University of Konstanz
Department of Economics

Indirect Estimation of α-Stable Garch Models

Giorgio Calzolari, Roxana Halbleib,
and Alessandro Parrini

Working Paper Series
2012-31

INDIRECT ESTIMATION OF α-STABLE GARCH MODELS*

Giorgio CALZOLARI[1], Roxana HALBLEIB[2], and Alessandro PARRINI[3]

November 23, 2012

Abstract

It is a well-known fact that financial returns exhibit conditional heteroscedasticity and fat tails. While the GARCH-type models are very popular in depicting the conditional heteroscedasticity, the α-stable distribution is a natural candidate for the conditional distribution of financial returns. The α-stable distribution is a generalization of the normal distribution and is described by four parameters, two of which deal with tail-thickness and asymmetry. However, practical implementation of α-stable distribution in finance applications has been limited by its estimation difficulties. In this paper, we propose an indirect approach of estimating GARCH models with α-stable innovations by using as auxiliary models GARCH-type models with Student's t distributed innovations. We provide comprehensive empirical evidence on the performance of the method within a series of Monte Carlo simulation studies and an empirical application to financial returns.

Keywords: Indirect Inference, α-stable Distribution, GARCH Models, Student's t Distribution.

*Suggestions from Marco J. Lombardi, of the European Central Bank, are gratefully acknowledged.
[1]Università di Firenze, Italy; email:calzolar@ds.unifi.it
[2]Universität Konstanz, Germany; email: roxana.halbleib@uni-konstanz.de.
[3]Vrije Universiteit Amsterdam, The Netherlands; email: parrini.ale@gmail.com

1 Introduction

Most of the financial returns exhibit conditional heteroskedasticity and heavy-tailedness. While the conditional heteroskedasticity is standardly captured by means of GARCH or stochastic volatility (SV) models (e.g. Bollerslev (1986), Ghysels et al. (1996)), depicting the empirically observed fat-thickeness of financial returns is not always straightforward. Although theoretically most of the GARCH and SV specifications can accommodate for fat-tailedness through their specification, in practice, in most of the cases, there is still excess kurtosis left in the standardized residuals. A very common solution to this problem is to assume a fat-tailed distribution for the standardized innovations of the GARCH models, and the Student's t is a natural candidate (e.g., Calzolari et al. (2003)). However, one drawback of the Student's t distribution is that it lacks in stability under aggregation, which is of particular importance in portfolio applications and risk management. A fat-tailed distribution that overcomes the drawbacks of the Student's t is the α-stable. Its theoretical foundations lay on the generalized central limit theorem. Moreover, similar to the Student's t distribution, the α-stable can be easily adapted to account for asymmetry in the underlying series. The main drawback of this specification is its estimation. The fact that, for most of the parameters constellations, the α-stable does not have a closed-form density specification or the theoretical moments simply do not exist, makes the estimation of its parameters a cumbersome task and limits the interest among academics and practitioners.

In this paper we show how GARCH models with α-stable innovations can be estimated by using the indirect inference (IndInf) method proposed by Gouriéroux et al. (1993). This estimation approach has already proved its adequacy in estimating the parameters of the stable distribution in Lombardi and Calzolari (2008), Lombardi and Calzolari (2009) and Garcia et al. (2011). In the GARCH context, the α-stable distribution is first mentioned by de Vries (1991) and Ghose and Kroner (1995), while the GARCH model with α-stable innovations is first proposed by McCulloch (1985) within a restricted framework and by Liu and Brorsen (1995*b*) within a more general context. The theoretical stationarity properties of GARCH models with α-stable innovations are studied by Panorska et al. (1995) and Mittnik et al. (2000). In what regards the estimation, Liu and Brorsen (1995*a*) propose the Maximum Likelihood (ML) approach, however for very specific values of the parameters and GARCH specification.

Our paper aims at alleviating the estimation problems in implementing the GARCH model with α-stable innovations under a very general parameter setting. Our implementation does not impose any parameter or model specification constraints and uses a GARCH specification with Student's innovations as an auxiliary model within the IndInf method. The choice of the auxiliary model is motivated by the fact that there is a rather natural correspondence between the two models: besides having the same number of parameters and a common GARCH specification for the conditional variance model, the degrees of freedom in the Student's t distribution is the direct counterpart of the parameter of stability or characteristic exponent in the stable distribution, as both measure the tail-thickeness of the distribution. Within a thorough Monte Carlo experiment and an empirical application to nine time series of financial returns of DJIA, SP500, IBM, sampled at different frequencies (daily, weekly, monthly), we provide valuable empirical evidence in favor of applying the IndInf method under very general model specification and parameter settings.

The rest of the paper is organized as follows: Section 2 gives a short introduction to α-stable family of distributions, Section 3 focuses on describing the model of interest, namely GARCH with α-stable innovations, Section 4 introduces the indirect inference method and Section 5 describes the estimation of the model of interest. Section 6 presents the results of a Monte Carlo experiment, while Section 7 shows results from estimating the model on real data. Section 8 concludes.

2 α-Stable Distributions

The stable family of distributions, which is also known under the name α-stable, constitutes a generalization of the Gaussian distribution by allowing for asymmetry and heavy tails. From a theoretical point of view, the use of models based on stable distributions is justified by the generalized version of the central limit theorem, in which the condition of finite variance is replaced by a much less restricting one concerning a regular behavior of the tails. It turns out that stable distributions are the only possible limiting laws for normalized sums of iid random variables (Feller (1966)). The lack of closed formulas for density and distribution functions (except for a few particular cases) has been, however, a major drawback of the stable distributions in applied fields.

