CORE DISCUSSION PAPER
2006/112

Estimation of Stable Distributions by Indirect Inference

René Garcia*, Eric Renault† and David Veredas‡

Decembre 2006

Abstract

This article deals with the estimation of the parameters of an α-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 α-stable distribution, with each parameter playing a similar role. To improve the properties of the estimator in finite sample, we use a variant of the method called Constrained Indirect Inference. 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. We provide an empirical application to hedge fund returns.

Keywords: Stable distribution, Indirect Inference, Constrained Indirect Inference, Skewed-t distribution.
JEL classification: C13, C15, G11.

*Author for correspondence. CIRANO and CIREQ, Université de Montréal; email: rene.garcia@umontreal.ca.
†CIRANO and CIREQ, University of North Carolina at Chapel Hill; email: renault@email.unc.edu.
‡ECARES, Université Libre de Bruxelles and CORE; email: dveredas@ulb.ac.be.

1 Introduction

The α-stable distribution has been widely used for fitting data in which extreme values are frequent. As shown in early work by Mandelbrot (1963) and Fama (1965a), it accommodates heavy-tailed financial series, and therefore produces more reliable measures of tail risk such as value at risk. The α-stable distribution is also able to capture skewness in a distribution, which is another characteristic 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 are taken into account to determine the optimal portfolio.[1] Stable processes have recently been used in the high-frequency microstructure literature as well as in consumption-based equilibrium asset pricing models.[2]

Our estimation procedure focuses on the unconditional distribution, which is often the distribution of interest in portfolio applications. Irrespective of the true data generating process, we aim at capturing the α-stable like behavior of the unconditional distribution of stationary time series often encountered in finance. In fact, the actual dynamics could well a highly persistent GARCH, since DeVries (1991) has shown that under certain conditions on the parameters of a GARCH-like process, the stable and GARCH processes are observationally equivalent from the viewpoint of the unconditional distribution.[3]

To estimate the parameters of an α-stable distribution we propose to use an indirect inference method. As well documented in Smith(1993) and Gouriéroux, Monfort and Renault (1993), indirect inference is particularly suited to situations where the model of interest is difficult to estimate but relatively easy to simulate. Therefore, it fits well the situation at hand. In most cases, the α-stable density function does not have a closed-form expression and is only characterized as an integral difficult to compute numerically. Therefore, ML estimation is not very appealing for practical use, even though Dumouchel (1973) has shown that the maximum likelihood (ML hereafter) estimator is consistent, asymptotically normal and reaches the Cramér-Rao efficiency bound. However, several methods are available to simulate α-stable random variables, such as the one described

[1] Basic references on the α-stable distribution are Feller (1971), Zolotarev (1986) and Samorodnitsky and Taqqu (1994). These properties motivate its use in the modelling of financial series in particular by Carr et al. (2002), Mittnik and Rachev (1993) and Mittnik, Paolella and Rachev (2000). For value-at-risk applications, see in particular Bassi et al. (1998) and Mittnik, Rachev and Paolella (1998). For portfolio allocation with stable distributions, see Fama (1965b), Bawa, Elton and Gruber (1979) and more recently Ortobelli, Huber and Schwartz (2002).

[2] Ait-Sahalia and Jacod (2004 a,b) proposed volatility estimators for some processes built from the sum of a stable process and another Levy process. Bidarkota, Dupoyet and McCulloch (2005) study a consumption-based asset pricing model with incomplete information and α-stable *i.i.d.* shocks.

[3] It is well-known that, except for the extreme case of the normal distribution, all the α-stable distributions have an infinite variance. However, it should be remembered that a highly persistent GARCH, with by definition finite conditional variances, may produce infinite moments at orders not much higher than two.

in Chambers, Mallows and Stuck (1976).

Indirect inference involves the use of an auxiliary model. Auxiliary parameters are recovered through the maximization of the pseudo-likelihood of a model based on the fictitious *i.i.d.* sampling in a skewed-t distribution, introduced independently in the literature by Fernandez and Steel (1998) and Hansen (1994).[4] It is a Student-t with an 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 α-stable distribution, with a clear and interpretable matching, parameter by parameter.

Our application of indirect inference is innovative in two respects. First, following an idea of McCulloch (1986) in the context of matching quantiles, we actually perform a constrained version of indirect inference, introducing an a priori constraint on one auxiliary parameter to match, namely the number of degrees of freedom of the Student-t. The theory for such constrained indirect inference has recently been developed in a general context by Calzolari, Fiorentini and Sentana (2004). Second, we stress in our application that the α-stable simulator need not take into account the actual dynamic features of the data.[5]

We show that, for a reasonable level of asymmetry, the pseudo-ML estimators of the four parameters of the skewed-t distribution are asymptotically normal even when the observations are generated by an α-stable distribution. [6] Consequently, the associated indirect inference estimators of the parameters of the α-stable distribution are asymptotically normal too. This is an important feature of the choice of the auxiliary parameters since the use of polynomials as in the conventional method of moments (see Gallant and Tauchen, 1999) could lead to asymptotically α-stable estimators.

Our extension of indirect inference can be seen as a generalization of the quantile approach, proposed by Fama and Roll (1971) and enhanced by McCulloch (1986). While McCulloch (1986) put forward four specific functions of theoretical quantiles of the α-stable distribution which are respectively well-focused on the four parameters of the α-stable distribution, we show that the four parameters of an auxiliary model provided by the skewed-t distribution have also a one-to-one correspondence with the parameters of the α-stable distribution.

A valid competing method for indirect inference consists in matching moments produced by the characteristic function (CF hereafter). This alternative is particularly relevant in the context of α-stable distributions since, by contrast with the probability

[4]During the course of this project, we were made aware by M. J. Lombardi that Lombardi, Calzolari and Gallo (2004) use the same auxiliary model to estimate a stable distribution. The two projects were conducted independently and differ in several respects.

[5]The use of a wrongly-specified simulator in indirect inference has not received much attention, except in Dridi, Guay and Renault (2005).

[6]According to our Monte Carlo experiments, the allowed level of asymmetry is actually consistent with the one produced by an α-stable distribution with support on the whole real line.

density function, the CF is available in closed form. All the methods based on the CF match the theoretical CF with its empirical counterpart, but in different ways. Since the empirical CF is a random variable with complex values, one can think about comparing (i) moments associated to real and imaginary components respectively (Press, 1972, Fielitz and Rozelle, 1981) or (ii) minimizing a distance between the empirical and the theoretical CF functions (Paulson, Holcomb and Leitch, 1975, Feuerverger and McDunnough, 1981, and Carrasco and Florens, 2002), or (iii) performing a regression analysis between the real and imaginary parts of the empirical and theoretical CF (Koutrevelis, 1980).

Except Carrasco and Florens (2002), all these methods suffer from the same drawback: the need to choose somewhat arbitrarily the frequencies of interest. Some authors, like Fielitz and Rozelle (1981), recommend to match only a few frequencies on the basis of Monte Carlo results, while others, like Feuerverger and McDunnough (1981), recommend on the contrary to use as many frequencies as possible. However, in the latter case, Carrasco and Florens (2002) have shown that, even asymptotically, matching a continuum of moment conditions introduces a fundamental singularity problem. They devise a theory for efficient continuous GMM (CGMM hereafter). It is an optimal generalized method of moments based on a continuum of moment conditions corresponding to the CF computed at all points. Along with likelihood-based methods, which as discussed are difficult to implement, their approach is the only one to provide an asymptotically efficient estimation method of the α-stable model. Therefore, it constitutes a good benchmark for assessing the performance of our indirect inference estimator.

We conduct a thorough comparison of our estimator with CGMM both in an extensive Monte Carlo study and in an empirical application. Not surprisingly, since our method delivers estimators close to the Cramér-Rao efficiency bound in the i.i.d. case [7], it does not appear to be clearly dominated by CGMM in terms of efficiency. More importantly, CGMM encounters some difficulties in its implementation (in particular the choice a of a regularization parameter) and its performance is much less robust than our simple indirect inference estimator. We also compare our method to the simple but inefficient quantile-based estimator of McCulloch (1986). Our estimates are close to those obtained with the quantile-based method. However, our estimators appear to have a much smaller variance, both asymptotically and in finite sample. [8]

Many of the properties of stable models are shared by GARCH models. In particular, both models share the facts that the unconditional distribution has fat tails and that the

[7]Since Dumouchel (1973) provides a way to compute the efficiency bound in the *i.i.d.* case, we are able to compare by Monte Carlo study the performance of our indirect inference estimator with the ML benchmark. At least in this *i.i.d.* setting, the efficiency loss appears mainly negligible given the finite sample improvement brought about by indirect inference.

[8]Another set of methods is based on the asymptotic tail behaviour of the distribution. The most known, the Hill estimator (Hill, 1975), is an ML estimator of a tail index via the Pareto distribution since the asymptotic tails of an α-stable distribution behave like a Pareto distribution. This approach has two drawbacks: the choice of the quantile from where the tail is considered is arbitrary and, more importantly, only the stability index α is estimated. Therefore, we do not include it in our comparison.

tail shape is invariant under addition (see Ghose and Kroner 1995 and de Vries 1991). We illustrate this observational equivalence by generating different GARCH(1,1) and IGARCH(1,1) with Gaussian and Student-t innovations and aggregating the generated processes to lower frequencies. We show that the unconditional density captures very well the variance and kurtosis through aggregation and memory. The tail index α remains relatively constant under aggregation while the estimated dispersion increases. As expected, the tail index and the dispersion are higher when the process is generated from a Student density that when it comes from a Gaussian probability distribution.

Our empirical application of the three selected estimators involves time series of monthly returns from indices of hedge funds. Several studies, in particular Fung and Hsieh (2002), Mitchell and Pulvino (2001), and Agarwal and Naik (2004), have put forward the nonlinear structure and option-like features of returns associated with hedge fund strategies. It means that these return distributions are likely to exhibit skewness and kurtosis and are therefore potentially well captured by stable distributions. We want to reemphasize that the stable distributions are preserved under convolution and that this property is appealing when considering portfolios of assets. Pension funds are now all considering the inclusion of hedge funds along with the traditional classes of assets in their portfolios and one needs a statistical framework that accomodates the non-normal characteristics of these funds. Results show that our estimator behaves well even in small samples (series on hedge funds are typically small), which is not always the case with the quantile or the continuous GMM methods.

Although we do not consider testing in this paper, indirect inference provides as a by-product specification tests about the matched characteristics, in our case the unconditional distribution. One can envision a battery of diagnostic tools. For example, the fact that the binding function can be interpreted parameter by parameter allows independent assessments of the ability of the stable model to capture the four relevant features of the data. One can also perform an omnibus test, by matching jointly McCulloch quantile-based functions and our skewed-t auxiliary parameters to obtain an automatic overidentification test. These tests could complement the work by Deo (2000) who proposed, in the context of m-dependent sequences, a goodness-of-fit test for stable distributions.

The structure of the paper is as follows. Section 2 describes the properties of α-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 α-stable distribution and the skewed-t distribution chosen as an 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. In Section 6, we discuss the use of the α-stable model and of our estimation strategy in the context of dependent processes. We compare and illustrate through simulations the relationship between the fat-tailed unconditional distributions produced by highly persistent GARCH models and an α-stable model. Section 7 is devoted to an empirical application to hedge funds while section 8 concludes. Theoretical proofs as well as numerical and computational issues are gathered in appendices.

2 The α-stable distributions and their efficient estimation methods

In general, stable distributions do not have closed form expressions for their density and distribution functions but can be described easily by their characteristic functions. The family of univariate α-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 recall in this section the interpretation of the four parameters as well as the strategies available in the literature for their efficient estimation. While efficient estimation will remain a benchmark when considering alternative moment-matching methods of estimation in Section 3, the interpretation of the parameters will be crucial to define well-suited moments to match.

2.1 Parameters and properties of α-stable distributions

The α-stable family of distributions is characterized by four parameters α, β, σ and μ, where α is the stability parameter, β the skewness parameter, σ the scale parameter, and μ the location parameter. These parameters define the natural logarithm of the characteristic function as

$$\ln \psi_\theta(t) = \ln E\left[\exp(it\,Y)\right] = i\mu t - \sigma^\alpha \left|t\right|^\alpha \left[1 - i\beta\ sign(t) w(t,\alpha)\right] \tag{2.1}$$

where

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

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

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

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

If β 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 $0 < \alpha < 2$:

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

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

Therefore, the stability parameter α 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, α-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 α-stable distribution for θ, an interior point of Θ:

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

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, Mittnik, 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 α-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 $p \geq \alpha$, $\frac{1}{n^{1/2}} \sum_{i=1}^{n} Y_i$ 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. This fact has been used by DuMouchel (1973) to develop the asymptotic theory of maximum likelihood in the context of a family of α-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 $\theta = (\alpha, \beta, \mu, \sigma)$ insofar as the domain of possible values of θ is limited in the following way:

$$\begin{aligned} \alpha &\in \ ]1,2[\text{ or } \alpha \in [\varepsilon, 1[\text{ for some } \varepsilon > 0, \\ & \quad \text{and} \\ |\beta| &< \min(\alpha, 2-\alpha). \end{aligned}$$

In particular, totally skewed-stable distribution ($|\beta| = 1$) are discarded as well as arbitrarily fat tails (α arbitrarily close to zero).

Because of the numerical difficulties associated with maximum likelihood estimation, we will not use it in this paper.[9] 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, the results 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 $Y \rightsquigarrow S(\alpha, \beta, \mu, \sigma)$. 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:

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

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

[9] We will not either compare our method to the ML estimation method where the density function of the stable distribution is approximated by using fast Fourier transforms of the characteristic function. Since it is based on a finite grid, one can argue that it suffers from a loss of efficiency with respect to the continuous GMM, to which we compare our results.

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

$$E\left[\exp(it_kY) - \psi_\theta(t_k)\right] = 0, \quad k = 1, \cdots K.$$

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

$$E\left[g_K(\theta, Y)\right] = 0$$

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

$$\mathrm{Re}h(t_k, Y, \theta) = \cos(t_kY) - \mathrm{Re}\psi_\theta(t_k)$$

and the imaginary parts

$$\mathrm{Im}\ h(t_k, Y, \theta) = \sin(t_kY) - \mathrm{Im}\psi_\theta(t_k)$$

of

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

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

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

A simple way to back out the coefficients of Σ_K is then:

$$\begin{aligned}
cov\left[\cos\left(tY\right), \cos\left(sY\right)\right] &= \frac{1}{2}\left[\mathrm{Re}\Psi_\theta(t+s)+\mathrm{Re}\Psi_\theta(t-s)\right] - \left(\mathrm{Re}\Psi_\theta(t)\right)\left(\mathrm{Re}\Psi_\theta(s)\right) \\
cov\left[\cos\left(tY\right), \sin\left(sY\right)\right] &= \frac{1}{2}\left[\mathrm{Im}\Psi_\theta(t+s)+\mathrm{Im}\Psi_\theta(t-s)\right] - \left(\mathrm{Re}\Psi_\theta(t)\right)\left(\mathrm{Im}\Psi_\theta(s)\right) \quad (2.2) \\
cov\left[\sin\left(tY\right), \sin\left(sY\right)\right] &= \frac{1}{2}\left[\mathrm{Re}\Psi_\theta(t+s)+\mathrm{Re}\Psi_\theta(t-s)\right] - \left(\mathrm{Im}\Psi_\theta(t)\right)\left(\mathrm{Im}\Psi_\theta(s)\right)
\end{aligned}$$

By using the empirical characteristic function $\hat{\phi}(t) = \frac{1}{n} \sum_{j=1}^{n} \exp\left(itY_j\right)$ as a consistent estimator of $\Psi_\theta\left(t\right)$, consistent estimators of the coefficients of Σ_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 Ω_K tends to the Cramer-Rao bound for estimation of θ. However, as first noticed by Carrasco and Florens (2002), this does not provide a way to estimate efficiently θ based on the characteristic function since the $2K$ moment conditions $E\left[g_K(\theta, Y)\right] = 0$ will suffer from a multicollinearity problem when K goes to infinity. In this latter case, one must think instead about the limit of $\sum_K^{-1}$ in terms of covariance operator. Since, for any t:

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

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

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

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

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

where $w(t,s) = E\left[h(t,Y,\theta)h(s,Y,\theta)\right]$. Then under standard regularity conditions, $\sqrt{n}\bar{h}_n(\theta) = \dfrac{1}{\sqrt{n}} \sum\limits_{j=1}^{n} h(Y_j, \theta)$ converges in $L^2(\mathbb{R},\Pi)$ towards a Gaussian process $\mathcal{N}\left[0, \Omega\right]$. This paves the way for defining optimal GMM for the continuum of moment conditions of interest. More precisely, the operator Ω is consistently estimated by the operator Ω_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 θ. 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 $\left|\left|\Omega^{-1/2}\bar{h}_n(\theta)\right|\right|$ with Ω replaced by its consistent estimator Ω_n. However, since Ω is a compact operator, its eigenvalues converge to zero and the operator Ω^{-1} is not continuous. Then a regularization scheme is needed. For instance, a Tykhonov regularized inverse of Ω_n is defined by:

$$\left(\Omega_n^{(\beta_n)}\right)^{-1} = \left(\beta_n Id + \Omega_n^2\right)^{-1} \Omega_n$$

where Id is the identity operator and $\beta_n > 0$ is a regularization parameter. Then, Carrasco and Florens (2000) propose to estimate θ as:

$$\hat{\theta}_n = \arg\min_{\theta \in \Theta} \left|\left| \left(\Omega_n^{(\beta_n)}\right)^{-1/2} \bar{h}_n(\theta) \right|\right|.$$

They show (see also Carrasco, Chernov, Florens and Ghysels (2004)) that for a sequence $\beta_n, n \in N$, of regularization parameters such that $\beta_n \longrightarrow \infty$ but $n\beta_n^{5/2} \longrightarrow \infty$ when $n \longrightarrow \infty, \hat{\theta}_n$ 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 procedure used to compute the sequence of operators $(\Omega_n^{(\beta_n)})^{-1/2}$. Carrasco, Florens and Renault (2005) provide a survey of estimation methods based on spectral decomposition and regularization. However, for this particular example, Carrasco, Chernov, Florens and Ghysels (2004) give a way to compute the objective function $\left\|\left(\Omega_n^{(\beta_n)}\right)^{-1/2} \bar{h}_n(\theta)\right\|^2$ without resorting to any spectral decomposition. They show that minimizing $\left\|\left(\Omega_n^{(\beta_n)}\right)^{-1/2} \bar{h}_n(\theta)\right\|^2$ is equivalent to minimizing:

$$\underline{\mathrm{v}}(\theta)' \left[\alpha_n Id_n + C^2\right]^{-1} \underline{\mathrm{v}}\,(\theta)$$

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

$$v_i(\theta) = \int \overline{h(t, Y_i, \tilde{\theta}_n)} \bar{h}_n(t, \theta)\Pi(dt)$$

and

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

Note however that the theoretical asymptotic result does not indicate how to select the regularization parameter β_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 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 moment-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 that will be strongly informative about the parameters of the stable distribution, 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).

