fixed \( \alpha \) and \( \beta \) equal to 0 and 0.5. Sd, Skw, Kur, Min and Max stand for standard deviation, skewness, kurtosis, minimum and maximum, respectively.
<table>
<thead>
<tr>
<th>$\alpha$</th>
<th>$\beta$</th>
<th>Mean</th>
<th>SD</th>
<th>Skw</th>
<th>Kur</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.7</td>
<td>0</td>
<td>0.60986</td>
<td>0.0014</td>
<td>0.22660</td>
<td>0.04423</td>
<td>0.25316</td>
<td>0.38054</td>
</tr>
<tr>
<td></td>
<td>0.25</td>
<td>0.69645</td>
<td>0.24162</td>
<td>0.02280</td>
<td>0.04660</td>
<td>0.10523</td>
<td>-0.1229</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>0.69545</td>
<td>0.49148</td>
<td>0.25550</td>
<td>0.03888</td>
<td>0.04777</td>
<td>0.05233</td>
</tr>
<tr>
<td></td>
<td>0.75</td>
<td>0.69766</td>
<td>0.74166</td>
<td>0.24010</td>
<td>0.03176</td>
<td>0.16320</td>
<td>-0.0118</td>
</tr>
<tr>
<td>1.1</td>
<td>0</td>
<td>1.09487</td>
<td>-0.0016</td>
<td>0.44360</td>
<td>0.05521</td>
<td>0.08517</td>
<td>0.05544</td>
</tr>
<tr>
<td></td>
<td>0.25</td>
<td>1.08969</td>
<td>0.24420</td>
<td>0.04161</td>
<td>0.05355</td>
<td>0.17913</td>
<td>0.16680</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>1.09152</td>
<td>0.49833</td>
<td>0.04669</td>
<td>0.05598</td>
<td>0.17414</td>
<td>0.07989</td>
</tr>
<tr>
<td></td>
<td>0.75</td>
<td>1.10867</td>
<td>0.74535</td>
<td>0.04318</td>
<td>0.04649</td>
<td>0.04545</td>
<td>0.02925</td>
</tr>
<tr>
<td>1.5</td>
<td>0</td>
<td>1.51070</td>
<td>-0.0081</td>
<td>0.36260</td>
<td>0.08626</td>
<td>0.04553</td>
<td>0.14055</td>
</tr>
<tr>
<td></td>
<td>0.25</td>
<td>1.50381</td>
<td>0.27923</td>
<td>0.07181</td>
<td>0.12075</td>
<td>0.22064</td>
<td>0.74860</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>1.50866</td>
<td>0.50169</td>
<td>0.06184</td>
<td>0.10783</td>
<td>0.07514</td>
<td>0.30276</td>
</tr>
<tr>
<td></td>
<td>0.75</td>
<td>1.53467</td>
<td>0.74529</td>
<td>0.04551</td>
<td>0.08060</td>
<td>0.05924</td>
<td>-0.0454</td>
</tr>
<tr>
<td>1.9</td>
<td>0</td>
<td>1.88381</td>
<td>0.43680</td>
<td>0.03229</td>
<td>0.34884</td>
<td>-0.0133</td>
<td>-0.0585</td>
</tr>
<tr>
<td></td>
<td>0.25</td>
<td>1.92062</td>
<td>0.26274</td>
<td>0.01875</td>
<td>0.21549</td>
<td>-0.2060</td>
<td>0.21114</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>1.87658</td>
<td>0.50982</td>
<td>0.03111</td>
<td>0.27608</td>
<td>-0.2467</td>
<td>-0.6353</td>
</tr>
<tr>
<td></td>
<td>0.75</td>
<td>1.88085</td>
<td>0.74721</td>
<td>0.04294</td>
<td>0.12172</td>
<td>0.24978</td>
<td>-0.5013</td>
</tr>
</tbody>
</table>

See footnote Table 5.