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 α, β, μ and σ suggests some regression-based summaries of a set of values $\Psi_\theta(t_k)$, $k = 1 \dots K$, 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 $\exp\left(-\frac{x^2}{2}\right)$ 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 $S(\alpha, \beta, \mu, \sigma)$, i.e. $P[Y < x_p] = p$. 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 μ and σ. The idea of McCulloch (1986) 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 α and β. Moreover, to get accurate estimators, it is intuitive to define the first auxiliary parameter $\phi_1 = \phi_1(\alpha,\beta)$ so it is well focused on α (for each β), and similarly for β (for each α) with the second auxiliary parameter $\phi_2 = \phi_2(\alpha, \beta)$. The definition of the two auxiliary parameters (ϕ_1, ϕ_2) must then be tightly related to the interpretation of the two structural parameters (α,β) that they intend to capture. Therefore, McCulloch (1986) proposes to define ϕ_1 as a measure of the relative sizes of the tails and the middle of the distribution:

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

A larger ϕ_1 means fatter tails and then a smaller α. McCulloch (1986) remarks that, since $\phi_1(\alpha, \beta)$ is a strictly decreasing function of α, for each β, the estimation of ϕ_1 will give us a strong fix on α. The function ϕ_2 is defined as a measure of the spread between the right part and the left part of the distribution:

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

A larger ϕ_2 means more weight on the right side and thus a larger β. In other words, since $\phi_2(\alpha, \beta)$ is a strictly increasing function of β, for each α, the estimation of ϕ_2 will be very informative about β.

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

$$\begin{cases} \phi_1\left(\hat{\alpha}, \hat{\beta}\right) = \hat{\phi}_1 \\ \phi_2\left(\hat{\alpha}, \hat{\beta}\right) = \hat{\phi}_2 \end{cases} \tag{3.3}$$

where

$$\begin{aligned} \hat{\phi}_1 &= \frac{\hat{x}_{0.95} - \hat{x}_{0.05}}{\hat{x}_{0.75} - \hat{x}_{0.25}} \\ \hat{\phi}_2 &= \frac{\hat{x}_{0.95} + \hat{x}_{0.05} - 2\hat{x}_{0.5}}{\hat{x}_{0.95} - \hat{x}_{0.05}}. \end{aligned} \tag{3.4}$$

McCulloch (1986) provides tables of values for the functions $\phi_1(\alpha,\beta)$ and $\phi_2(\alpha,\beta)$, given a grid of values of (α,β), 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 in (3.3), 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 α increases to 2, irrespective of the value of β. 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):

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

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 μ and proportional to σ, it is natural to define an estimator of the scale parameter σ as:

$$\hat{\sigma} = \frac{\hat{x}_{0.75}-\hat{x}_{0.25}}{\hat{\phi}_3} \tag{3.5}$$

(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 μ nor σ and can be tabulated for a grid of values of (α,β). The needed estimation $\hat{\phi}_3$ of the auxiliary parameter ϕ_3 is deduced from the previous estimation (3.3) of (α,β)

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

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

Finally, to back out the location parameter μ, 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 (α, β) independent of μ and σ. Unfortunately, although well defined when $\alpha = 1$, this function goes to $(-\infty)$ (resp. $+\infty$), for all β, when α 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

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

knowing that

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

From the estimation $\hat{\phi}_4 = \phi_4\left(\hat{\alpha}, \hat{\beta}\right)$, we then deduce the estimator of μ from previously defined estimators of (α, β, σ):

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

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 $\hat{\Psi} = \left(\hat{\Psi}_1, \hat{\Psi}_2, \hat{\Psi}_3, \hat{\Psi}_4\right)'$ of a vector $\Psi =$plim$\hat{\Psi}$ of auxiliary parameters. This estimator is defined by:

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

where f and g are known functions defined by:

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

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 Ψ defined as:

$$\Psi(\theta) = \begin{bmatrix} \phi_1(\theta_1, \theta_2) \\ \phi_2(\theta_1, \theta_2) \\ \theta_3 \\ \theta_4 \end{bmatrix}. \tag{3.10}$$

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:

$$\gamma = \Gamma(\theta). \tag{3.11}$$

In some respect, resorting to the more parsimonious vector Ψ of auxiliary parameters (instead of γ) 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 Ψ_i of Ψ is conceived to be directly informative about the corresponding component θ_i of θ.

A second important element is the additional constraint on ϕ_1 (or Ψ_1) introduced by McCulloch (1986). Although immaterial asymptotically when the unknown true value α^0 of α 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 $\left(\hat{\Psi}_2, \hat{\Psi}_3, \hat{\Psi}_4\right)$ which does not allow to identify the four unknown structural parameters. A relevant solution for this problem is to use the constrained indirect Inference 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:

$$\ln\left(-\ln|\phi(t)|^2\right) = \ln(2\sigma^\alpha) + \alpha\ln|t| \tag{3.12}$$

and (for $\alpha \neq 1$):

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

for the real and imaginary parts of $\phi(t)$. These two equations lead to

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

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

$$y = m + \alpha w_\kappa + \varepsilon_\kappa, \quad k = 1, 2, \cdots, K$$

where $(t_k; k = 1, 2, \cdots, \kappa)$ is an appropriate set of real numbers and $m = \ln(2\sigma^\alpha)$. This regression model provides estimates of σ and α. Given these estimates, one can use the regression model[10]

$$z_l = \mu u_l - \beta\sigma^\alpha\tan\left(\frac{\Pi\alpha}{2}\right) sgn(u_l)\,|u_l|^\alpha + \eta_l \qquad l = 1, 2, \cdots, L$$

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

[10]The function z is equal to $\arctan(\mathrm{Im}\phi_n(u)/\mathrm{Re}\phi_n(u)) + \pi k_n(u)$ 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.

Therefore, this two-step procedure provides estimates of the four parameters of the stable law[11] 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 and the regression method provide accurate estimates of the characteristic exponent α. However, they note that in general the estimates of the skewness parameter β are not as good as the estimates of the stability index α. The mean squared errors as well as the biases for both methods are relatively large. This is especially true when α is close to 2, as already explained in the previous section. The regression method does not therefore improve significantly over the McCulloch (1986) method.[12]

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) and Bauwens and Laurent (2005) for the multivariate extension).

Let us consider the skewed-t density function:

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

where

$$h(\nu) = \frac{2\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} \text{ and } g_\omega(y, \gamma) = \begin{cases} \frac{1}{\gamma^2} \text{ if } y \geq \omega \\ \gamma^2 \text{ if } y < \omega. \end{cases}$$

While the possibly non integer degrees of freedom ν of a Student distribution capture the thickness of the tails as α does for stable distributions, location ω and scale parameter λ can easily be introduced to match the two parameters μ and σ. Finally, the skewed-t extension allows one to accommodate skewness through an additional parameter γ which should be informative about β.

[11] 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_ℓ.

[12] However, they also provide bootstrapping results based on samples drawn from stock-market data and recommend the regression method based on these results.

Assuming that $Y \rightsquigarrow S(\alpha, \beta, \mu, \sigma)$ we define a vector $\phi(\alpha, \beta, \mu, \sigma)$ of four auxiliary parameters:

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

characterized by:

$$\phi(\alpha, \beta, \mu, \sigma) = \underset{(\nu,\gamma,\omega,\lambda)}{\arg\max} E\left[\log\ l(Y; \nu, \gamma, \omega, \lambda)\right].$$

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 marginal distribution is the stable one $S(\alpha, \beta, \mu, \sigma)$, irrespective of the possibly dynamic structure of the data generating process. 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: $\nu \longleftrightarrow \alpha, \gamma \longleftrightarrow \beta, \omega \longleftrightarrow \mu,\ \lambda \longleftrightarrow \sigma$. To see this, we prove four results.

Proposition 3.1: *For any real number* a, $\phi_3(\alpha, \beta, \mu + a, \sigma) = \phi_3(\alpha, \beta, \mu, \sigma) + a$, *or:* $\omega(\mu + a) = \omega(\mu) + a$

Proposition 3.1 confirms that the auxiliary parameter $\omega = \phi_3$ should inform us very well on the location parameter μ.

Proposition 3.2: *For any* $a > 0, \phi_4(\alpha, \beta, a\mu, a\sigma) = a\phi_4(\alpha, \beta, \mu, \sigma)$, *or:* $\lambda(a\sigma) = a\lambda(\sigma)$.

Proposition 3.2 indicates that the auxiliary parameter $\lambda = \phi_4$ should be very informative about the scale parameter σ.

Proposition 3.3:

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

or

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

Proposition 3.3 shows that the auxiliary parameter $\gamma = \phi_3$ should capture the skewness parameter β.

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 $Y \rightsquigarrow S(\alpha, \beta, \mu, \sigma)$, $h(\nu)$ *is the* log *quasi likelihood function and* $(\nu, \gamma, \omega, \lambda) = \phi(\alpha, \beta, \mu, \sigma)$

In particular

$$\begin{cases} \phi_1(\alpha, \beta, \mu + a, \sigma) = \phi_1(\alpha, \beta, \mu, \sigma) \text{ for all } a, \\ \phi_1(\alpha, \beta, a\mu, a\sigma) = \phi_1(\alpha, \beta, \mu, \sigma) \text{ for all } a > 0, \\ \phi_1(\alpha, -\beta, \mu, \sigma) = \phi_1(\alpha, \beta, \mu, \sigma) \end{cases}$$

or

$$\begin{cases} \nu(\mu + a) = \nu(\mu) \\ \nu(a\sigma) = \nu(\sigma) \\ \nu(-\beta) = \nu(\beta) \end{cases}$$

Proposition 3.4 confirms that the auxiliary parameter $\nu = \phi_1$ should correspond to the tail parameter α. 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 $\left(\hat{\nu}, \hat{\gamma}, \hat{\omega}, \hat{\lambda}\right)$ of the skewed-t parameters should be very informative about the structural parameters $(\alpha, \beta, \mu, \sigma)$. Indirect inference estimators $\left(\hat{\alpha}, \hat{\beta}, \hat{\mu}, \hat{\sigma}\right)$ of the latter could be simply computed as solutions of the following equations

$$\begin{cases} \hat{\nu} = \phi_1\left(\hat{\alpha}, \hat{\beta}, \hat{\mu}, \hat{\sigma}\right) \\ \hat{\gamma} = \phi_2\left(\hat{\alpha}, \hat{\beta}, \hat{\mu}, \hat{\sigma}\right) \\ \hat{\omega} = \phi_3\left(\hat{\alpha}, \hat{\beta}, \hat{\mu}, \hat{\sigma}\right) \\ \hat{\lambda} = \phi_4\left(\hat{\alpha}, \hat{\beta}, \hat{\mu}, \hat{\sigma}\right) \end{cases}$$

Of course, the binding functions ϕ_1, ϕ_2, ϕ_3, and ϕ_4 are not known in closed form and must be recovered though simulations. Note that the binding functions involve only marginal distributions. Therefore, simulations can be performed with *i.i.d.* sampling in the stable distribution, without accounting for the possibly dynamic features of the data. However, in finite sample, and most likely for larger values of the true unknown α, $\hat{\nu}$ may be off the theoretical range corresponding to stable distributions exactly as ϕ_1 was in the previously described quantile-based approach. To see this, note that since $\lim\limits_{\nu=+\infty} h'(\nu) = 0, \nu \longrightarrow +\infty$ is always a solution of the equation in proposition 3.4 which defines the pseudo-true value of ν. Intuitively, for α 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 ϕ_1 by imposing on it an upper bound as McCulloch (1986) did for his auxiliary parameter which provided information about α. We choose to impose $\nu \leq 2$ that is to redefine ϕ_1 as $\phi_1 = \min(\nu, 2)$. Theoretically, the constraint on ν does not have to be strictly 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

$$L_n^*(\Psi) = L_n(\phi) + \delta(2-\nu) \tag{4.1}$$

where $\Psi = (\phi', \delta')'$ and $\phi = (\nu, \gamma, \omega, \lambda)$ is the vector of auxiliary parameters corresponding to the skewed-t quasi-likelihood function

$$\begin{aligned} L_n(\phi) &= n \log\left[\frac{h(\nu)}{\sqrt{\pi\nu}} \cdot \frac{1}{\lambda\left(\gamma + \frac{1}{\gamma}\right)}\right] \\ &\quad - \frac{\nu+1}{2} \sum_{i=1}^{n} \log\left[1 + \frac{1}{\nu}\left(\frac{Y_i - \omega}{\lambda}\right)^2 g_\omega(Y_i, \lambda)\right] \end{aligned} \tag{4.2}$$

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 δ identical to zero (no constraint on ν) and second with $\delta \geq 0$ (inequality constraint on ν).

The estimator $\hat{\Psi}$ of the pseudo-true value of Ψ is then defined by the first-order conditions

$$\left\{ \begin{array}{l} \dfrac{\partial L_n}{\partial(\gamma, \omega, \lambda)'}\left(\hat{\phi}\right) = 0 \\ \dfrac{\partial L_n}{\partial \nu}\left(\hat{\phi}\right) = \hat{\delta} \end{array} \right. \tag{4.3}$$

jointly with the complementary slackness restriction $\hat{\delta}(2 - \hat{\nu}) = 0$

Let us denote $Y_i^h(\theta), h = 1, \cdots, H, \ i = 1, \cdots n$, the components of H simulated paths of an α-stable process for a given value $\theta = (\alpha, \beta, \mu, \sigma)$ of structural parameters.

As already mentioned, the simulations are performed with *i.i.d.* draws in the stable distribution.

The simulated path $\left(Y_i^h(\theta)\right)_{1\leq i\leq n}$ defines a simulated criterion function

$$L_n^{*h}(\Psi\,|\theta) = L_n^h(\phi\,|\theta) + \delta(2-\nu) \tag{4.4}$$

where $L_n^h(\phi\,|\theta)$ is computed as in (4.2) but with simulated data $\left(Y_i^h(\theta)\right)_{1\leq i\leq n}$ 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

$$\begin{cases} \dfrac{\partial L_n^h}{\partial(\gamma,\omega,\lambda)'}\left[\hat{\phi}^h(\theta)\,|\theta\right] = 0 \\ \dfrac{\partial L_n^h}{\partial\nu}\left[\hat{\phi}^h(\theta)\,|\theta\right] = \hat{\delta}^h(\theta) \\ \hat{\delta}^h(\theta)\cdot\left(2-\hat{\nu}^h(\theta)\right) = 0 \end{cases} \tag{4.5}$$

Let us then consider the average estimator over the H simulated paths

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

The main idea of indirect inference is to choose the estimator $\hat{\theta}$ of structural parameters θ in order to match $\tilde{\Psi}_H(\theta)$ 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:

$$\hat{\phi}^u = \hat{\phi}_H^u\left(\hat{\theta}^u\right) \tag{4.6}$$

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 $h = 1,\cdots,H$. 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 overidentified problem since both $\tilde{\Psi}_H(\theta)$ and $\hat{\Psi}$ entail five free parameters while the unknown θ 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:

$$\hat{\theta}^c = \arg\min_{\theta}\left(\tilde{\Psi}_H(\theta)-\hat{\Psi}\right)' W\left(\tilde{\Psi}_H(\theta)-\hat{\Psi}\right) \tag{4.7}$$

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 θ will go through the Kuhn-Tucker multiplier $\hat{\delta} = \frac{\partial \ln}{\partial \nu}\left(\hat{\phi}\right)$. Therefore, constrained indirect inference will not suffer from the weak identification problem about β that is currently encountered with competing estimation methods when the true unknown value of α 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 $\left(1 + \frac{1}{H}\right)$.

Generally speaking, standard theory of indirect inference (see Gourieroux, Monfort and Renault, 1993) can be applied insofar as the information content of auxiliary parameters is sufficient to identify the structural parameters and as the estimator $\hat{\Psi}$ of auxiliary parameters is root-n asymptotically normal. We show in the appendix that the latter property is fulfilled by $\hat{\Psi}$ solution of (4.3), at least when the pseudo-true value of the asymmetry coefficient γ^2 does not exceed $(2+\sqrt{5})^{1/2}$. According to our Monte Carlo study, this constraint does not appear to be binding for application to the α-stable model. This maintained asymptotic normality 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, the standard theory of overidentification tests in the indirect inference setting will provide $\mathcal{X}^2$-based goodness of fit tests for stable observations. As mentioned by Deo (2000), there are not many statistical tests that are both formal and simple to implement for the goodness-of-fit of stable distributions. While Deo (2000) proposes some $\mathcal{X}^2$-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 $\widehat{\theta}$ 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 α-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 α, namely 0.3, 0.7, 1, 1.3, 1.5, 1.7 ,1.9, 1.95 and 1.99.[13] We keep the other parameters μ, σ and β fixed and set them equal to 0, 0.5 and 0 respectively. The simulation experiment is divided in two parts.[14] First we estimate the unconstrained skewed-t distribution, i.e. solving

$$\widehat{\phi}^u = \arg\max_{\phi} L_n(\phi), \tag{5.1}$$

where $\ln(\phi)$ is as defined in (4.2). Since the true values of ϕ are unknown, we can only check by Monte Carlo the behavior of the first four moments of the estimators. The results are reported in Table 1. Concerning ω, λ and γ, the Monte-Carlo skewness and kurtosis coefficients of their estimators are fairly close to zero and three respectively, which confirms the finite sample validity of the asymptotic normal approximation of the distribution of the estimators. Moreover, in accordance with the theoretical analysis of section 3.3, we check that these estimators do not vary too much in terms of mean and standard deviation when the value of α is changed.

In contrast, $\widehat{\nu}$ is ill-behaved in finite sample, especially when α approaches to 2. Of course, one would expect $\widehat{\nu}$ to be an increasing function of α, since these two parameters are supposed to describe the tails. However, when α approaches 2, $\widehat{\nu}$ appears to be attracted towards infinity, as if, in finite sample, one had the spurious feeling to reach normality corresponding to the limit case $\alpha = 2$ and $\nu = +\infty$. On the contrary, even for $\alpha = 1.99$, population variance is infinite, which should correspond to ν smaller than two. Intuitively, $\widehat{\nu}$ is attracted too much towards very large values in finite sample. This intuition is confirmed by examination of the variance, skewness and kurtosis of $\widehat{\nu}$, which feature some weird values when α approaches 2. Figure 1 represents the kernel density of the Monte Carlo distribution of $\widehat{\nu}$ for several values of α. For a sample of 1000 observations, when $\alpha \geq 1.5$ the estimator $\widehat{\nu}$ exhibits serious departures from normality. Notice that for α 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. For $\alpha = 1.99$, the density of $\widehat{\nu}$ becomes bimodal, which shows even more that besides the true unknown pseudo-true value, $\widehat{\nu}$ is attracted towards $+\infty$.

While the kernel densities represented in Figure 1 are all about a sample size of 1000, the role of sample size in the dependence of $\widehat{\nu}$ in α is investigated in Table 2. Clearly, the closer is α to the value 2, the more observations are needed to get approximate normality for the distribution of $\widehat{\nu}$. No less than 10000 observations are needed for fairly approaching normality when $\alpha = 1.9$. Even more observations would be necessary for larger values of α.

We also show in Figure 2 the link between the degrees of freedom of the auxiliary model and the stability index α. We plot the estimated $\widehat{\nu}$ as a function of α for a given set of values for β (indicated in the legend of the Figure). One can conclude from the

[13] In Appendix B, we explain how to simulate from an α-stable distribution

[14] We also carried out the same experiment for different values of β and obtained similar results.

figure that the relation between $\widehat{\nu}$ and α appears to be exponential, confirming that as $\alpha \to 2$ we get closer to a Gaussian distribution within the α-stable family and therefore $\widehat{\nu} \to \infty$. Moreover, since $\widehat{\nu}$ is not always below two, it means that the true process may be an infinite variance process, yet the estimated pseudo-true skewed-t distribution has a finite second moment. Finally, β does not have a significant impact on $\widehat{\nu}$ 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 θ we generate 500 samples of 1000 observations for different values of θ and estimate for each sample $\widehat{\theta}$ by indirect inference. The values of μ and σ are set to 0 and 0.5, while α takes the values 1.5 and 1.9 and β the values 0 and 0.75. The results, shown in Table 3, indicate that for $\alpha = 1.9$ the skewness and kurtosis for μ and σ 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 α comes close to 2.

Therefore we propose to use a constrained version of the skewed-t distribution. Since the estimate $\widehat{\nu}$ is ill-behaved in finite sample when $\alpha \to 2$ because it is attracted by ∞, we impose an upper bound on ν, i.e. $\nu < \nu^c$. We choose $\nu^c = 2$ but in fact even a larger bound could fulfill the purpose.[15] However, choosing a larger bound, say of 10, is not innocuous in terms of finite sample performance of constrained indirect estimators of α and β. Generally speaking, a tighter bound will provide a more accurate estimator of α, while deteriorating the estimation of β. We choose to maintain the bound at 2, which leads to the maximization program:

$$\hat{\Psi}^c = \arg\max_{\Psi} L_n(\phi) + \delta(2 - \nu), \tag{5.2}$$

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

If $\widehat{\nu}^c$, the estimate of ν under (5.2), is smaller than two the multiplier $\hat{\delta}$ is zero, because of the slackness condition. On the contrary, if $\widehat{\nu}^c$ reaches its upper bound, the multiplier is not zero. In the former case, 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 $\widehat{\nu}^c$ who does not play a role and the multiplier is the parameter at work with constrained indirect inference. To check that the estimated multiplier is well-behaved in finite sample, we draw in Figure 3 the densities of $\hat{\delta}$ when α is getting closer to 2. It can be seen that, contrary to $\widehat{\nu}$ (see Figure 1), the distributions are much closer to a normal. This explains why $\hat{\delta}$ is a more relevant auxiliary parameter than $\widehat{\nu}$ in such a case. The role of sample size is studied 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 α and

[15] The constrain $\nu < 2$ is sufficient but not necessary, since what matters is to prevent $\widehat{\nu}$ from approaching ∞.

β, while setting the location and scale parameters, μ and σ, to 0 and 0.5 respectively.[16] We choose four different positive values, namely $\{0, 0.25, 0.5, 0.75\}$, for the skewness parameter β. Finally the stability index α takes the values $\{0.7, 1.1, 1.7, 1.9\}$. We generate 500 samples of 1000 observations for the resulting values of θ. 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 variance of the moments to match and possibly the finite sample bias and variance of the resulting estimator. We choose H= $\{1, 2, 5\}$ and report the results in Tables 5, 6 and 7 respectively.[17]

To conduct the constrained indirect inference estimation procedure, we first estimate the set of auxiliary parameters from the generated series for a given set of stable parameters. If the number of degrees of freedom $\hat{\nu}$ is estimated at a value above 2, we switch to estimating the Lagrange multiplier $\hat{\delta}$. Thereafter, we conduct the indirect inference procedure to match this $\hat{\delta}$, along with the three other auxiliary parameters. Therefore, we either match $\left(\hat{\nu}, \hat{\gamma}, \hat{\lambda}, \hat{w}\right)$ if $\hat{\nu}$ is less than 2, or $\left(\hat{\delta}, \hat{\gamma}, \hat{\lambda}, \hat{w}\right)$ is $\hat{\nu}$ is greater or equal to 2. For the weighting matrix W, we choose a (4 x 4) identity matrix in each case, whether we match the set of auxiliary parameters with $\hat{\gamma}$ or $\hat{\delta}$ respectively. With the 500 estimated parameters we compute some basic statistics: mean, standard deviation, skewness, kurtosis, minimum and maximum, so we can assess 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 α-stable distribution. The first one is the continuous GMM method of Carrasco and Florens (2002) based on the characteristic function with a regularization parameter β_n equal to 10^{-6}.[18] The second one is the empirical quantile method of McCulloch (1986).[19] 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 θ. The skewness for all parameters is close to 0 and the kurtosis close to 3. Thanks to the constraint imposed on ν in the auxiliary model, the estimator behaves well even when α approaches 2. However, when α is equal to 1.1 and is therefore close to 1, the parameter μ is badly estimated since theoretically it becomes infinite. This is a feature which is shared by all estimation methods of the α-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,

[16] 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 $\sigma = 0.5$.

[17] In Appendix C, we comment on some numerical aspects related to the implementation of the Monte Carlo procedure.

[18] 10^{-6} is the same value that Carrasco and Florens (2002) chose for their Monte Carlo study with the α-stable distribution.

[19] We use a GAUSS procedure written by J. Huston McCulloch and available in his web page http://www.econ.ohio-state.edu/jhm/jhm.html

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 σ. For values of α less than 1, continuous GMM overestimates the value of σ, while it underestimates it for values of α 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 $\alpha = 1.1$. Estimates for other parameters , for example β, suffer when α 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.[20]

The empirical quantile method does not seem to suffer from any systematic bias except for β 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 α-stable distribution. It certainly provides an improvement in terms of efficiency over classical quantile methods and even over the more recently proposed continuous GMM method.

To judge the efficiency of the indirect inference procedure, we also compare in Table 10 the empirical standard deviations of α and β to the asymptotic Cramer-Rao bounds reported in DuMouchel (1975) for a set of parameter values. It can be seen that the constrained indirect inference procedure produces standard deviations that are close to the asymptotic lower bounds. Interestingly, when α is getting closer to 2 (starting with 1.5 in our table), the empirical standard deviations produced by the Monte Carlo indirect inference procedure are smaller than the asymptotic bounds. This is a finite sample phenomenon. As we increase the number of observations from 1000 to 5000 the indirect inference standard deviation becomes higher than the lower bound while remaining close to it.

6 Issues regarding Dependent Processes

Our method has been applied to *i.i.d.* stable processes. In most financial applications, stable distributions have been used to characterize unconditional distributions of asset

[20] As noted above, continuous GMM has been performed with a fixed ad hoc regularization coefficient. An endogenous choice of β_n could improve results, namely suppress the bias for σ. However, in the Monte Carlo study performed by Carrasco and Florens (2002) for the α-stable distribution, the estimated σ does not change significantly when β_n is selected in an ad hoc way or endogenously.

returns exhibiting skewness and excess kurtosis. This is partly rooted in the Generalized Central Limit Theorem (GCLT) of Feller (1971), which states that if the distribution of a sum of *i.i.d.* random variables exists then it must be a member of the stable Paretian class of distributions. Ibragimov and Linnik (1971) have generalized the GCLT to allow for some forms of temporal dependence. Therefore, looking at say monthly returns, which are aggregated for higher-frequency daily returns, the data may appear as being generated by a stable distribution.

Several papers have investigated the relationship between stable processes and processes with conditional heteroskedasticity such as GARCH and IGARCH. De Vries (1991) has shown that under certain conditions on the parameters of a GARCH-like process, the stable and GARCH processes are observationally equivalent from the viewpoint of the unconditional distribution. Ghose and Kroner (1995) establish that many of the properties of stable models are shared by GARCH models. In particular, both models share the facts that the unconditional distribution has fat tails and that the tail shape is invariant under addition.[21] However, they identify distinctive properties, namely the clustering in volatility that is not present in stable distributions and the distributions of the extreme values, captured by their tail indices. Indeed, even though both models are heavy-tailed, GARCH models allow for bounded second and higher moments, while non-Gaussian stable laws exhibit infinite variance. Groenendijk, Lucas and de Vries (1995) show however that it is not always the case that the tail shapes can be used to discriminate between the competing models. More recently, Deo (2000, 2002) has devised an estimation procedure for the tail index α as well as a goodness-of-fit test that is valid in the presence of m-dependence in the series.

To illustrate the results put forward in this literature, we generate GARCH(1,1) series for the same set of parameter values chosen by Ghose and Kroner (1995) (see the top part of Table 2, p. 234). The models involve both Gaussian and Student densities, aggregated over every 5, 10 and 20 periods.[22] The main difference being that instead of computing the Hill estimator to measure the tail index, we will apply our indirect inference method to estimate α, along with the other parameters.[23]

We simulate 500 samples of 1000 observations from the following GARCH(1,1) model:

$$
\begin{aligned}
y_t &= \varepsilon_t, \ \varepsilon_t \sim D(0, h_t) \\
h_t &= \delta_0 + \delta_1 h_{t-1} + \delta_2 \varepsilon_{t-1}^2.
\end{aligned}
\tag{6.3}
$$

[21] However strong GARCH models are not closed under aggregation and therefore do not share the self-similarity property of the stable distribution. A weak GARCH model is always close under aggregation (Drost and Nijman, 1993) but since our main interest is tail behavior, we consider only the strong version.

[22] If the original process is thought to be daily, aggregation over 5, 10 and 20 periods would represent weekly, biweekly and monthly series.

[23] The tail measure by the Hill estimator is not the same as α. The Hill estimator is the ML estimator of a Pareto density that is often used to estimate the decay of the asymptotic tail of a density but only focusing on the tail area, independently of the rest of the density. Moreover, the Hill estimator has support equal to the positive real numbers while α is bounded between zero and two.

where D is either Gaussian or Student with 5 degrees of freedom. Table 11 shows the mean, standard deviation and kurtosis of the 500 simulated processes. We only show results for the second and fourth moments as the first and third moments do not play any significant role in this simulation. Table 11 shows that all processes exhibit empirically excess kurtosis. Excess kurtosis also increases with the memory of the model (the closer the sum of δ_1 and δ_2 is to one) and does not vanish under aggregation. This is a finite sample artifact, since we know from Diebold (1988) that stationary GARCH models converge to normality under temporal aggregation. If we assume that the original frequency of the generated data is daily, aggregation over 20 observations results in monthly data. For a sample size of 1000 daily observations, we obtain 50 monthly observations. This is indeed a small sample, but small samples are characteristic of hedge fund data that will be used in our empirical application in the next section. Of course, conditional Student t_5 distributions exhibit stronger kurtosis. Regarding the variance, we see that, as expected, it increases with the memory of the model and the aggregation frequency.

Tables 12 and 13 show the mean and standard deviation of 500 estimates of α and σ obtained through our constrained indirect inference procedure. Based on the empirical moments, it is clearly seen that the stable density captures very well the increases of variance and kurtosis through aggregation and memory. The tail index remains relatively constant under aggregation while the estimated dispersion increases -and in all cases the standard errors increase with the aggregation as the sample size decreases. As expected, the tail index and the dispersion are higher when the process is generated from a Student density that when it comes from a Gaussian probability distribution. Last, the higher the memory of the model, in particular for the last three cases, the lower the tail index and the higher the dispersion as well.

We have illustrated that stable distributions could serve as a good statistical tool to capture the unconditional distribution of asset returns at low frequency, even if the true DGP was a conditionally heteroscedastic process like GARCH. Of course, one might argue that what matters for financial applications is the conditional distribution and that the estimation method developed here will not be useful to capture the conditional dependence in mean and variance. Although we do not intend to show in detail how our method can be extended to characterize conditional distributions, we will argue that it should be possible to introduce dependence in our indirect inference procedure. First, it should be stressed that GARCH-like processes, that is processes exhibiting conditional heteroskedasticity, can be built from stable random variables as shown by de Vries (1991). Second, Deo (2002) has used this conditionally heteroscedastic process with marginal stable distribution, which has infinite dependence, to estimate tail indices with an estimation procedure that accommodates dependence of order m. He concludes that his procedure provides quite good estimates of the tail index. As mentioned in Section 4, it does not seem too difficult to extend the general theorems provided in Appendix A to include m-dependence. Given the results in Deo (2002), it can be hoped that it will behave well in presence of processes with infinite dependence. Of course, if the goal is to characterize a conditional distribution, one should change the auxiliary model to incorporate the skewed-t distribution into a GARCH-model. A important gain of estimating the parameters of a stable process could be to use the aggregation property of stable processes for portfolio applications.

7 An Empirical Application to Hedge Funds

In this section we will apply our indirect inference estimation procedure to a set of hedge fund indices. Several studies, in particular Fung and Hsieh (2002), Mitchell and Pulvino (2001), and Agarwal and Naik (2004), have put forward the nonlinear structure and option-like features of returns associated with hedge fund strategies. It means that these return distributions are likely to exhibit skewness and kurtosis and are therefore potentially well captured by stable distributions.

Hedge fund returns are computed from the TASS database, which provides returns and net asset value data on 4,606 funds starting February 1977. However, reliable indices can only be computed starting in the mid-nineties. Therefore, our sample starts in January 1996 and ends in March 2004.[24] Other information includes for each fund an entry date, an exit date (if any), first reporting date, reasons for a fund death if necessary, lock-up periods. This information is useful to correct the data for two well-known biases associated with hedge fund data. The first is a backfilling bias, whereby the database backfills the historical return data of a fund before its entry into the database. The bias comes from the fact that the entry generally follows a period of good returns. The second bias is a survivorship bias. Many funds disappear from the database during the sample period for various reasons such as fund liquidation, fund not reporting any longer to the database, no answer from the fund managers, merger with another fund, to name a few. Not all these reasons have the same consequences in terms of monetary loss for the investor.

We built three types of indices of hedge funds based on different methodologies used by various index producing firms. The Standard & Poor's Hedge Fund Index (SP) is based on a selection of styles and strategies to construct a representative index. [25] The Hedge Fund Research (HFR) indexes are equally weighted and therefore weight relatively more the returns of small hedge funds. The Credit Suisse First Boston/Tremont (TRE) index is a value weighted (i.e. valued by the net asset value of the fund), which gives relatively more weight to the large funds. We follow the respective rules for constructing these three indexes from the individual funds. We report the gross returns (letter O

[24]We chose to start in 1996 to have a reasonable representation for all categories of funds. Also, prior to 1994, the TASS database did not give any information on funds that disappeared from the database.

[25]The styles are Arbitrage, Event-Driven and Directional/Tactical. The strategies for arbitrage are equity market neutral, fixed income arbitrage and convertible arbitrage, for Event-Driven, merger arbitrage, distressed and special situations, for Directional/Tactical, equity long/short, managed futures and macro. Different numbers of funds are selected in each category to arrive at a total of 40 funds.

after each index mnemonic) associated with the three indices, but we also construct series that are free from the two above-mentioned biases. For the backfilling bias, we eliminate all data that precede the date of entry of the fund in the database. The returns are corrected for the survivorship bias by applying a loss of 25 % when the indicated reasons for not reporting are fund liquidation, fund not reporting to TASS, managers not answering requests, and other. In all other cases we did not apply any loss. Summary statistics for the corrected returns for both biases (C after each index mnemonic) and gross returns are reported in Table 14.

These descriptive statistics show that the monthly mean returns are much lower when the two bias corrections are applied. The corrected returns still have a positive mean but barely spectacular. As expected, the TRE index exhibits the highest returns, HFR and SP ranking second and third. Limiting the included categories in the S&P index lowers the mean return but it also reduces the standard error. The standard errors of the HFR and TRE indices are roughly fifty per cent larger than the SP standard errors. The corrected indexes exhibit higher standard errors than the respective gross return indexes. This is explained by the fact that adding losses for the survivorship bias tends to increase dispersion. The corrections make also the fund indexes less Gaussian since both the skewness and kurtosis increase in absolute value. For the gross returns, the skewness is almost always positive. However, when the corrections are applied, there is a clear distinction between the equally weighted (SP and HFR) and the value-weighted TRE index. The skewness of the former becomes negative, while for the latter, it remains positive and not significantly different. This is explained by a maximum return which is roughly double the minimum return in absolute value. This tends to show that the large funds are less susceptible to be affected by the survivorship bias and that they provide a better performance than the smaller funds. Finally, all the indexes show excess kurtosis, but the effect is more pronounced when the two corrections are applied. To control for the exclusion of certain categories in the SP index we also constructed HFR and TRE indexes with the same styles and strategies than the SP. They are denoted with an added SP mnemonic. This tends to increase skewness towards the positive values and to lower kurtosis for the equally weighted HFR and increase kurtosis for the value-weighted TRE.

Before estimating the parameters of a stable distribution for each series, we apply Engle's (1982) Lagrange Multiplier (LM) test to each series to test for the presence of correlation in the second moments. Ghose and Kroner (1995) check by simulation that the test behaves well to discriminate between stable and conditionally heteroscedastic processes. The values of the LM statistic for each of the hedge fund indices are reported in Table 14. We reject the presence of ARCH effects in 8 of the 10 fund indices. Ghose and Kroner (1995) also compute the Hill estimator to measure the tail thickness and use it as a second indicator to differentiate stable distributions from GARCH-type models. We compute the values of the Hill estimator for both the left and right tails of the series. The estimated values are also reported in Table 14. The values obtained for the left and the right tails are strikingly different. While the left tail values are often lower than two, the values in the right tails are always greater than two. This evidence is consistent with an asymmetric stable distribution since it admits different asymptotic left and right tails (see section 2.1). It is not consistent with a Gaussian or Student-t GARCH model.