<table>
<thead>
<tr>
<th>$\alpha$</th>
<th>$\beta$</th>
<th>Mean</th>
<th>SD</th>
<th>Skw</th>
<th>Kurt</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.7</td>
<td>0</td>
<td>0.70071</td>
<td>0.06660</td>
<td>0.22383</td>
<td>0.84186</td>
<td>0.27325</td>
<td>0.16414</td>
</tr>
<tr>
<td>0.75</td>
<td>0.25</td>
<td>0.69764</td>
<td>0.24530</td>
<td>0.23100</td>
<td>0.64434</td>
<td>0.18449</td>
<td>-0.1662</td>
</tr>
<tr>
<td>0.75</td>
<td>0.5</td>
<td>0.69705</td>
<td>0.49614</td>
<td>0.24555</td>
<td>0.03864</td>
<td>0.09840</td>
<td>0.02058</td>
</tr>
<tr>
<td>0.75</td>
<td>0.75</td>
<td>0.69905</td>
<td>0.74414</td>
<td>0.23590</td>
<td>0.03044</td>
<td>0.13177</td>
<td>-0.0994</td>
</tr>
<tr>
<td>1.1</td>
<td>0</td>
<td>0.10933</td>
<td>0.02080</td>
<td>0.04171</td>
<td>0.05348</td>
<td>0.04876</td>
<td>0.00668</td>
</tr>
<tr>
<td>1.1</td>
<td>0.25</td>
<td>0.10914</td>
<td>0.24970</td>
<td>0.03796</td>
<td>0.05225</td>
<td>0.08205</td>
<td>0.07372</td>
</tr>
<tr>
<td>1.1</td>
<td>0.5</td>
<td>0.10546</td>
<td>0.49079</td>
<td>0.03081</td>
<td>0.05115</td>
<td>0.09664</td>
<td>0.01681</td>
</tr>
<tr>
<td>1.1</td>
<td>0.75</td>
<td>0.10876</td>
<td>0.75229</td>
<td>0.03878</td>
<td>0.04233</td>
<td>0.01686</td>
<td>-0.0191</td>
</tr>
<tr>
<td>1.5</td>
<td>0</td>
<td>1.61636</td>
<td>-0.0087</td>
<td>0.15830</td>
<td>0.09958</td>
<td>-0.78811</td>
<td>-0.05277</td>
</tr>
<tr>
<td>1.5</td>
<td>0.25</td>
<td>1.60899</td>
<td>0.29539</td>
<td>0.08943</td>
<td>0.14234</td>
<td>0.07741</td>
<td>0.06266</td>
</tr>
<tr>
<td>1.5</td>
<td>0.5</td>
<td>1.60537</td>
<td>0.50892</td>
<td>0.04980</td>
<td>0.11594</td>
<td>0.06101</td>
<td>0.00962</td>
</tr>
<tr>
<td>1.5</td>
<td>0.75</td>
<td>1.53610</td>
<td>0.75847</td>
<td>0.04100</td>
<td>0.07680</td>
<td>0.23251</td>
<td>-0.1363</td>
</tr>
<tr>
<td>1.9</td>
<td>0</td>
<td>1.91296</td>
<td>0.04446</td>
<td>0.02901</td>
<td>0.38864</td>
<td>-0.0339</td>
<td>0.0382</td>
</tr>
<tr>
<td>1.9</td>
<td>0.25</td>
<td>1.90222</td>
<td>0.25615</td>
<td>0.03032</td>
<td>0.19695</td>
<td>-0.1984</td>
<td>-0.1690</td>
</tr>
<tr>
<td>1.9</td>
<td>0.5</td>
<td>1.87498</td>
<td>0.49797</td>
<td>0.03176</td>
<td>0.28403</td>
<td>-0.1910</td>
<td>-0.6245</td>
</tr>
<tr>
<td>1.9</td>
<td>0.75</td>
<td>1.90291</td>
<td>0.75551</td>
<td>0.04969</td>
<td>0.12304</td>
<td>0.20658</td>
<td>-0.5419</td>
</tr>
</tbody>
</table>

Note: See footnote Table 5.