Table 15 shows the estimated parameters of the stable distribution for the various indices . Our main conclusions are as follows. First, estimated parameters with our constrained indirect inference method are in line with the empirical moments. Consider, for example, both TREC indices. They exhibit the highest values for the empirical kurtosis (5.52 for TREC and 6.2 for TREC-SP), and their $\hat{\alpha}$ is respectively equal to 1.70 and 1.67, the smallest estimated values for the α parameter. For the skewness parameter β, the sign and magnitude observed in the data is generally respected in the estimates. An exception seems to be for the two TREC indices. Finally the means are generally in line with the empirical moments.

Moreover, when the data are close to normal, which is reflected in the empirical moments by a kurtosis close to 3, the estimates for α are close to 2 (as in SPO and HFRO), which is what we want. In this case, the skewness parameter β becomes irrelevant and the estimates appear to be close to 1. Note that in these cases the empirical quantile method reaches the upper bound of 2 for $\hat{\alpha}$, and the β parameter is then set to zero. Observe also that the means tend to be underestimated by the empirical quantile method and overestimated by the indirect inference method. In general, as noted in the Monte Carlo experiment, the values estimated with the quantile method are close to the constrained indirect inference estimates.

The CGMM appears to be much more unstable than the two other methods. When it works, as for example in the case of the SPC index, it provides estimates that are very much in line with the two other methods and the empirical moments. There are cases however, most notably when the data are getting close to normal, where the estimated values for the mean parameter become unreasonable. As noted in the Monte Carlo study, continuous GMM has been performed with a fixed ad hoc regularization coefficient. An endogenous choice of β_n could improve results.

Finally, a comparison of the estimates provided by the constrained indirect inference and the skewed-t is also instructive. When $\hat{\alpha}$ gets closer to 2, the number of degrees of freedom becomes high, as expected. It is much lower for the lowest estimates of α for the TREC and TREC-SP indexes. The skewness parameter γ is also in line with the estimates of β. However, it is worth to recall that an $\hat{\alpha}$ close to 2, but not 2, implies infinite variance while its $\hat{\nu}$ counterpart typically takes values around 9, which implies finite first eight moments and, in particular, finite variance. The standard deviation estimates of the skewed-t are much closer to the corresponding empirical moment. However, rather surprisingly, the estimates of the mean parameter are close to 0.5 for all series. Figure 4 shows nonparametric densities from simulated data of a stable and a skewed-t densities evaluated at the estimated parameters. Indeed, the mode of the skewed-t is systematically on the left of the mode of the stable density. Nevertheless, the essential features of both densities are similar in the sense that when one density is very spiked and has very long tails, the other density also mimics this aspects. This is an intuitive indication that the skewed-t density is indeed a good auxiliary model for the stable.

8 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 used a variant of the indirect inference method called constrained indirect inference. We also showed that this new method for estimating stable distributions proved very useful for capturing the skewness and kurtosis present in hedge funds returns series. Computation of values at risk based on this indirect inference method may also deliver more reliable estimates.

Besides a specific application of constrained indirect inference associated to the skewed-t based QML estimator, this paper provides a unified framework for estimation and goodness of fit tests of stable stochastic processes through a variety of popular instrumental parameters like quantiles or selected values of the characteristic function. Any instrumental parameter for which a consistent asymptotically normal estimator is available is a suitable candidate. For finite sample performance however, it is recommended to choose a reduced number of parameters which are well informative about the structural ones. This paper contains both theoretical arguments and Monte-Carlo evidence to show that specific (constrained) functions of quantiles as considered by McCulloch (1986) as well as skewed-t (constrained) QML estimators are good candidates.

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:

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

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

We first show that:

$$\begin{cases} Q\left[(\nu,\gamma,\omega+a,\lambda);(\alpha,\beta,\mu+a,\sigma)\right]=Q\left[(\nu,\gamma,\omega,\lambda);(\alpha,\beta,\mu,\sigma)\right] \text{ for all } a, \\ Q\left[(\nu,\gamma,\omega a,a\lambda);(\alpha,\beta,a\mu,a\sigma)\right]=Q\left[(\nu,\gamma,\omega,\lambda);(\alpha,\beta,\mu,\sigma)\right] \text{ for all } a>0, \\ Q\left[(\nu,1/\gamma,2\mu-\omega,\lambda);(\alpha,-\beta,\mu,\sigma)\right]=Q\left[(\nu,\gamma,\omega,\lambda);(\alpha,\beta,\mu,\sigma)\right] \end{cases} \tag{A.1}$$

To see this, it is sufficient to notice that:

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

$$l\left(y+a;\nu,\gamma,\omega+a,\lambda\right)=l\left(y;\nu,\gamma,\omega,\lambda\right).$$

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

$$l\left(ay;\nu,\gamma,a\omega,\lambda\right)=l\left(y;\nu,\gamma,\omega,\lambda\right).$$

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

$$l\left(2\mu-y;\nu,1/\gamma,2\mu-w,\lambda\right)=l\left(y;\nu,\gamma,\omega,\lambda\right).$$

This proves (A.1).

Suppose for the moment that we can also prove that, for maximizing the quasi-likelihood function, the first-order condition with respect to ν is tantamount to the equation stated by proposition 3.4:

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

Then, by joint application of (A.1) and (A.2), we see clearly that the changes $Y \longrightarrow Y+a$, $Y \longrightarrow aY$ and $Y \longrightarrow 2\mu-Y$ will have the effects on the 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 the quasi-likelihood maximization with respect to λ and ν. The partial derivatives of the expected log-quasi-likelihood function with respect to λ and ν are:

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

and

$$\begin{aligned}\frac{\partial Q}{\partial \nu} &= \frac{h'(\nu)}{h(\nu)} - \frac{1}{2\nu} - \frac{1}{2}E\left[\log\left[1+\frac{1}{\nu}\left(\frac{Y-\omega}{\lambda}\right)^2 g_w(Y,\gamma)\right]\right] \\ &\quad -\frac{\nu+1}{2\lambda^2}\left(-\frac{1}{\nu^2}\right) E\left[\frac{(Y-\omega)^2 g_\omega(Y,\gamma)}{1+\frac{1}{\nu}\left(\frac{Y-\omega}{\lambda}\right)^2 g_w(Y,\gamma)}\right].\end{aligned}$$

This leads to the following first order conditions, for λ and ν respectively:

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

and:

$$\begin{aligned}&\frac{h'(\nu)}{h(\nu)} - \frac{1}{2\nu} - \frac{1}{2}E\left[\log\left[1+\frac{1}{\nu}\left(\frac{Y-\omega}{\lambda}\right)^2 g_w(Y,\gamma)\right]\right] \\ &\qquad + \frac{\nu+1}{2\nu}E\left[\frac{\frac{1}{\nu}\left(\frac{Y-\omega}{\lambda}\right)^2 g_\omega(Y,\gamma)}{1+\frac{1}{\nu}\left(\frac{Y-\omega}{\lambda}\right)^2 g_w(Y,\gamma)}\right] = 0\end{aligned} \tag{A.4}$$

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

Q.E.D

Proof of asymptotic normality

As explained in section 4, we have only to show that the estimator $\hat{\Psi}$ of the pseudo-true value Ψ^0 of Ψ defined by first-order conditions (4.3) is root n asymptotically normal. For sake of expositional simplicity, we address the asymptotic normality issue separately for each component of $\hat{\Psi}$, by considering the corresponding first-order conditions in (4.3). A closer argument to check the joint asymptotic normality of $\hat{\Psi}$ would be easy to settle at the cost of cumbersome matrix notations. Moreover, we consider only the case of *i.i.d.* sampling, with application of the Lindeberg-Levy central limit theorem. Extensions using convenient central limit theorems for stationary dependent data with high-level mixing assumptions would be straightforward.

Starting from formula (4.2) for the quasi-loglikelihood function, the estimator $\hat{w}$ is characterised as solution of:

$$\frac{\partial L_n}{\partial w}\left(\hat{\Psi}\right) = \frac{\hat{\nu}+1}{2}\sum_{i=1}^{n}\frac{\frac{1}{\hat{v}}g_w\left(y_i,\hat{\gamma}\right)2\frac{(y_i-\hat{w})}{\hat{\lambda}^2}}{1+\frac{1}{\hat{v}}g_w\left(y_i,\hat{\gamma}\right)\left(\frac{y_i-\hat{w}}{\hat{\lambda}}\right)^2} = 0$$

$$\Leftrightarrow \hat{w} = \left[\sum_{i=1}^{n}\frac{y_i g_w\left(y_i,\hat{\gamma}\right)}{1+\frac{1}{\hat{\nu}}g_w\left(y_i,\hat{\gamma}\right)\left(\frac{y_i-\hat{w}}{\hat{\lambda}}\right)^2}\right]\times\left[\sum_{i=1}^{n}\frac{g_w\left(y_i,\hat{\gamma}\right)}{1+\frac{1}{\hat{\nu}}g_w\left(y_i,\hat{\gamma}\right)\left(\frac{y_i-\hat{w}}{\hat{\lambda}}\right)^2}\right]^{-1} \tag{A.5}$$

The variables

$$\frac{g_w\left(y_i,\gamma\right)}{1+\frac{1}{\gamma}g_w\left(y_i,\gamma\right)\left(\frac{y_i-w}{\lambda}\right)^2}$$

are $i.i.d$ and bounded (absolute value smaller than max $\left[\gamma^2,\frac{1}{\gamma^2}\right]$). Thus, by the uniform strong law of large numbers:

$$\frac{1}{n}\sum_{i=1}^{n}\frac{g_w\left(y_i,\hat{\gamma}\right)}{1+\frac{1}{\hat{\nu}}g_w\left(y_i,\hat{\gamma}\right)\left(\frac{y_i-\hat{w}}{\hat{\lambda}}\right)}$$

is asymptotically a positive constant. Consider now the sequence of variables

$$Z_i = \frac{y_i g_w\left(y_i,\gamma\right)}{1+\frac{1}{\gamma}g_w\left(y_i,\gamma\right)\left(\frac{y_i-w}{\lambda}\right)^2}$$

Since:

$$|Z_i| \leq |Y_i|\,Max\left(\gamma^2,\frac{1}{\gamma^2}\right)$$

Z_i is $i.i.d.$ integrable as Y_i (since $\alpha > 1$) and, by the Lindeberg-Levy central limit theorem, $\frac{1}{\sqrt{n}}\sum_{i=1}^{n}(Z_i - EZ_i)$ is asymptotically normal. Therefore: $\sqrt{n}\left(\hat{w}-w^0\right)$ is asymptotically normal.

From (4.2), $\hat{\gamma}$ is characterized as solution of the first-order condition:

$$\begin{aligned}\frac{1-\frac{1}{\hat{\gamma}^2}}{\hat{\gamma}+\frac{1}{\hat{\gamma}}} &= \frac{\hat{\nu}+1}{2n}\sum_{i/y_{i\geq 0}}\frac{\frac{1}{\hat{\nu}}2\hat{\gamma}(\frac{y_i-\hat{w}}{\hat{\lambda}})^2}{1+\frac{\hat{\gamma}^2}{\hat{\nu}}\left(\frac{y_i-\hat{w}}{\hat{\lambda}}\right)^2} \\ &\quad +\frac{\hat{\nu}+1}{2n}\sum_{i/y_{i<0}}\frac{\frac{1}{\hat{\nu}}\left(\frac{-2}{\hat{\gamma}^3}\right)(\frac{y_i-\hat{w}}{\hat{\lambda}})^2}{1+\frac{1}{\hat{\nu}\hat{\gamma}^2}\left(\frac{y_i-\hat{w}}{\hat{\lambda}}\right)^2}\end{aligned} \tag{A.6}$$

Lindeberg-Levy central limit theorem will apply to the sum of the RHS of (4.2) since variables of the form

$$\frac{(y_i - w)^2}{1 + a\left(\frac{y_i - w}{\lambda}\right)^2}$$

are integrable when taking values in the bounded interval $\left[0, \frac{1}{a}\right]$. Therefore, by the delta-theorem, equation (A.2) will ensure asymptotic normality for $\hat{\gamma}$ insofar as:

$$g(\gamma) = \frac{1 - \frac{1}{\gamma^2}}{\gamma + \frac{1}{\gamma}}$$

is a C^1 diffeomorphism in the neighborhood of the pseudo-true value. But:

$$g'(\gamma) = \frac{\frac{1}{\gamma^2} + \frac{4}{\gamma^2} - 1}{\left(\gamma + \frac{1}{\gamma}\right)^2} < 0$$

insofar as $\frac{1}{\gamma^2} > \sqrt{5} - 2$ that is $\gamma^2 < 2 + \sqrt{5}$.

Therefore, asymptotic normality of $\hat{\gamma}$ is guaranteed insofar as the pseudo true-value of γ is smaller than $\left(2 + \sqrt{5}\right)^{1/2}$.

From (4.2), $\hat{\lambda}$ is characterized as solution:

$$\hat{\lambda}^2 = \frac{\hat{\nu} + 1}{2\hat{\nu}} \frac{1}{n} \sum_{i=1}^{n} \frac{(y_i - \hat{w})^2 g_{\hat{w}}(y_i, \hat{\gamma})}{1 + \frac{1}{\hat{\nu}}\left(\frac{y_i - \hat{w}}{\hat{\lambda}}\right)^2 g_{\hat{w}}(y_i, \hat{\gamma})} \tag{A.7}$$

The positive variables:

$\frac{(y_i - w)^2 g_{\hat{w}}(y_i, \gamma)}{1 + \frac{1}{\hat{\nu}}\left(\frac{y_i - w}{\hat{\lambda}}\right)^2 g_{\hat{w}}(y_i, \gamma)}$ are $i.i.d$ and bounded by $\lambda^2 \nu$. Therefore $\hat{\lambda}^2$ is asymptotically normal by the Lindeberg-Levy central limit theorem and so is $\hat{\lambda}$ by the delta theorem since the pseudo true value of λ is positive.

Finally:

$$\begin{aligned} \frac{\partial L_n}{\partial \nu}\left(\hat{\Psi}\right) &= n\frac{h'(\hat{\nu})}{h(\hat{\nu})} - \frac{n}{2\hat{\nu}} \\ &\quad - \frac{1}{2}\sum_{i=1}^{n} \log\left[1 + \frac{1}{\hat{\nu}} g_w(y_i, \hat{\gamma})\left(\frac{y_i - \hat{w}}{\hat{\lambda}}\right)^2\right] \\ &\quad + \frac{\hat{\nu} + 1}{2} \sum_{i=1}^{n} \left[\frac{\frac{1}{\hat{\nu}^2 \hat{\lambda}^2}(y_i - \hat{\gamma})(y_i, \hat{w})^2}{1 + \frac{1}{\hat{\nu}} g_w(y_i, \hat{\gamma})\left(\frac{y_i - \hat{w}}{\hat{\lambda}}\right)^2}\right] \end{aligned} \tag{A.8}$$

By plugging (A.7) into (A.8), we deduce the last first-order condition:

$$\hat{\delta} = \frac{\partial L_n}{\partial \nu}\left(\hat{\Psi}\right) = n\frac{h'\left(\hat{\nu}\right)}{h\left(\hat{\nu}\right)} - \frac{1}{2}\sum_{i=1}^{n}\log\left[1 + \frac{1}{\hat{\nu}}g_w\left(y_i, \hat{\gamma}\right)\left(\frac{y_i - \hat{w}}{\hat{\lambda}}\right)^2\right] \tag{A.9}$$

Consider the sequence if *i.i.d.* variables:

$$V_i = \log\left[1 + \frac{1}{\nu}g_w\left(y_i, \gamma\right)\left(\frac{y_i - w}{\gamma}\right)^2\right].$$

They are integrable since, first, $V_i \longrightarrow 0$ when $y_i \longrightarrow w$ and second, when $|y_i - w| \longrightarrow \infty$, $w_i \sim 2\log|y_i - w|$ which is dominated (for large $|y_i - w|$) by $|y_i - w|$ which is integrable. Therefore, $\left(\hat{\delta}, \hat{\nu}\right)$ are jointly asymptotically normal by application to (A.9) of Lindeberg-Levy central limit theorem and the delta theorem.

Appendix B: Simulating a α-stable distribution

The use of indirect inference or its constrained version necessitates to simulate an α-stable process. For simulating we adopt the method proposed by Chambers, Mallows and Stuck (1976) which is fast and easy to implement. [26]

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 $\alpha \neq 1$,

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

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

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

To generate z and y, we draw two independent uniform$(0,1)$ random variables U_1 and U_2 and set $z = \pi\left(U_1 - \frac{1}{2}\right)$ 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

$$\begin{array}{rl} \sigma X + \mu \sim S_\alpha(\beta, \sigma, \mu) & if \quad \alpha \neq 1, \\ \sigma X + \frac{2}{\pi}\beta\sigma\ln\sigma + \mu \sim S_\alpha(\beta, \sigma, \mu) & if \quad \alpha = 1. \end{array} \tag{B.3}$$

See Samorodnitsky and Taqqu, (1994, p 43).

[26]The GAUSS procedure for simulating from the α-stable process has been written by J. Huston McCulloch and available in his web page http://www.econ.ohio-state.edu/jhm/jhm.html.

Appendix C: Numerical aspects

As 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 α-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).

Second, two of the four parameters in θ_α are constrained. To avoid using a constrained optimization algorithm, we reparametrize the initial parameters. In general if a parameter ϑ 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:

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

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

Acknowledgements

We are grateful to Marine Carrasco, François Lamontagne, Marco Lombardi, Marc Paolella, Mark Steel and Casper de Vries for useful discussions and remarks as well as the audiences at the CIRANO-CIREQ Conference in Financial Econometrics 2004, ESEM 2004, Canadian Econometric Study Group 2004, York University, Conference on Heavy tails and stable Paretian distributions 2005 organized by the Deutsche Bundesbank and the Econometrics Seminar at Tilburg University. Marine Carrasco kindly provided us with the computer code for estimating α-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 and later a Marie Curie postdoctoral fellow at Tilburg University. The third author thanks CIRANO, CIREQ, IFM2 and the European Community's Human Potential Programme under contract HPRN-CT-2002-00232, [MICFINMA] for financial support. This text presents research results of the Belgian Program on Interuniversity Poles of Attraction initiated by the Belgian State, Prime Minister's Office, Science Policy Programming. The scientific responsibility is assumed by the authors.

References

[1] Agarwal, V., and Naik, N.Y. (2004), "Risks and Portfolio Decisions involving Hedge Funds", *Review of Financial Studies*, 17, 63-98.