<table>
<thead>
<tr>
<th>( \alpha )</th>
<th>( \beta )</th>
<th>Mean</th>
<th>( \text{Std} )</th>
<th>( \text{Skw} )</th>
<th>( \text{Kurt} )</th>
<th>( \text{Min} )</th>
<th>( \text{Max} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \alpha = 0.7 )</td>
<td>( \beta = 0 )</td>
<td>0.70246</td>
<td>0.00453</td>
<td>0.05523</td>
<td>0.09917</td>
<td>0.14482</td>
<td>0.02651</td>
</tr>
<tr>
<td>( \beta = 0.25 )</td>
<td>0.65042</td>
<td>0.23274</td>
<td>0.05489</td>
<td>0.09699</td>
<td>-0.0247</td>
<td>0.12688</td>
<td>2.94714</td>
</tr>
<tr>
<td>( \beta = 0.5 )</td>
<td>0.70403</td>
<td>0.39871</td>
<td>0.05703</td>
<td>0.09563</td>
<td>0.20096</td>
<td>-0.0653</td>
<td>2.97157</td>
</tr>
<tr>
<td>( \beta = 0.75 )</td>
<td>0.70149</td>
<td>0.75205</td>
<td>0.05755</td>
<td>0.08797</td>
<td>0.11734</td>
<td>0.06381</td>
<td>3.25004</td>
</tr>
<tr>
<td>( \alpha = 1.1 )</td>
<td>( \beta = 0 )</td>
<td>1.10646</td>
<td>-0.0018</td>
<td>0.06514</td>
<td>0.10478</td>
<td>0.08502</td>
<td>0.05904</td>
</tr>
<tr>
<td>( \beta = 0.25 )</td>
<td>1.16748</td>
<td>0.24723</td>
<td>0.06675</td>
<td>0.10334</td>
<td>0.45218</td>
<td>-0.0899</td>
<td>2.81269</td>
</tr>
<tr>
<td>( \beta = 0.5 )</td>
<td>1.10532</td>
<td>0.50653</td>
<td>0.06172</td>
<td>0.10160</td>
<td>0.40956</td>
<td>-0.1326</td>
<td>2.65589</td>
</tr>
<tr>
<td>( \beta = 0.75 )</td>
<td>1.10788</td>
<td>0.75612</td>
<td>0.05670</td>
<td>0.08860</td>
<td>0.46866</td>
<td>-0.1299</td>
<td>2.58816</td>
</tr>
<tr>
<td>( \alpha = 1.5 )</td>
<td>( \beta = 0 )</td>
<td>1.50394</td>
<td>0.00380</td>
<td>0.07491</td>
<td>0.15502</td>
<td>0.07944</td>
<td>-0.0121</td>
</tr>
<tr>
<td>( \beta = 0.25 )</td>
<td>1.50259</td>
<td>0.25017</td>
<td>0.07460</td>
<td>0.14791</td>
<td>0.10290</td>
<td>-0.0623</td>
<td>3.05506</td>
</tr>
<tr>
<td>( \beta = 0.5 )</td>
<td>1.50042</td>
<td>0.50430</td>
<td>0.07654</td>
<td>0.14268</td>
<td>0.15854</td>
<td>-0.1292</td>
<td>3.21382</td>
</tr>
<tr>
<td>( \beta = 0.75 )</td>
<td>1.50143</td>
<td>0.75472</td>
<td>0.07766</td>
<td>0.12883</td>
<td>0.02569</td>
<td>-0.1980</td>
<td>2.71052</td>
</tr>
<tr>
<td>( \alpha = 1.9 )</td>
<td>( \beta = 0 )</td>
<td>1.89979</td>
<td>0.00267</td>
<td>0.05785</td>
<td>0.61545</td>
<td>-0.4663</td>
<td>0.01405</td>
</tr>
<tr>
<td>( \beta = 0.25 )</td>
<td>1.89617</td>
<td>0.24182</td>
<td>0.05614</td>
<td>0.59452</td>
<td>-0.3373</td>
<td>-0.4358</td>
<td>2.56999</td>
</tr>
<tr>
<td>( \beta = 0.5 )</td>
<td>1.88859</td>
<td>0.43803</td>
<td>0.05572</td>
<td>0.52585</td>
<td>-0.3873</td>
<td>-0.8727</td>
<td>2.54725</td>
</tr>
<tr>
<td>( \beta = 0.75 )</td>
<td>1.88562</td>
<td>0.63586</td>
<td>0.05408</td>
<td>0.44541</td>
<td>-0.3426</td>
<td>-1.5231</td>
<td>2.89383</td>
</tr>
</tbody>
</table>

See footnote Table 5.
<table>
<thead>
<tr>
<th>α</th>
<th>β</th>
<th>Mean</th>
<th>Sd</th>
<th>Skw</th>
<th>Kur</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>α = 0.7</td>
<td>β = 0</td>
<td>0.00729</td>
<td>-0.0066</td>
<td>0.05014</td>
<td>0.12599</td>
<td>0.21597</td>
<td>0.00068</td>
</tr>
<tr>
<td></td>
<td>β = 0.25</td>
<td>0.07297</td>
<td>0.23859</td>
<td>0.05503</td>
<td>0.15120</td>
<td>0.04247</td>
<td>-0.3194</td>
</tr>
<tr>
<td></td>
<td>β = 0.5</td>
<td>0.04666</td>
<td>0.50730</td>
<td>0.06440</td>
<td>0.09177</td>
<td>0.19822</td>
<td>-0.3600</td>
</tr>
<tr>
<td></td>
<td>β = 0.75</td>
<td>0.01351</td>
<td>0.68729</td>
<td>0.06764</td>
<td>0.08998</td>
<td>0.26223</td>
<td>0.27189</td>
</tr>
<tr>
<td>α = 1.1</td>
<td>β = 0</td>
<td>0.09615</td>
<td>0.00356</td>
<td>0.07207</td>
<td>0.12800</td>
<td>0.04143</td>
<td>-0.3075</td>
</tr>
<tr>
<td></td>
<td>β = 0.25</td>
<td>0.09813</td>
<td>0.24569</td>
<td>0.07408</td>
<td>0.11230</td>
<td>0.12544</td>
<td>-0.2572</td>
</tr>
<tr>
<td></td>
<td>β = 0.5</td>
<td>0.10362</td>
<td>0.50553</td>
<td>0.08408</td>
<td>0.09280</td>
<td>0.10391</td>
<td>-0.1444</td>
</tr>
<tr>
<td></td>
<td>β = 0.75</td>
<td>0.11116</td>
<td>0.79433</td>
<td>0.09761</td>
<td>0.10293</td>
<td>0.36719</td>
<td>0.48664</td>
</tr>
<tr>
<td>α = 1.5</td>
<td>β = 0</td>
<td>1.49728</td>
<td>0.00052</td>
<td>0.08622</td>
<td>0.15408</td>
<td>0.26324</td>
<td>0.03276</td>
</tr>
<tr>
<td></td>
<td>β = 0.25</td>
<td>1.50519</td>
<td>0.26462</td>
<td>0.09771</td>
<td>0.16563</td>
<td>0.56144</td>
<td>0.26814</td>
</tr>
<tr>
<td></td>
<td>β = 0.5</td>
<td>1.51200</td>
<td>0.53616</td>
<td>0.10665</td>
<td>0.18244</td>
<td>0.60535</td>
<td>0.60485</td>
</tr>
<tr>
<td></td>
<td>β = 0.75</td>
<td>1.51366</td>
<td>0.78814</td>
<td>0.11552</td>
<td>0.16116</td>
<td>0.38357</td>
<td>-0.2304</td>
</tr>
<tr>
<td>α = 1.9</td>
<td>β = 0</td>
<td>1.88483</td>
<td>0.00718</td>
<td>0.10682</td>
<td>0.39920</td>
<td>-0.5890</td>
<td>0.03027</td>
</tr>
<tr>
<td></td>
<td>β = 0.25</td>
<td>1.88017</td>
<td>0.11624</td>
<td>0.10815</td>
<td>0.38410</td>
<td>-0.4547</td>
<td>0.20840</td>
</tr>
<tr>
<td></td>
<td>β = 0.5</td>
<td>1.86602</td>
<td>0.21183</td>
<td>0.11228</td>
<td>0.46063</td>
<td>-0.6868</td>
<td>0.25370</td>
</tr>
<tr>
<td></td>
<td>β = 0.75</td>
<td>1.85600</td>
<td>0.35634</td>
<td>0.11023</td>
<td>0.46777</td>
<td>-0.5977</td>
<td>0.38863</td>
</tr>
</tbody>
</table>

Table 9: Simulation Results using Empirical Quantiles

See footnote Table 5.