[2] Ait-Sahalia, Y. and Jacod, J. (2004), "Fisher's Information for Discretely Sampled Lévy Processes", mimeo Princeton Univerisity.

[3] Ait-Sahalia, Y. and Jacod, J. (2004), "Volatility Estimators for Discretely Sampled Lévy Processes", mimeo Princeton Univerisity.

[4] Akgiray, V., and Lamoureux, C.G. (1989), "Estimation of Stable-Law Parameters: A Comparative Study", *Journal of Business and Economic Statistics*, 7, 85-93.

[5] Bassi, F., Embrechts, P. and Kafetzaki, M. (1998), "Risk Management and Quantile Estimation", in *A Practical Guide to Heavy Tails*, edited by R. Adler *et al*, Birkhuser, Boston, 111-130.

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

[7] Bawa, V.S., Elton, E.J. and Gruber, M.J. (1979), "Simple Rules for Optimal Portfolio Selection in Stable Paretian Markets", *Journal of Finance*, 34, 1041-1047.

[8] Bidarkota, P.V., Dupoyet, B.V. and McCulloch, J.H. (2005), "Asset Pricing with Incomplete Information under Stable Shocks", mimeo Florida International University.

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

[10] Carr, P., Geman, H., Madan, D. B. and Yor, M. (2002), "The Fine Structure of Asset Returns: An Empirical Investigation", *Journal of Business*, 75, 305-332.

[11] Carrasco, M. and Florens, J.P. (2000), "Generalization of GMM to a continuum of moment conditions", *Econometric Theory*, 16, 797-834.

[12] Carrasco, M., Chernov, M., Florens, J.P. and Ghysels, E. (2004), "Efficient estimation of jump diffusions and general dynamic models with a continuum of moment conditions", Discussion Paper, University of North Carolina.

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

[14] Carrasco, M., Florens, J.P. and Renault. E. (2005), "Linear Inverse Problems in Structural Econometrics: Estimation based on Spectral Decomposition and Regularization", forthcoming in *Handbook of Econometrics,* vol. 6, J. Heckman (ed.), North Holland.

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

[16] Coppejans, M. and Gallant, A.R. (2002), "Cross-Validated SNP Density Estimates", *Journal of Econometrics*, 110, 27-65.

[17] Deo, R.S. (2000), "On Estimation and Testing Goodness of Fit for m-dependent Stable sequences", *Journal of Econometrics*, 99, 349-372.

[18] Deo, R.S. (2002), "On testing the adequacy of stable processes under conditional heteroscedasticity", *Journal of Empirical Finance*, 9, 257-270.

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

[20] Diebold, F.X. (1988), *Empirical Modelling of Exchange Rates Dynamics*, Berlin: Springer Verlag.

[21] Dridi, R. Guay, A. and R. Renault E. (2005) "Indirect Inference and Calibration of Dynamic Stochastic General Equilibrium Models, mimeo Université de Montreal.

[22] Drost, F.C. and Nijman T.E. (1993), "Temporal Aggregation of GARCH processes", *Econometrica*, 61/4, 909-1027.

[23] Duffee, G. and Stanton R. (2005) "EMM and indirect inference estimation of near unit-root processes", mimeo Haas School of Business, U.C. Berkeley.

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

[25] DuMouchel, W.H. (1975), “Stable Distributions in Statistical Inference: 2. Information from Stably Distributed Samples”, *Journal of the American Statistical Association,* 70, n^o 350, 386-393.

[26] Engle, R.F. (1982), “Autoregressive Conditional Heteroscedasticity and Estimates of U.K. Inflation”, *Econometrica,* 50, 987-1008.

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

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

[29] Fama, E. and Roll, R. (1971), “Parameter Estimates for Symmetric Stable Distributions”, *Journal of the American Statistical Association*, 66, 331-338.

[30] Feller, W. (1971), *An Introduction to Probability Theory and its Applications*, Vol 2, 2nd edn. New York: John Wiley and Sons.

[31] Fernández, C. and Steel, M. (1998), “On Bayesian Modelling of Fat Tails and Skewness”, *Journal of the American Statistical Association*, 93, 359-371.

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

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

[34] Fung, W. and Hsieh, D.A. (2002), “Benchmarks of Hedge Fund Performance: Information Content and Measurement Biases”, *Financial Analyst Journal*, 58, 22-34.

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

[36] Gallant, A.R. and Tauchen, G. (1999), “The Relative Efficiency of Method of Moments Estimators”, *Journal of Econometrics*, 92, 149-172.

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

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

[39] Gouriéroux, C. Touzi, N. and Renault, E. (1999) ”Calibration by Simulation for Small Sample Bias Correction”, in *Simulation Based Inference in Econometrics Methods*, edited by J. Geweke and R. Mariano.

[40] Groenendijk, P.A., Lucas, A. and de Vries, C.G. (1995), "A note on the relationship between GARCH and symmetric stable processes", *Journal of Empirical Finance*, 2, 253-264.

[41] Hansen, B.E. (1994), "Autoregressive Conditional Density Estimation", *International Economic Review*, vol 35, 3, 705-730.

[42] Hill, B.M. (1975), "A Simple General Approach to Inference About the Tail of a Distribution", *Annals of Statistics*, 3, 1163-1174.

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

[44] Koutrouvelis, I.A. (1980) "Regression-type Estimation of the Parameters of Stable Laws", *Journal of the American Statistical Association*, 75, 918-928.

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

[46] Mandelbrot, B. (1963), "The Variation of Speculative Prices", *Journal of Business*, 36, 394-419.

[47] McCulloch, J.H. (1986), "Simple Consistent Estimators of Stable Distribution Parameters", *Communications in Statistics - Simulation and Computation*, 15, 1109-1136.

[48] Mitchell, M.L., and Pulvino, T.C. (2006), "Characteristics of Risk and Return in Risk Arbitrage", Forthcoming in *Journal of Finance*

[49] Mittnik, S., Paolella, M.S. and Rachev, S.T. (2000), "Diagnosis and treating the fat tails in financial returns data", *Journal of Empirical Finance*, 7, 389-416.

[50] Mittnik, S. and Rachev, S.T. (1993), "Modeling Asset Returns with Alternative Stable Distributions", *Econometric Reviews*, 12(3), 261-330.

[51] Mittnik, S., Rachev, S.T., Paolella, M.S. (1998), "Stable Paretian Modeling in Finance: Some Empirical and Theoretical Aspects", in *A Practical Guide to Heavy Tails*, edited by R. Adler, R. Feldman and M.S. Taqqu, Birkhuser, Boston, 79-110.

[52] Mittnik, S., Rachev, S.T, Doganoglu, T. and Chenyao, D. (1999), "Maximum Likelihood Estimation of Stable Paretian Models", *Mathematical and Computer Modelling*, 29, 275-293.

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

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

[55] Press, J.S. (1972). "Estimation of Univariate and Multivariate Stable Distributions", *Journal of the American Statistical Association*, 67 (340), 842-846.

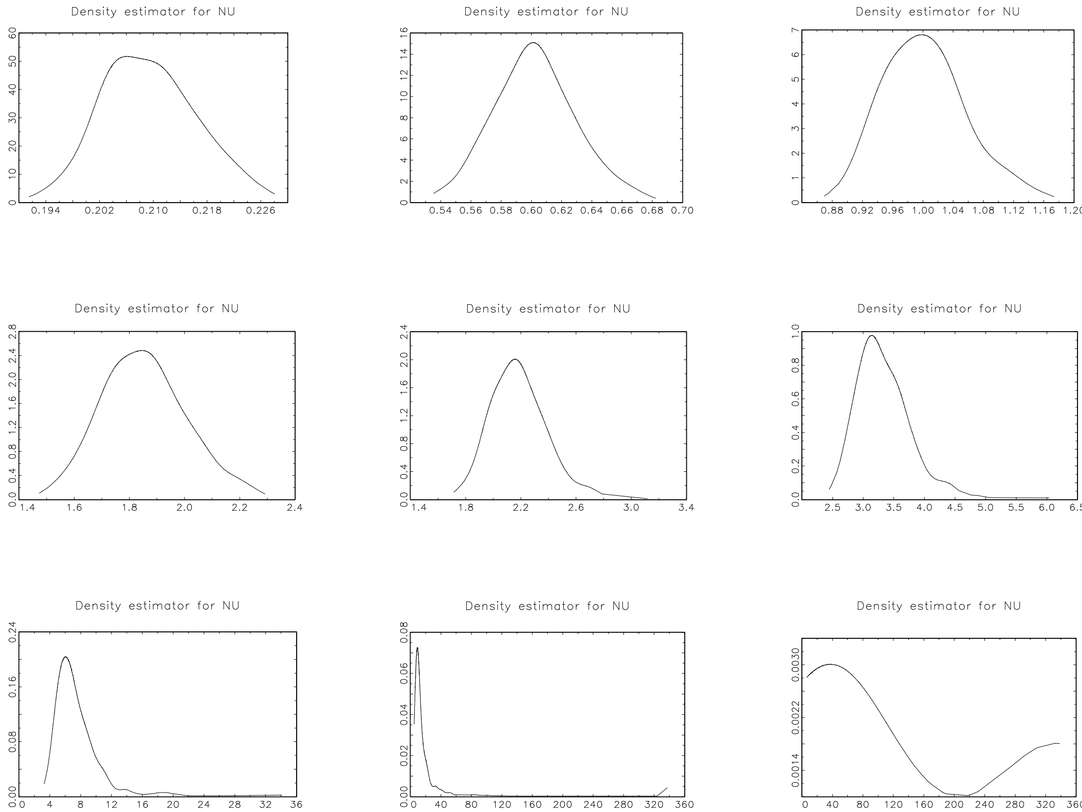
[56] Samorodnitsky, G. and Taqqu, M.S. (1994), *Stable Non-Gaussian Random Processes*. New York: Chapman and Hall.

[57] Singleton, K. (2001). "Estimation of Affine Pricing Models Using the Empirical Characteristic Function", *Journal of Econometrics*, 102, 111-141.

[58] Smith, A. A. (1993). "Estimating Nonlinear Time Series Models Using Simulated Vector Autoregressions", *Journal of Applied Econometrics* 8, S63-S84.

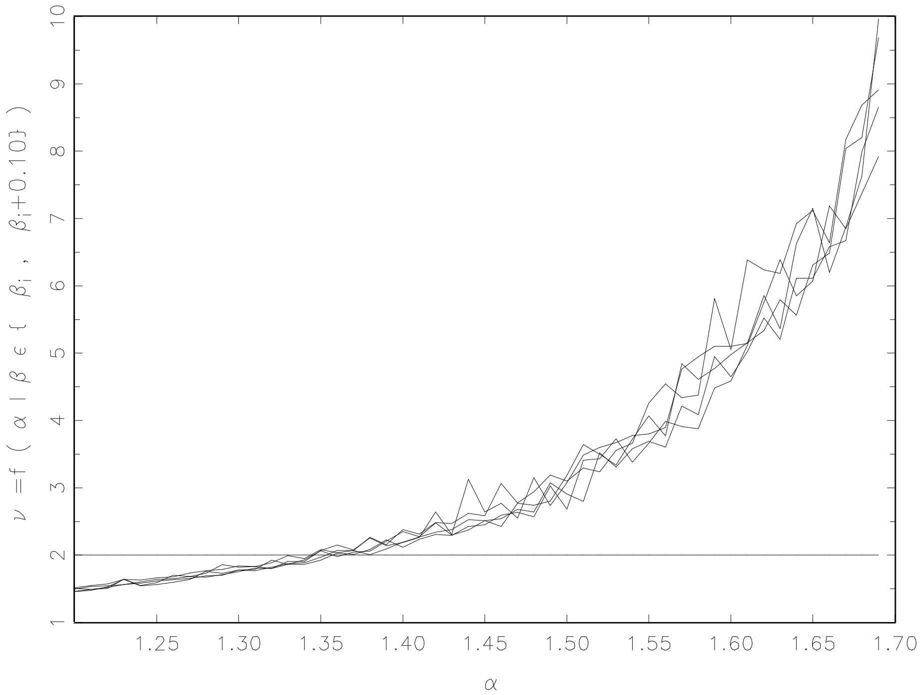
[59] Zolotarev, V.M. (1954). "Vyrazhenie plotnosti ustoichivogo raspredeleniya s pokazatelem α bolshim edinitsy cherez plotnost s pokazatelem $1/\alpha$", *Doklady Akademii Nauk USSR (N.S.)* 98, 735-738.

[60] Zolotarev, V.M. (1986). *One dimensional Stable Distributions*, Amer. Math. Soc. Transl. of Math. Monographs, vol 65, Amer. Math. Soc., Providence, RI. (Transl of the original 1983 Russian).



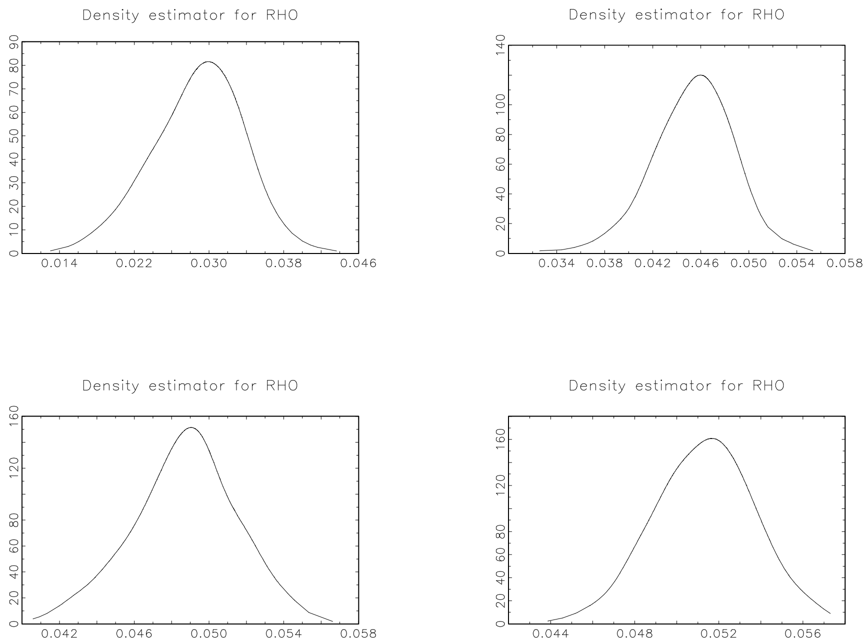
Reading in rows. From top to bottom and from left to right the true α are 0.3, 0.7, 1, 1.3, 1.5, 1.7, 1.9, 1.95 and 1.99.

Figure 1: Kernel Densities for $\widehat{\nu}$



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

Figure 2: $\widehat{\nu}$ sensibility to α and β



Reading in rows. From top to bottom and from left to right the true α are 1.7, 1.9, 1.95 and 1.99.

Figure 3: Kernel Densities for $\hat{\delta}$

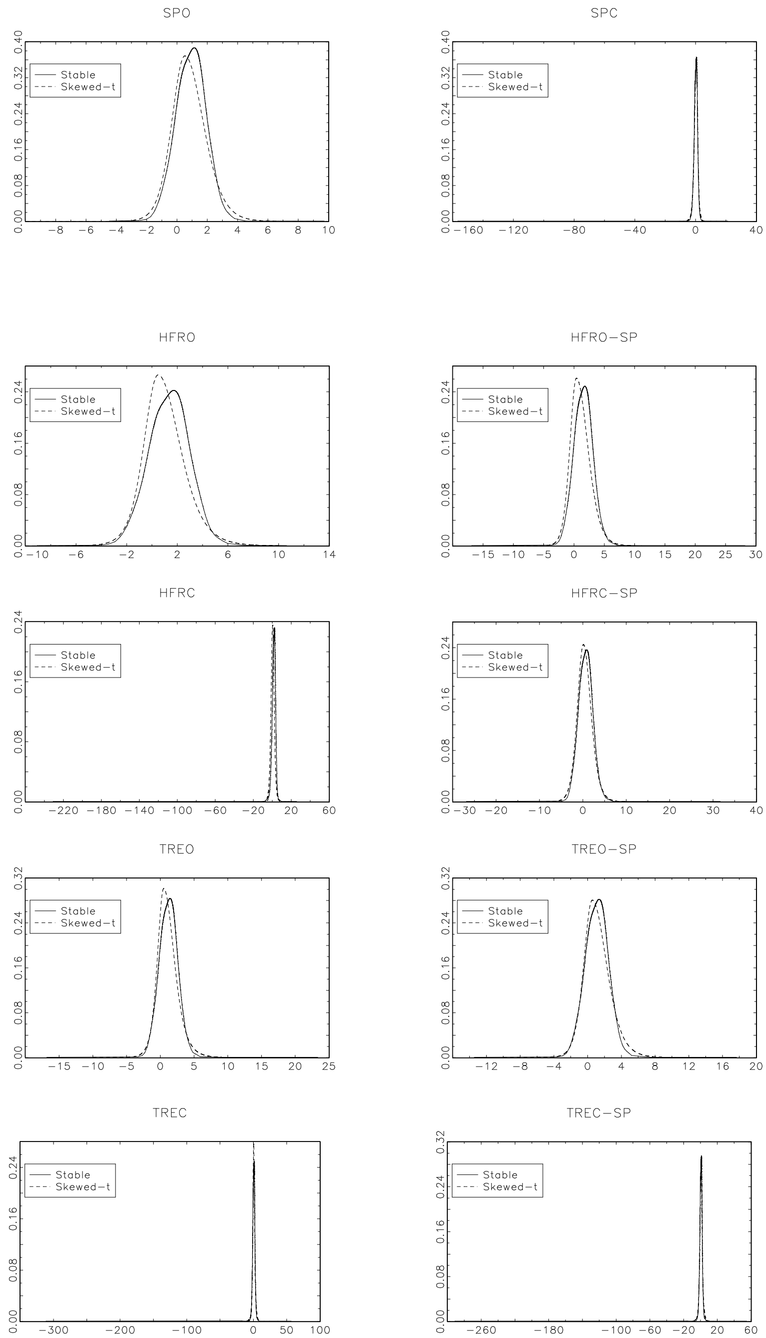


Figure 4: Kernel densities for the Stable and the Skewed-t densities evaluates at the empirical estimates.

Table 1: Simulation Results for $\hat{\phi}$

	Mean		Sd		Skw		Kur		Min		Max	
	$\widehat{\nu}$	$\widehat{\gamma}$	$\widehat{\nu}$	$\widehat{\gamma}$	$\widehat{\nu}$	$\widehat{\gamma}$	$\widehat{\nu}$	$\widehat{\gamma}$	$\widehat{\nu}$	$\widehat{\gamma}$	$\widehat{\nu}$	$\widehat{\gamma}$
$\alpha = 0.3$	0.2094	1.0097	0.0069	0.0311	0.0414	-0.064	2.8717	3.0436	0.1915	0.9053	0.2280	1.1088
$\alpha = 0.7$	0.6022	1.0011	0.0271	0.0337	0.1890	0.1641	2.9371	3.0676	0.5354	0.9137	0.6823	1.1141
$\alpha = 1$	1.0026	0.9996	0.0557	0.0369	0.2454	-0.026	3.0054	3.0559	0.8697	0.8919	1.1733	1.1144
$\alpha = 1.3$	1.8533	1.0051	0.1529	0.0458	0.2742	0.2072	2.8104	3.2842	1.4741	0.8654	2.2915	1.1738
$\alpha = 1.5$	2.1833	1.0027	0.2016	0.0451	0.5870	-0.081	3.8085	2.7993	1.7148	0.8749	3.1170	1.1190
$\alpha = 1.7$	3.3412	1.0008	0.4647	0.0470	1.3859	0.1439	7.1682	2.9253	2.4443	0.8795	6.0397	1.1604
$\alpha = 1.9$	7.5570	1.0024	3.3227	0.0524	3.4285	0.1867	22.054	2.8751	3.2734	0.8668	34.097	1.1832
$\alpha = 1.95$	26.953	1.0005	59.161	0.0517	4.6297	-0.039	23.802	2.7246	4.7906	0.8612	336.99	1.1707
$\alpha = 1.99$	146.36	1.0036	142.75	0.0515	0.4927	-0.063	1.3611	2.7554	6.2895	0.8489	337.99	1.1541
	$\widehat{\lambda}$	$\widehat{\omega}$	$\widehat{\lambda}$	$\widehat{\omega}$	$\widehat{\lambda}$	$\widehat{\omega}$	$\widehat{\lambda}$	$\widehat{\omega}$	$\widehat{\lambda}$	$\widehat{\omega}$	$\widehat{\lambda}$	$\widehat{\omega}$
$\alpha = 0.3$	0.0817	0.0002	0.0114	0.0026	0.0619	0.1463	3.2397	3.0407	0.0500	-0.008	0.1224	0.0091
$\alpha = 0.7$	0.3894	0.0005	0.0277	0.0211	0.0509	-0.097	2.8330	2.9211	0.3201	-0.078	0.4657	0.0562
$\alpha = 1$	0.4986	-0.001	0.0268	0.0331	-0.043	0.0027	3.5540	3.1587	0.4106	-0.105	0.5974	0.0930
$\alpha = 1.3$	0.5724	-0.001	0.0243	0.0460	-0.051	-0.099	3.0005	3.1720	0.4955	-0.143	0.6429	0.1574
$\alpha = 1.5$	0.5849	-0.001	0.0230	0.0459	0.1553	-0.015	2.9353	2.8992	0.5195	-0.139	0.6561	0.1344
$\alpha = 1.7$	0.6121	-0.001	0.0230	0.0523	0.1341	0.0860	3.2199	2.8318	0.5527	-0.127	0.6978	0.1426
$\alpha = 1.9$	0.6477	-0.001	0.0266	0.0615	-0.115	-0.017	3.0180	2.8757	0.5627	-0.210	0.7197	0.1457
$\alpha = 1.95$	0.6667	0.0003	0.0279	0.0596	0.1114	0.0169	2.6477	2.7914	0.5945	-0.153	0.7453	0.2066
$\alpha = 1.99$	0.6903	-0.002	0.0236	0.0622	-0.066	0.2484	3.7866	2.8251	0.5862	-0.178	0.7634	0.2032

The top part of the Table are results for $\widehat{\nu}$ and $\widehat{\gamma}$, for each statistic the first column is for $\widehat{\nu}$ and the second for $\widehat{\gamma}$. Bottom part are for $\widehat{\lambda}$, first column, and $\widehat{\omega}$, second column. μ, β and σ 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 ν to the sample size and α

α		N=100	200	500	1000	5000	10000
1	Sd	0.2298	0.1353	0.0828	0.0569	0.0243	0.0173
	Skw	0.9070	0.6197	0.3841	0.3316	0.0549	0.0245
	Kur	4.4173	3.6194	3.2138	3.0485	3.0959	3.0330
1.3	Sd	0.5489	0.2970	0.1736	0.1171	0.0503	0.0366
	Skw	2.0973	0.9792	0.6197	0.3542	0.3323	0.2224
	Kur	11.535	4.8421	3.3634	3.5739	3.0352	2.9433
1.5	Sd	21.281	0.5676	0.3381	0.2004	0.0879	0.0635
	Skw	21.895	2.2189	1.2433	0.8046	0.3184	0.1396
	Kur	486.06	12.361	7.8417	4.9488	3.3858	2.8486
1.7	Sd	62.268	18.792	0.7770	0.4968	0.1931	0.1326
	Skw	4.7446	19.820	1.3933	0.9682	0.4147	0.1893
	Kur	23.851	442.04	6.2004	4.5652	3.2662	3.0705
1.9	Sd	145.39	105.47	38.823	15.467	0.8323	0.5825
	Skw	0.8854	2.2574	7.1886	3.0826	0.9205	0.7833
	Kur	1.8296	6.2649	56.530	16.532	4.4781	4.1518
1.95	Sd	158.77	145.75	102.03	58.269	2.5094	1.4378
	Skw	0.2557	0.9314	2.4688	5.5896	2.7023	0.7988
	Kur	1.1004	1.9366	7.3776	34.003	19.298	3.7519
1.99	Sd	156.38	148.11	150.54	143.28	86.330	55.369
	Skw	-0.306	-0.045	0.0496	0.4791	2.3712	1.2311
	Kur	1.1185	1.2754	1.0784	1.3418	7.4779	5.8909

Table 3: Simulation Results Using Indirect Inference (h=1)

		Mean		Sd		Skw		Kur		Min		Max	
		$\hat{\alpha}$	$\hat{\beta}$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\alpha}$	$\hat{\beta}$
$\alpha = 1.5$	$\beta = 0$	1.47593	-0.1169	0.04659	0.09317	0.03458	0.28248	2.94244	3.14196	1.34871	-0.4077	1.61019	0.18980
	$\beta = 0.75$	1.49288	0.70918	0.05816	0.11778	0.26090	0.57982	3.16809	3.06127	1.37248	0.40753	1.66455	1
$\alpha = 1.9$	$\beta = 0$	1.82005	0.16182	0.79555	0.21406	-3.1243	0.16499	22.9914	2.66551	1.19065	-0.5400	2	0.84090
	$\beta = 0.75$	1.85128	0.73587	0.08693	0.21192	-0.1666	-1.1151	1.90220	6.72215	1.74232	-1	2	1
		$\hat{\sigma}$	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\mu}$
$\alpha = 1.5$	$\beta = 0$	0.50084	-0.0438	0.01657	0.04324	0.13835	0.12029	3.45820	3.29292	0.45010	-0.1742	0.55768	0.12496
	$\beta = 0.75$	0.50465	0.00084	0.02011	0.09918	0.56922	1.59753	3.78103	5.51234	0.45319	-0.1502	0.57278	0.36556
$\alpha = 1.9$	$\beta = 0$	0.50257	0.11925	0.01378	0.09326	-0.1399	1.32183	3.48464	7.39204	0.45518	-0.0373	0.55010	0.73419
	$\beta = 0.75$	0.50213	-0.0662	0.22463	2.25457	-12.920	-21.993	241.835	489.327	0.08336	-50.208	0.54151	3.21255

The top part of the Table are results for $\widehat{\alpha}$ and $\widehat{\beta}$, for each statistic the first column is for $\widehat{\alpha}$ and the second for $\widehat{\beta}$. Bottom part are for $\widehat{\sigma}$ and $\widehat{\mu}$. μ and σ 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 δ to the sample size and α

α		N=100	200	500	1000	5000	10000
1.7	Sd	0.3967	0.1851	0.0070	0.0045	0.0022	0.0015
	Skw	4.2254	9.7938	-0.338	-0.180	-0.053	0.0413
	Kur	18.954	97.537	3.4076	2.7903	2.9022	2.9681
1.9	Sd	0.0094	0.0068	0.0044	0.0031	0.0015	0.0010
	Skw	-0.714	-0.436	-0.297	-0.149	-0.230	-0.179
	Kur	3.9675	3.4135	2.9608	2.8653	3.3693	3.0314
1.95	Sd	0.0084	0.0058	0.0038	0.0028	0.0012	0.0008
	Skw	-0.585	-0.131	-0.377	-0.252	-0.123	0.1181
	Kur	3.7231	2.9503	3.1520	3.4036	3.1047	2.9901
1.99	Sd	0.0067	0.0053	0.0034	0.0023	0.0010	0.0007
	Skw	-0.233	-0.406	-0.125	0.0145	-0.007	-0.020
	Kur	2.9394	3.2607	2.7799	3.2751	3.4280	3.1732

Table 5: Simulation Results Using Constrained Indirect Inference ($h = 1$)

α	β	Mean		Sd		Skw		Kur		Min		Max	
		$\hat{\alpha}$	$\hat{\beta}$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\alpha}$	$\hat{\beta}$
$\alpha = 0.7$	$\beta = 0$	0.69914	-0.0017	0.02525	0.04299	0.33602	0.12594	3.36826	2.95982	0.63857	-0.1316	0.78464	0.13094
	$\beta = 0.25$	0.69497	0.24104	0.02538	0.04665	0.22867	-0.1459	2.97425	3.17725	0.62643	0.06239	0.77675	0.40479
	$\beta = 0.5$	0.69434	0.49150	0.02610	0.04185	0.07106	-0.0683	3.13936	2.81487	0.63421	0.36364	0.78663	0.59944
	$\beta = 0.75$	0.69614	0.73893	0.02485	0.03611	0.13553	-0.0905	3.02754	3.12066	0.64368	0.61574	0.78098	0.84255
$\alpha = 1.1$	$\beta = 0$	1.09009	-0.0017	0.04792	0.05536	0.07621	0.09207	3.08678	3.05441	0.95731	-0.1698	1.29828	0.16084
	$\beta = 0.25$	1.08352	0.24102	0.04470	0.05768	0.06737	0.06999	2.98146	2.99368	0.96108	0.07511	1.22325	0.42096
	$\beta = 0.5$	1.08703	0.49088	0.04577	0.05858	0.12693	0.01076	2.85982	2.84175	0.98156	0.32852	1.22310	0.66039
	$\beta = 0.75$	1.10803	0.74394	0.04975	0.44567	0.01838	-0.0079	3.23507	2.90922	0.96207	0.62497	1.25874	0.86740
$\alpha = 1.5$	$\beta = 0$	1.50170	-0.0021	0.05033	0.09211	0.04064	0.09035	2.93934	3.07248	1.37432	-0.2652	1.67563	0.32244
	$\beta = 0.25$	1.50130	0.25655	0.04557	0.10691	-0.0118	0.09777	3.34033	3.25015	1.39257	0.00467	1.59846	0.59641
	$\beta = 0.5$	1.50708	0.51479	0.04110	0.09308	-0.0590	0.06233	3.25741	3.42154	1.38046	0.22334	1.64909	0.88257
	$\beta = 0.75$	1.53243	0.79842	0.04467	0.08517	0.05734	-0.0475	3.13595	2.91921	1.40657	0.55105	1.70385	1
$\alpha = 1.9$	$\beta = 0$	1.87921	0.04882	0.03315	0.33591	-0.0274	-0.0279	2.69800	3.10226	1.77821	-1	2	1
	$\beta = 0.25$	1.92028	0.24223	0.01856	0.21779	-0.2136	0.25077	3.11422	2.87579	1.85782	-0.5600	2	1
	$\beta = 0.5$	1.90426	0.54333	0.03307	0.36026	-0.1035	-0.5706	2.91289	2.97721	1.87128	-0.3994	2	1
	$\beta = 0.75$	1.89619	0.74200	0.04439	0.13274	0.32779	-0.4841	3.34266	2.81680	1.78481	0.26729	2	1
		$\hat{\sigma}$	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\mu}$
$\alpha = 0.7$	$\beta = 0$	0.48263	-0.0120	0.03766	0.04614	0.34151	-0.1432	3.43387	3.23663	0.37455	-0.1746	0.62195	0.12574
	$\beta = 0.25$	0.49226	-0.0148	0.03858	0.05549	0.11130	-0.1233	2.85877	3.11552	0.38308	-0.2080	0.60631	0.15544
	$\beta = 0.5$	0.49301	-0.0227	0.03589	0.07385	0.08120	-0.3027	2.99377	3.03826	0.39986	-0.2983	0.63774	0.13673
	$\beta = 0.75$	0.49509	-0.0324	0.03078	0.08957	0.14562	-0.6298	3.00485	4.02902	0.41749	-0.4415	0.59792	0.14833
$\alpha = 1.1$	$\beta = 0$	0.49236	0.04282	0.02961	0.67272	0.04983	-12.466	3.35908	269.656	0.38048	-12.715	0.57896	5.74438
	$\beta = 0.25$	0.48469	0.12022	0.02796	2.83622	-0.0202	14.4712	2.90884	310.540	0.38609	-20.113	0.56873	56.1417
	$\beta = 0.5$	0.49932	-1.6968	0.02639	40.0957	0.18733	-22.171	2.78871	494.399	0.43056	-895.74	0.56708	23.8952
	$\beta = 0.75$	0.49403	-0.4700	0.02559	12.9944	0.11719	-15.637	2.80770	271.913	0.42889	-241.97	0.57679	45.2599
$\alpha = 1.5$	$\beta = 0$	0.49206	0.01898	0.01174	0.03762	-0.0086	0.09610	3.16859	3.15158	0.42454	-0.1105	0.55566	0.17740
	$\beta = 0.25$	0.50086	0.01512	0.02155	0.05323	0.08922	0.07767	3.29561	3.08291	0.44003	-0.1414	0.57723	0.17886
	$\beta = 0.5$	0.49766	0.03273	0.02173	0.05620	0.04091	0.04769	2.95778	3.37743	0.44048	-0.1352	0.57519	0.21874
	$\beta = 0.75$	0.49971	0.02951	0.02248	0.06141	0.00396	0.16876	2.98700	2.88719	0.42740	-0.1470	0.56912	0.21856
$\alpha = 1.9$	$\beta = 0$	0.48897	0.03911	0.01716	0.05099	-0.1066	0.10991	3.09820	3.01462	0.43645	-0.0713	0.53902	0.14810
	$\beta = 0.25$	0.49462	0.02141	0.01782	0.03811	0.01182	0.02014	2.68891	3.10787	0.44516	-0.0748	0.53877	0.09631
	$\beta = 0.5$	0.49915	0.00928	0.01623	0.04081	-0.0332	0.08530	3.20408	3.06938	0.45020	-0.1143	0.55800	0.20017
	$\beta = 0.75$	0.49294	0.02981	0.01956	0.04752	0.07312	0.31112	3.36650	3.25271	0.41250	-0.1028	0.55546	0.20135

See legend Table 3.

Table 6: Simulation Results Using Constrained Indirect Inference ($h = 2$)

α	β	Mean		Sd		Skw		Kur		Min		Max	
		$\hat{\alpha}$	$\hat{\beta}$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\alpha}$	$\hat{\beta}$
$\alpha = 0.7$	$\beta = 0$	0.69980	-0.0014	0.02266	0.04323	0.25316	0.18054	3.05471	2.97109	0.64673	-0.1379	0.77046	0.12585
	$\beta = 0.25$	0.69645	0.24468	0.02280	0.04600	0.10523	-0.1229	2.94400	3.18648	0.63182	0.04382	0.77290	0.39651
	$\beta = 0.5$	0.69545	0.49448	0.02555	0.03888	0.04777	0.05233	3.03555	2.74086	0.63410	0.38557	0.78069	0.61758
	$\beta = 0.75$	0.69796	0.74166	0.02401	0.03176	0.16932	-0.0118	2.90120	2.89013	0.64023	0.64374	0.78909	0.82604
$\alpha = 1.1$	$\beta = 0$	1.09487	-0.0016	0.04436	0.05521	0.08517	0.05544	2.98115	3.06311	0.98492	-0.2137	1.27277	0.19269
	$\beta = 0.25$	1.08969	0.24429	0.04161	0.05355	0.17913	0.16680	2.90983	3.06728	0.98875	0.09117	1.21269	0.40767
	$\beta = 0.5$	1.09152	0.49833	0.04869	0.05598	0.17414	0.07898	3.61143	2.94021	0.99571	0.33583	1.35919	0.64022
	$\beta = 0.75$	1.10867	0.74535	0.04318	0.04049	0.04545	0.02925	3.17190	2.92667	0.98222	0.63630	1.25351	0.86662
$\alpha = 1.5$	$\beta = 0$	1.51070	-0.0081	0.03936	0.08620	0.04593	0.14055	2.93026	3.09121	1.28452	-0.3073	1.64648	0.30770
	$\beta = 0.25$	1.50381	0.27923	0.07181	0.12075	0.22964	0.74800	2.93053	2.98429	1.34820	-0.1079	1.49842	0.76966
	$\beta = 0.5$	1.50106	0.50169	0.06184	0.10783	0.07514	0.30276	3.29082	2.95614	1.41426	0.26227	1.66708	0.90999
	$\beta = 0.75$	1.53467	0.74529	0.04551	0.08006	0.05924	-0.0454	3.39137	2.95405	1.43521	0.58915	1.64904	1
$\alpha = 1.9$	$\beta = 0$	1.88381	0.04368	0.03229	0.34884	-0.0133	-0.0585	2.62977	3.10054	1.79687	-1	2	1
	$\beta = 0.25$	1.92002	0.26274	0.01875	0.21549	-0.2060	0.21114	2.71186	2.87838	1.86035	-0.8084	2	1
	$\beta = 0.5$	1.89765	0.50982	0.03111	0.27608	-0.2467	-0.6353	2.61594	2.94058	1.81107	-0.6905	2	1
	$\beta = 0.75$	1.89985	0.74731	0.04294	0.12712	0.24978	-0.5013	3.22778	3.09165	1.79377	0.28295	2	1
		$\hat{\sigma}$	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\mu}$
$\alpha = 0.7$	$\beta = 0$	0.48846	-0.0072	0.03490	0.04770	0.39928	-0.1225	3.60897	3.23624	0.37173	-0.1589	0.61814	0.13201
	$\beta = 0.25$	0.49541	-0.0124	0.03408	0.05269	0.04819	-0.1314	2.87677	3.24073	0.40574	-0.2082	0.60176	0.16016
	$\beta = 0.5$	0.49596	-0.0176	0.03225	0.07350	0.08467	-0.3816	2.90084	3.54838	0.41288	-0.3107	0.60122	0.16932
	$\beta = 0.75$	0.49592	-0.0254	0.02837	0.08367	0.18867	-0.6998	3.03114	3.66701	0.42635	-0.3938	0.61197	0.14884
$\alpha = 1.1$	$\beta = 0$	0.49742	0.04210	0.02609	0.29655	0.06462	-4.6667	2.84856	81.2470	0.40881	-4.0893	0.57421	1.97180
	$\beta = 0.25$	0.49069	-0.0224	0.02361	5.80801	0.06300	-19.093	2.87730	416.011	0.41479	-123.94	0.55301	28.0671
	$\beta = 0.5$	0.49275	-0.3983	0.02366	10.5808	0.02948	-20.478	2.80057	440.398	0.42774	-229.48	0.57487	11.7277
	$\beta = 0.75$	0.49578	0.01229	0.02213	3.55400	0.14009	-6.4207	2.98572	141.137	0.43959	-54.921	0.56370	31.9651
$\alpha = 1.5$	$\beta = 0$	0.49607	0.01620	0.02359	0.05177	0.02460	0.06796	3.05103	2.98950	0.43354	-0.1513	0.56961	0.19014
	$\beta = 0.25$	0.50478	0.00753	0.01905	0.04877	0.15889	0.20619	3.09613	3.18032	0.44668	-0.1621	0.56830	0.18598
	$\beta = 0.5$	0.49985	0.02153	0.01934	0.05233	0.03834	0.18295	3.19436	3.52115	0.42902	-0.1322	0.56326	0.24172
	$\beta = 0.75$	0.50328	0.02120	0.01875	0.05477	0.08697	0.17702	3.37295	3.47780	0.44813	-0.1716	0.56718	0.26079
$\alpha = 1.9$	$\beta = 0$	0.49172	0.03579	0.01645	0.04869	0.00233	0.10981	3.37304	3.55536	0.44442	-0.0822	0.54170	0.14187
	$\beta = 0.25$	0.49759	0.01454	0.01575	0.03263	-0.0968	0.01406	2.71510	3.09501	0.45654	-0.0838	0.53699	0.08601
	$\beta = 0.5$	0.49428	0.02060	0.01472	0.03903	-0.0011	0.13643	2.81550	3.22588	0.45273	-0.0785	0.53716	0.23055
	$\beta = 0.75$	0.49532	0.01984	0.01708	0.04227	0.07087	0.23456	3.08340	3.03199	0.44703	-0.0833	0.55988	0.18993

See legend Table 3.

Table 7: Simulation Results Using Constrained Indirect Inference ($h = 5$)

α	β	Mean		Sd		Skw		Kur		Min		Max	
		$\hat{\alpha}$	$\hat{\beta}$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\alpha}$	$\hat{\beta}$
$\alpha = 0.7$	$\beta = 0$	0.70071	0.00066	0.02238	0.04186	0.27335	0.10441	3.09070	2.91673	0.64529	-0.1283	0.76912	0.12737
	$\beta = 0.25$	0.69764	0.24539	0.02310	0.04434	0.18449	-0.1062	3.02296	3.17429	0.63630	0.05973	0.77945	0.38414
	$\beta = 0.5$	0.69705	0.49614	0.02455	0.03864	0.09840	0.02058	3.12645	2.91200	0.63883	0.39487	0.77498	0.60894
	$\beta = 0.75$	0.69905	0.74414	0.02359	0.03044	0.13177	-0.0994	2.92730	3.03739	0.64260	0.64114	0.76732	0.85088
$\alpha = 1.1$	$\beta = 0$	1.09939	0.00208	0.04171	0.05248	0.04876	0.00668	3.00714	3.08991	1.00337	-0.1941	1.26044	0.17982
	$\beta = 0.25$	1.09147	0.24970	0.03796	0.05225	0.08205	0.07372	2.88214	2.92212	1.00259	0.10405	1.20477	0.40128
	$\beta = 0.5$	1.09486	0.49979	0.03981	0.05115	0.09604	0.01681	3.00581	3.07064	1.00251	0.32725	1.24811	0.63706
	$\beta = 0.75$	1.10876	0.75529	0.03878	0.04233	0.01686	-0.0191	3.05151	3.02877	1.00870	0.62158	1.22721	0.86645
$\alpha = 1.5$	$\beta = 0$	1.61636	-0.0007	0.15830	0.09258	-0.7881	0.05277	3.55923	2.92514	1.28452	-0.3408	1.77736	0.37286
	$\beta = 0.25$	1.50889	0.29539	0.08943	0.14234	0.07741	0.06266	2.95583	2.90733	1.40957	-0.0575	1.65926	0.48329
	$\beta = 0.5$	1.50573	0.50382	0.04990	0.11594	0.06101	0.00962	3.04216	3.10867	1.33055	0.25042	1.63743	0.85000
	$\beta = 0.75$	1.53610	0.75847	0.04100	0.07680	0.23251	-0.1363	3.03184	2.94886	1.43634	0.60247	1.66595	1
$\alpha = 1.9$	$\beta = 0$	1.91296	0.04446	0.02901	0.38864	-0.0339	-0.0382	2.73966	3.14470	1.81998	-1	2	1
	$\beta = 0.25$	1.90222	0.25615	0.03032	0.19695	-0.1984	-0.1696	2.80229	3.14547	1.81613	-0.9168	2	1
	$\beta = 0.5$	1.87499	0.49797	0.03176	0.28403	-0.1910	-0.6242	2.82220	3.28527	1.77608	-0.5831	2	1
	$\beta = 0.75$	1.90291	0.75551	0.04266	0.12304	0.20058	-0.5419	3.04685	3.17162	1.80491	0.31181	2	1
		$\hat{\sigma}$	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\mu}$
$\alpha = 0.7$	$\beta = 0$	0.49925	-0.0063	0.03269	0.04437	0.23332	-0.2048	3.08615	3.08155	0.41658	-0.1586	0.59932	0.10343
	$\beta = 0.25$	0.49797	-0.0074	0.03111	0.05014	0.02654	-0.1057	2.92822	3.14844	0.41502	-0.1645	0.60212	0.16391
	$\beta = 0.5$	0.49854	-0.0114	0.02919	0.06721	0.01907	-0.3664	2.85359	3.19597	0.42312	-0.3460	0.60783	0.14463
	$\beta = 0.75$	0.49790	-0.0182	0.02615	0.07905	0.18026	-0.6417	2.95854	3.47849	0.43240	-0.3239	0.57920	0.14245
$\alpha = 1.1$	$\beta = 0$	0.49216	0.02163	0.02411	1.26213	0.05333	-17.301	3.13336	380.187	0.43295	-26.244	0.57332	8.94518
	$\beta = 0.25$	0.49333	0.03452	0.02281	1.89113	-0.0096	-14.394	2.81751	275.141	0.42893	-36.175	0.55469	10.1563
	$\beta = 0.5$	0.49521	0.23407	0.02195	1.87679	0.13469	13.2331	2.70495	234.624	0.43274	-0.9963	0.56313	34.8600
	$\beta = 0.75$	0.50104	0.67595	0.02403	5.76526	0.17314	18.9709	2.82491	398.196	0.43905	-1.3514	0.56953	122.435
$\alpha = 1.5$	$\beta = 0$	0.49670	0.08436	0.02554	0.03302	-0.0709	0.02734	3.19652	3.06548	0.40846	-0.0843	0.57379	0.13049
	$\beta = 0.25$	0.50705	0.00359	0.01878	0.04511	0.03505	0.18700	2.93850	2.74076	0.45623	-0.1268	0.56640	0.13473
	$\beta = 0.5$	0.50236	0.01794	0.01836	0.05027	-0.0438	0.12913	3.04681	3.22343	0.44177	-0.1305	0.55049	0.17614
	$\beta = 0.75$	0.50497	0.01451	0.01821	0.04897	0.04005	0.11516	2.90116	4.01455	0.44330	-0.1699	0.55090	0.22975
$\alpha = 1.9$	$\beta = 0$	0.49736	0.00215	0.01354	0.03779	0.00819	0.08750	2.88224	3.31736	0.46009	-0.0670	0.53931	0.07439
	$\beta = 0.25$	0.49771	0.03975	0.01464	0.05196	0.01784	0.00851	2.94627	3.06366	0.45508	-0.1471	0.53735	0.18762
	$\beta = 0.5$	0.49452	0.01273	0.01321	0.03372	-0.0367	0.05731	3.12125	2.97261	0.45160	-0.0837	0.53472	0.18361
	$\beta = 0.75$	0.49779	0.01508	0.01561	0.03950	0.18322	0.30731	3.38417	3.02739	0.44790	-0.0836	0.55220	0.17931

See legend Table 3.

Table 8: Simulation Results Using CGMM

α	β	Mean		Sd		Skw		Kur		Min		Max	
		$\hat{\alpha}$	$\hat{\beta}$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\alpha}$	$\hat{\beta}$
$\alpha = 0.7$	$\beta = 0$	0.70246	0.00453	0.05523	0.09917	0.10428	0.02051	2.92881	2.98363	0.54271	-0.3031	0.92322	0.33654
	$\beta = 0.25$	0.69942	0.25274	0.05449	0.09799	-0.0287	0.12488	2.94741	3.04557	0.53148	-0.0197	0.86657	0.65019
	$\beta = 0.5$	0.70403	0.49871	0.05703	0.09563	0.20096	-0.0653	2.97157	3.22779	0.51523	0.17162	0.89751	0.82418
	$\beta = 0.75$	0.70149	0.75205	0.05750	0.08797	0.11734	0.06381	3.25004	3.03277	0.52674	0.44702	0.95060	1
$\alpha = 1.1$	$\beta = 0$	1.10646	-0.0018	0.06514	0.10478	0.08502	0.05904	2.92651	3.09517	0.88966	-0.3445	1.34268	0.33340
	$\beta = 0.25$	1.10748	0.24723	0.06475	0.10334	0.45218	-0.0899	2.81269	3.00468	1.00015	-0.0744	1.34398	0.59337
	$\beta = 0.5$	1.10532	0.50653	0.06172	0.10180	0.46986	-0.1326	2.65589	3.07094	1.01123	0.12173	1.33320	0.80897
	$\beta = 0.75$	1.10788	0.75642	0.06570	0.08800	0.48686	-0.1229	2.58816	2.77712	1.01350	0.47181	1.33662	1
$\alpha = 1.5$	$\beta = 0$	1.50394	0.00300	0.07491	0.15502	0.07944	-0.1021	2.90075	2.94933	1.27609	-0.4999	1.73336	0.49171
	$\beta = 0.25$	1.50259	0.25017	0.07660	0.14791	0.10920	-0.0023	3.05500	2.82810	1.28174	-0.4995	1.80542	0.72411
	$\beta = 0.5$	1.50042	0.50436	0.07654	0.14268	0.15854	-0.1292	3.21382	3.01327	1.22978	0.04270	1.78497	0.94274
	$\beta = 0.75$	1.50143	0.75472	0.07766	0.12883	0.02569	-0.1980	2.71052	2.83864	1.27747	0.32673	1.73554	1
$\alpha = 1.9$	$\beta = 0$	1.89579	0.00267	0.05785	0.61545	-0.4663	0.01405	3.17735	2.08033	1.68857	-1	2	1
	$\beta = 0.25$	1.89617	0.24182	0.05614	0.59452	-0.3373	-0.4358	2.56499	2.36644	1.72319	-1	2	1
	$\beta = 0.5$	1.88859	0.43803	0.05572	0.52585	-0.3873	-0.8827	2.84725	3.22289	1.70699	-1	2	1
	$\beta = 0.75$	1.88562	0.63096	0.05408	0.44541	-0.3426	-1.5221	2.89383	5.37557	1.70502	-1	2	1
		$\hat{\sigma}$	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\mu}$
$\alpha = 0.7$	$\beta = 0$	0.61677	-0.0049	0.04339	0.12403	0.08100	0.05620	2.99654	3.48691	0.48646	-0.4386	0.77329	0.49165
	$\beta = 0.25$	0.61840	-0.0141	0.04315	0.13209	0.19957	-0.5865	3.31673	3.99939	0.49262	-0.6136	0.78981	0.33557
	$\beta = 0.5$	0.61257	-0.0330	0.04207	0.20235	0.07586	-1.7218	3.01417	11.7298	0.47571	-1.9383	0.75397	0.38686
	$\beta = 0.75$	0.61257	-0.0448	0.04204	0.31682	0.16892	-8.1163	2.88960	14.9402	0.50101	-6.2499	0.75397	0.47161
$\alpha = 1.1$	$\beta = 0$	0.46443	-0.0076	0.03417	0.87989	-0.0054	0.73063	2.74153	21.4386	0.37091	-6.3376	0.56670	6.29337
	$\beta = 0.25$	0.46730	0.59779	0.03452	1.65256	0.20921	2.30911	3.05457	7.81348	0.36158	-0.9676	0.59808	8.69041
	$\beta = 0.5$	0.46805	0.98966	0.03396	2.40559	0.13749	1.81419	3.11319	5.28035	0.35914	-1.3099	0.58887	9.91476
	$\beta = 0.75$	0.46740	1.40612	0.03320	3.30409	0.16813	1.67018	2.83572	5.13296	0.37911	-1.6046	0.58193	17.5146
$\alpha = 1.5$	$\beta = 0$	0.35310	0.00097	0.02031	0.06647	0.14148	-0.0090	2.95166	3.18155	0.27786	-0.2332	0.45002	0.23373
	$\beta = 0.25$	0.35286	-0.0005	0.02553	0.06647	0.12364	0.34933	2.75394	3.63643	0.28308	-0.1918	0.42999	0.32893
	$\beta = 0.5$	0.35203	0.00448	0.02535	0.07431	0.04423	0.45102	2.83478	3.28991	0.28048	-0.2005	0.43532	0.26673
	$\beta = 0.75$	0.35191	0.00644	0.02526	0.08696	0.39504	0.65770	3.55127	3.84472	0.28365	-0.1995	0.46983	0.39964
$\alpha = 1.9$	$\beta = 0$	0.26834	-0.0011	0.01801	0.03648	-0.0139	0.04827	2.87775	3.04473	0.21239	-0.1046	0.31949	0.12637
	$\beta = 0.25$	0.26847	0.00057	0.01802	0.03717	0.05325	0.05794	3.04424	3.03652	0.21285	-0.1281	0.33277	0.12620
	$\beta = 0.5$	0.26671	-0.0010	0.01774	0.03673	0.17326	0.00541	3.19706	2.99391	0.21670	-0.1152	0.32882	0.10988
	$\beta = 0.75$	0.26778	-0.0014	0.01785	0.03643	0.28610	0.02897	3.08036	2.99949	0.21947	-0.1209	0.33558	0.10727

See legend Table 3.

Table 9: Simulation Results Using Empirical Quantiles

α	β	Mean		Sd		Skw		Kur		Min		Max	
		$\hat{\alpha}$	$\hat{\beta}$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\alpha}$	$\hat{\beta}$
$\alpha = 0.7$	$\beta = 0$	0.69739	-0.0066	0.05014	0.12590	0.21597	0.00495	3.24569	2.81922	0.55716	-0.4457	0.89282	0.38660
	$\beta = 0.25$	0.70279	0.23858	0.05503	0.11520	0.04247	-0.3194	3.01759	2.82425	0.51594	-0.1346	0.87992	0.53750
	$\beta = 0.5$	0.70466	0.50739	0.06440	0.09177	0.19822	-0.3600	3.30088	3.40595	0.52512	0.18476	0.98048	0.81046
	$\beta = 0.75$	0.70135	0.80729	0.06764	0.08998	0.26223	0.27189	3.17566	2.55525	0.50711	0.49054	0.99332	1
$\alpha = 1.1$	$\beta = 0$	1.09615	0.00356	0.07207	0.12800	0.04143	-0.0375	3.21410	2.78511	0.84297	-0.3875	1.32831	0.39826
	$\beta = 0.25$	1.09813	0.24569	0.07408	0.11230	0.12454	-0.2572	3.24376	2.87124	0.85374	-0.1545	1.36240	0.54963
	$\beta = 0.5$	1.10362	0.50553	0.08408	0.09280	0.10291	-0.1444	2.80454	3.12140	0.85475	0.21036	1.40068	0.80700
	$\beta = 0.75$	1.11116	0.79433	0.09764	0.10293	0.30719	0.48604	3.04301	2.40930	0.83550	0.55593	1.44142	1
$\alpha = 1.5$	$\beta = 0$	1.49728	0.00052	0.08622	0.15408	0.26334	0.03276	2.95087	3.47837	1.27739	-0.6284	1.83906	0.71356
	$\beta = 0.25$	1.50519	0.26462	0.09771	0.16563	0.55144	0.26814	3.46410	4.14887	1.27023	-0.3556	1.96258	1
	$\beta = 0.5$	1.51320	0.53616	0.10665	0.18244	0.60535	0.60485	3.32202	2.39938	1.21266	0	2	1
	$\beta = 0.75$	1.51366	0.78814	0.11552	0.16116	0.38357	-0.2304	3.01588	3.98180	1.22153	0.00023	2	1
$\alpha = 1.9$	$\beta = 0$	1.88483	0.00718	0.10982	0.39920	-0.5890	0.03027	2.32392	4.15334	1.50585	-1	2	1
	$\beta = 0.25$	1.88017	0.11624	0.10815	0.38410	-0.4547	0.20840	2.05118	4.03050	1.56879	-1	2	1
	$\beta = 0.5$	1.88602	0.21183	0.11228	0.40663	-0.6808	0.25370	2.53809	3.32461	1.53240	-1	2	1
	$\beta = 0.75$	1.88500	0.35634	0.11022	0.40777	-0.5977	0.38863	2.30347	2.01720	1.54524	-1	2	1
		$\hat{\sigma}$	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\mu}$
$\alpha = 0.7$	$\beta = 0$	0.48011	0.00356	0.05505	0.13201	0.34061	-0.2475	3.30803	3.27211	0.30115	-0.6036	0.67273	0.35201
	$\beta = 0.25$	0.48801	0.01059	0.06619	0.13941	0.07724	-0.7199	2.92247	7.55662	0.28803	-1.0916	0.70233	0.45907
	$\beta = 0.5$	0.49314	-0.0595	0.08086	0.35424	0.27471	-7.5434	2.81885	98.1057	0.28850	-5.2759	0.77788	0.40285
	$\beta = 0.75$	0.47942	-0.2328	0.07560	2.15322	0.16892	-29.007	3.12168	890.343	0.26879	-66.374	0.72052	0.58795
$\alpha = 1.1$	$\beta = 0$	0.49281	0.15793	0.03486	4.37611	0.04542	19.1429	2.92317	469.885	0.37329	-15.744	0.60093	113.239
	$\beta = 0.25$	0.49536	-1.7429	0.03676	42.4589	0.10256	-19.611	3.05209	442.188	0.38885	-1043.5	0.62406	210.590
	$\beta = 0.5$	0.49588	-0.2179	0.04133	17.1100	0.12611	-0.2654	2.99236	101.934	0.36434	-251.04	0.63187	-915.16
	$\beta = 0.75$	0.49300	18.5730	0.04100	516.121	0.11450	28.3667	3.19387	846.183	0.35929	-915.16	0.64711	15649.3
$\alpha = 1.5$	$\beta = 0$	0.49717	0.00085	0.02818	0.07124	0.09323	-0.2927	2.91915	3.59259	0.41958	-0.3195	0.62924	0.22724
	$\beta = 0.25$	0.49953	0.00462	0.02955	0.07139	0.08656	0.26372	3.03492	3.14169	0.40454	-0.1864	0.60016	0.25175
	$\beta = 0.5$	0.49799	0.00543	0.02920	0.08479	0.19406	0.36305	3.26891	3.57071	0.41503	-0.2878	0.60630	0.32177
	$\beta = 0.75$	0.49762	-0.0007	0.02945	0.10540	0.10433	0.27163	3.01588	3.98180	0.40026	-0.3473	0.59542	0.47220
$\alpha = 1.9$	$\beta = 0$	0.49895	0.00022	0.02712	0.04059	0.01713	-0.0197	2.92259	2.99586	0.40682	-0.1325	0.57580	0.13591
	$\beta = 0.25$	0.49799	-0.0038	0.02710	0.04005	0.03818	-0.0030	2.89880	3.09531	0.41499	-0.1322	0.59038	0.12151
	$\beta = 0.5$	0.50022	-0.0081	0.02831	0.04269	0.13708	0.13264	2.93005	3.13437	0.42373	-0.1454	0.58900	0.13618
	$\beta = 0.75$	0.49991	-0.0151	0.03011	0.04403	0.16740	-0.0957	3.09234	3.04984	0.41636	-0.1545	0.63315	0.12255

See legend Table 3.

Table 10: Comparison of Finite Sample Indirect Inference Standard Deviations (Sd. Ind. Inf.) with Asymptotic Deviations Corresponding to the Cramér-Rao bounds (Sd. Asympt.) for α and β

α	β	Sd. Ind. Inf.		Sd. Asympt.	
		$\hat{\alpha}$	$\hat{\beta}$	$\hat{\alpha}$	$\hat{\beta}$
1	0	0.049	0.058	0.039	0.055
	0.5	0.045	0.058	0.015	0.053
1.3	0	0.072	0.076	0.045	0.075
	0.5	0.077	0.072	0.044	0.066
1.5	0	0.048	0.096	0.049	0.096
	0.5	0.044	0.097	0.047	0.085
1.9	0	0.026	0.369	0.036	2.287
	0.5	0.015	0.364	0.035	0.268

Table 11: Empirical Standard deviation and Kurtosis of the simulated GARCH processes

GARCH model	distn	1		5		10		20	
		Sd	Kur	Sd	Kur	Sd	Kur	Sd	Kur
$\delta_1 = 0.1$, $\delta_2 = 0.4$	Normal	0.4481	3.0830	1.8712	3.966	3.2154	4.1601	4.3465	4.0714
	t_5	0.4490	7.3609	1.6471	7.8251	2.7954	6.8723	3.8312	6.2914
$\delta_1 = 0.2$, $\delta_2 = 0.4$	Normal	0.5001	3.3841	1.6251	5.6451	2.7051	5.8799	3.5962	5.6221
	t_5	0.5033	9.2001	1.3502	10.628	2.2098	9.5810	2.9969	8.7135
$\delta_1 = 0.1$, $\delta_2 = 0.8$	Normal	0.7238	3.1852	1.9432	3.9021	3.2094	4.049	4.3450	3.9800
	t_5	0.7251	7.6987	1.7654	7.2485	2.8451	6.5741	3.8990	6.0552
$\delta_1 = 0.2$, $\delta_2 = 0.7$	Normal	0.7495	3.9001	1.745	5.5252	2.7623	5.6610	3.6623	5.4392
	t_5	0.7501	10.211	1.5258	9.9633	2.3284	9.1366	3.3132	8.3241
$\delta_1 = 0.05$, $\delta_2 = 0.9$	Normal	1.4032	3.1222	3.5340	3.6051	5.7514	3.7412	7.7562	3.8254
	t_5	1.4111	7.5611	3.4444	6.8010	5.5674	5.9662	7.5499	5.4552
$\delta_1 = 0.05$, $\delta_2 = 0.95$	Normal	3.7330	3.7512	6.0751	3.9102	8.0281	3.8251	10.127	3.7308
	t_5	3.5912	8.0741	5.8332	6.4541	7.7271	5.7799	9.8432	5.3432
$\delta_1 = 0.1$, $\delta_2 = 0.9$	Normal	2.1616	4.1857	3.5131	4.3551	4.6100	4.2077	5.7754	4.0344
	t_5	1.8612	8.9587	3.0214	7.3891	4.0271	6.6388	5.1633	6.1037
$\delta_1 = 0.2$, $\delta_2 = 0.8$	Normal	1.7927	6.2410	2.8901	6.3082	3.7521	5.9133	4.6781	5.4741
	t_5	1.4241	12.233	2.3091	10.413	3.0714	9.2874	3.9371	8.3142

Mean, standard deviation (Sd) and kurtosis (Kur) of the 500 simulated Gaussian and Student-t_5 GARCH processes for all the frequencies.

Table 12: Simulation Results when the DGP is a GARCH model. Results for $\hat{\alpha}$

GARCH model	distn	1	5	10	20
$\delta_1 = 0.1$, $\delta_2 = 0.4$	Normal	1.9132 [0.049]	1.8314 [0.0772]	1.9560 [0.0341]	1.9014 [0.0958]
	t_5	1.7449 [0.0382]	1.8320 [0.0689]	1.9001 [0.0760]	1.8529 [0.1281]
$\delta_1 = 0.2$, $\delta_2 = 0.4$	Normal	1.9040 [0.0316]	1.8641 [0.0689]	1.8153 [0.0813]	1.8987 [0.0775]
	t_5	1.7172 [0.0502]	1.7329 [0.0747]	1.9013 [0.0503]	1.7331 [0.1365]
$\delta_1 = 0.1$, $\delta_2 = 0.8$	Normal	1.9132 [0.0348]	1.9256 [0.0424]	1.9168 [0.0351]	1.9072 [0.1095]
	t_5	1.7277 [0.0418]	1.8423 [0.0718]	1.8947 [0.0787]	1.7929 [0.0807]
$\delta_1 = 0.2$, $\delta_2 = 0.7$	Normal	1.8395 [0.0419]	1.8471 [0.0754]	1.7186 [0.1128]	1.9011 [0.0535]
	t_5	1.6539 [0.0637]	1.6109 [0.0929]	1.6476 [0.1154]	1.5794 [0.1754]
$\delta_1 = 0.05$, $\delta_2 = 0.9$	Normal	1.9361 [0.0248]	1.9136 [0.0538]	1.8537 [0.0714]	1.7992 [0.1112]
	t_5	1.7459 [0.0382]	1.7785 [0.0652]	1.7599 [0.0962]	1.8071 [0.1333]
$\delta_1 = 0.05$, $\delta_2 = 0.95$	Normal	1.7526 [0.0992]	1.6941 [0.1102]	1.6891 [0.0962]	1.5987 [0.1983]
	t_5	1.5882 [0.1260]	1.6085 [0.1721]	1.5014 [0.1777]	1.5942 [0.1630]
$\delta_1 = 0.1$, $\delta_2 = 0.9$	Normal	1.6803 [0.1950]	1.6047 [0.1658]	1.6854 [0.1432]	1.5401 [0.1631]
	t_5	1.5443 [0.1437]	1.5671 [0.1270]	1.4482 [0.1183]	1.4260 [0.2244]
$\delta_1 = 0.2$, $\delta_2 = 0.8$	Normal	1.6992 [0.0569]	1.5760 [0.1604]	1.4951 [0.1644]	1.7519 [0.1896]
	t_5	1.4423 [0.1607]	1.4827 [0.1407]	1.3012 [0.2081]	1.4782 [0.2241]

Mean and standard deviation of 500 estimated α's when the DGP is a 1000 observations GARCH model with Gaussian or Student-t_5 density and for different aggregation frequencies.

Table 13: Simulation Results when the DGP is a GARCH model. Results for $\hat{\sigma}$

GARCH model	distn	1	5	10	20
$\delta_1 = 0.1,\ \delta_2 = 0.4$	Normal	0.3073 [0.0135]	0.6579 [0.0482]	0.9487 [0.0878]	1.3024 [0.1784]
	t_5	0.2602 [0.0083]	0.6249 [0.0466]	0.8924 [0.0837]	1.3362 [0.1964]
$\delta_1 = 0.2,\ \delta_2 = 0.4$	Normal	0.3374 [0.0129]	0.7223 [0.0503]	1.0351 [0.1027]	1.4718 [0.2083]
	t_5	0.2825 [0.0106]	0.6574 [0.0537]	1.0179 [0.0867]	1.4288 [0.2045]
$\delta_1 = 0.1,\ \delta_2 = 0.8$	Normal	0.6800 [0.0311]	1.4731 [0.1003]	2.0944 [0.1943]	2.8884 [0.4403]
	t_5	0.5618 [0.0269]	1.3437 [0.1136]	2.0097 [0.1758]	2.9488 [0.4071]
$\delta_1 = 0.2,\ \delta_2 = 0.7$	Normal	0.6369 [0.0389]	1.3575 [0.1112]	1.9211 [0.2127]	2.8364 [0.3744]
	t_5	0.5106 [0.0278]	1.1527 [0.119]	1.7455 [0.1803]	2.4972 [0.3994]
$\delta_1 = 0.05,\ \delta_2 = 0.9$	Normal	0.9621 [0.0427]	2.1098 [0.1439]	2.9741 [0.3041]	4.0430 [0.5512]
	t_5	0.8079 [0.0381]	1.9341 [0.1744]	2.8520 [0.2510]	3.9414 [0.6127]
$\delta_1 = 0.05,\ \delta_2 = 0.95$	Normal	3.5478 [1.3464]	7.7671 [2.9350]	10.7642 [2.9182]	14.790 [3.9942]
	t_5	2.6441 [0.6771]	6.0090 [1.7830]	8.9185 [2.3221]	13.028 [3.6017]
$\delta_1 = 0.1,\ \delta_2 = 0.9$	Normal	2.5326 [1.2519]	5.3294 [3.1886]	7.6305 [2.2275]	10.4077 [3.5646]
	t_5	1.5962 [0.3591]	4.0038 [0.8087]	5.4858 [1.2710]	7.8055 [2.0849]
$\delta_1 = 0.2,\ \delta_2 = 0.8$	Normal	1.4910 [0.3524]	3.1236 [0.8672]	4.3848 [0.9342]	6.6421 [1.7574]
	t_5	0.9867 [0.1240]	2.2070 [0.3454]	3.1700 [0.5084]	4.6427 [1.0487]

Mean and standard deviation of 500 estimated σ's when the DGP is a 1000 observations GARCH model with Gaussian or Student-t_5 density and for different aggregation frequencies.

Table 14: Descriptive Statistics

	SPO	SPC	HFRO	HFRO-SP	HFRC	HFRC-SP	TREO	TREO-SP	TREC	TREC-SP
Mean	0.8035	0.1806	0.9965	1.0744	0.3706	0.3921	1.0727	1.1577	0.5704	0.6853
Sd	1.1914	1.3283	1.7378	1.7037	2.0015	1.8301	1.6067	1.6461	1.9063	1.8829
Skw	-0.070	-0.742	0.1972	0.4806	-0.376	0.1445	0.6159	0.6277	0.6123	1.0078
Kur	3.6345	4.6757	4.2321	3.6157	5.2811	3.8936	4.5247	4.5089	5.5162	6.2887
Min	-3.373	-4.798	-5.010	-3.595	-7.874	-5.268	-3.249	-3.068	-4.664	-3.441
Max	3.6055	3.2017	6.3352	6.1364	6.1410	5.9667	6.9508	7.4271	7.8880	8.5677
LM ARCH	2.9621	10.801	5.7598	8.7414	1.5487	9.1414	10.074	9.7865	14.621	13.587
Hill (left tail)	1.5187	2.2412	1.7100	1.5307	2.7721	2.3878	1.3521	1.4921	1.7382	2.7790
Hill (right tail)	5.0001	5.7074	3.6464	3.2327	2.8630	2.7298	3.7644	4.2715	2.1851	2.1971

Descriptive Statistics of the hedge fund returns. LM ARCH stands for the Lagrange multiplier test for the presence of autocorrelation (up to order four) in the second moment. The test is based in the R^2. To compare with a $\chi^2_4 = 11.1$. Hill (left tail) and Hill (right tail stand for the Hill estimators for the left and right tail respectively.

Table 15: Estimation Results

	SPO	SPC	HFRO	HFRO-SP	HFRC	HFRC-SP	TREO	TREO-SP	TREC	TREC-SP
Constrained Indirect Inference										
$\hat{\alpha}$	1.9725	1.7858	1.9872	1.9132	1.8120	1.8980	1.9197	1.9502	1.6972	1.6712
	[0.0548]	[0.0417]	[0.0987]	[0.0397]	[0.0002]	[0.0112]	[0.0529]	[0.0468]	[0.0087]	[0.0136]
$\hat{\beta}$	0.9635	−0.322	0.8795	0.9501	−0.534	0.9201	0.9321	0.9235	−0.009	0.0345
	[0.4872]	[0.0677]	[0.7154]	[0.2141]	[0.0934]	[0.0693]	[0.2455]	[0.4354]	[0.0426]	[0.0483]
$\hat{\sigma}$	0.7199	0.7659	1.1515	1.1143	1.2080	1.1712	0.9779	0.9850	1.1237	0.9482
	[0.0669]	[0.0362]	[0.1158]	[0.0907]	[0.0477]	[0.0335]	[0.0217]	[0.0610]	[0.0273]	[0.0310]
$\hat{\mu}$	0.9401	0.1412	1.3656	1.6069	0.6197	0.7158	1.2719	1.1263	0.6580	0.6735
	[0.1486]	[0.0595]	[0.1157]	[0.2109]	[0.1408]	[0.0911]	[0.1004]	[0.2554]	[0.0532]	[0.0544]
Empirical Quantiles										
$\hat{\alpha}$	2.0000	1.7756	2.0000	1.9169	1.7965	1.8895	1.9491	1.9943	1.6971	1.6340
	[0.0000]	[0.0228]	[0.0000]	[0.1081]	[0.0790]	[0.0482]	[0.0412]	[0.0104]	[0.0257]	[0.0521]
$\hat{\beta}$	0.0000	−0.755	0.0000	1.0000	0.3214	1.0000	1.0000	0.5500	−0.125	−0.210
	[0.0000]	[0.0654]	[0.0000]	[0.0918]	[0.3379]	[0.2827]	[0.5028]	[0.5104]	[0.0427]	[0.1592]
$\hat{\sigma}$	0.8301	0.7731	1.2266	1.1460	1.2126	1.1883	1.0467	1.1263	1.0648	1.0588
	[0.0227]	[0.0130]	[1.2429]	[0.0572]	[0.0427]	[0.0218]	[0.0454]	[0.0228]	[0.0149]	[0.0173]
$\hat{\mu}$	0.6102	0.1139	0.7790	0.8900	0.4485	0.4045	0.9551	1.0459	0.4556	0.5150
	[0.3587]	[0.0180]	[0.7792]	[0.1146]	[0.0907]	[0.0736]	[0.0482]	[0.0582]	[0.0098]	[0.0501]
CGMM										
$\hat{\alpha}$	1.0230	1.6977	0.9899	1.6709	1.8697	1.8610	0.9903	1.7872	1.0003	0.9854
	[0.3361]	[0.3755]	[0.2633]	[0.3695]	[0.3533]	[0.2088]	[0.3722]	[0.0242]	[0.3309]	[0.3684]
$\hat{\beta}$	1.0000	−0.322	0.4264	1.0000	−0.609	0.9360	0.3541	1.0000	0.0033	−1.000
	[0.4833]	[0.5251]	[0.4611]	[0.4182]	[0.6169]	[0.0895]	[0.4589]	[0.0307]	[0.5015]	[0.8035]
$\hat{\sigma}$	0.8358	0.6829	0.9961	1.1276	1.4982	1.3817	0.8474	1.0703	0.9708	1.1654
	[0.0546]	[1.8019]	[1.6522]	[0.0939]	[12.378]	[0.0534]	[1.7274]	[0.0643]	[0.1653]	[0.0622]
$\hat{\mu}$	52.729	0.1925	−26.09	1.3385	0.2730	0.5027	−18.82	1.2635	7.5007	22.441
	[17.312]	[26.297]	[18.754]	[10.983]	[35.235]	[3.3912]	[15.966]	[0.0508]	[12.473]	[19.420]
Skewed-t										
$\hat{\nu}$	9.5524	6.6707	6.7279	9.7960	6.0731	8.0049	5.4009	7.2914	3.7503	4.0305
	[2.1581]	[1.1524]	[1.2354]	[2.2501]	[1.0924]	[1.7549]	[0.6487]	[1.6203]	[0.6801]	[1.4998]
$\hat{\gamma}$	1.1818	0.8515	1.2295	1.3134	0.9893	1.1109	1.2846	1.2923	1.0151	1.0342
	[0.1541]	[0.1393]	[0.2019]	[0.1749]	[0.1379]	[0.1915]	[0.2427]	[0.2257]	[0.1438]	[0.1410]
$\hat{\lambda}$	1.0393	1.0792	1.4123	1.4342	1.6261	1.5703	1.2250	1.3295	1.3473	1.3489
	[0.1299]	[0.1294]	[0.1533]	[0.1710]	[0.1755]	[0.1756]	[0.1702]	[0.1650]	[0.1607]	[0.1674]
$\hat{\omega}$	0.5202	0.5149	0.4831	0.3993	0.4152	0.1054	0.5018	0.5419	0.4863	0.5318
	[0.2259]	[0.3202]	[0.4055]	[0.3119]	[0.3997]	[0.4631]	[0.4015]	[0.3862]	[0.3349]	[0.3333]

Entries are estimated parameters for Constrained Indirect Inference, Empirical Quantiles, Continuous GMM and Skewed-t density. In square brackets bootstrapped, 20 repetitions of sampling with replacement, standard errors.