A random variable X is said to have a stable distribution if and only if, for any positive numbers c_1 and c_2, there exists a positive number c and a real number d such that

$$cX + d \stackrel{d}{=} c_1X_1 + c_2X_2 \tag{1}$$

where X_1 and X_2 are independent and have the same distribution as X and $\stackrel{d}{=}$ stands for equality in distribution. If $d = 0$, X is said to be strictly stable. In order to show that the stable distribution is a generalization of the normal, let the variable $X \sim N(\mu, \sigma^2)$. The sum of n independent copies of X is $N(n\mu, n\sigma^2)$ distributed and $[X_1 + X_2 + \cdots + X_n]/c - d \stackrel{d}{=} X$, where $c = \sqrt{n}$ and $d = (\sqrt{n} - 1)\mu$.

The most concrete way to describe all possible stable distributions is by means of their characteristic function. The characteristic function of a stable random variable is of the form

$$\phi_1(t) := \begin{cases} \exp\left\{\mathrm{i}\delta_1 t - \gamma^\alpha |t|^\alpha \left[1 - \mathrm{i}\beta \operatorname{sgn}(t) \tan\frac{\pi\alpha}{2}\right]\right\} & \alpha \neq 1 \\ \exp\left\{\mathrm{i}\delta_1 t - \gamma |t| \left[1 - \mathrm{i}\beta \frac{2}{\pi} \operatorname{sgn}(t) \ln|t|\right]\right\} & \alpha = 1 \end{cases} \tag{2}$$

where $\operatorname{sgn}(t) = t/|t|$ for $t \neq 0$ (and 0 for $t = 0$), $\alpha \in]0, 2]$ is the index of stability or characteristic exponent that describes the tail-thickness of the distribution (small values correspond to thick tails), $\beta \in [-1, 1]$ describes the degree of asymmetry of the distribution, $\gamma \in \mathbb{R}^+$ is the scale parameter and $\delta_1 \in \mathbb{R}$ is the location parameter. When $\beta = 0$, the distribution is symmetric; when $\beta > 0$ the distribution turns out to be right skewed and for $\beta < 0$ the distribution is left skewed. The case $\beta = 1$ corresponds to a perfect positive skewness: the distribution has density zero on the negative semi-axis and positive values on the positive one. Conversely, when $\beta = -1$ the distribution is totally skewed to the left.

The stable distribution is characterized by four parameters $(\alpha, \beta, \gamma, \delta_1)$ and is denoted as $\mathrm{S}_1(\alpha, \beta, \gamma, \delta_1)$.

Note that when $\alpha = 1$, $\phi_1(t)$ is not continuous in the parameters. This is a source of problems for what concerns estimation and inferential purposes. An alternative way to write the characteristic function that overcomes the problem of discontinuity is the following

$$\phi_0(t) = \begin{cases} \exp\left\{\mathrm{i}\delta_0 t - \gamma^\alpha |t|^\alpha \left[1 + \mathrm{i}\beta \,\mathrm{sgn}\,(t) \tan\frac{\pi\alpha}{2} \left(|\gamma t|^{1-\alpha} - 1\right)\right]\right\} & \alpha \neq 1 \\ \exp\left\{\mathrm{i}\delta_0 t - \gamma|t| \left[1 + \mathrm{i}\beta\frac{2}{\pi} \,\mathrm{sgn}\,(t) \ln(\gamma|t|)\right]\right\} & \alpha = 1 \end{cases} \tag{3}$$

In this case the distribution is denoted by $\mathrm{S}_0(\alpha, \beta, \gamma, \delta_0)$. Expression (3) is more cumbersome, and the analytic properties have less intuitive meaning. Despite that, it is much more useful for what concerns statistical applications and, unless otherwise stated, we will refer to it in what follows.

The correspondence between δ_0 in S_0 and δ_1 in S_1 is given by:

$$\delta_0 = \begin{cases} \delta_1 + \beta\gamma \tan\frac{\pi\alpha}{2} & \text{if } \alpha \neq 1 \\ \delta_1 + \beta\frac{2}{\pi}\gamma \ln\gamma & \text{if } \alpha = 1 \end{cases} \tag{4}$$

On the basis of the above relationship, a $\mathrm{S}_0(\alpha, \beta, 1, 0)$ corresponds to a $\mathrm{S}_1(\alpha, \beta, 1, -\beta \tan\frac{\pi\alpha}{2})$, provided that $\alpha \neq 1$. Note also that when $\beta = 0$, the immaginary term (the asymmetry factor) in equations (2) and (3) disappears and the two parameterizations coincide.

Let $Z \sim S_0(\alpha, \beta, 1, 0)$. Then:

$$X = \begin{cases} \gamma(Z - \beta \tan \frac{\pi\alpha}{2}) + \delta_0 & \textit{if } \alpha \neq 1 \\ \gamma Z + \delta_0 & \textit{if } \alpha = 1 \end{cases} \tag{5}$$

is $\mathrm{S}_0(\alpha, \beta, \gamma, \delta_0)$ distributed. If, on the other hand, $Z \sim \mathrm{S}_1(\alpha, \beta, 1, 0)$, then:

$$X = \begin{cases} \gamma Z + \delta_1 & \textit{if } \alpha \neq 1 \\ \gamma(Z + \beta \frac{2}{\pi} \ln \gamma) + \delta_1 & \textit{if } \alpha = 1 \end{cases} \tag{6}$$

is $S_1(\alpha, \beta, \gamma, \delta_1)$ distributed. Z is thus the standardized version of X. In the sequel, the standardized α-stable distribution $\mathrm{S}_k(\alpha, \beta, 1, 0)$ is denoted by $\mathrm{S}_k(\alpha, \beta)$ for $k = 0, 1$. The characteristic function of a standardized α-stable distribution, symmetric around zero, reduces to

$$\phi_k(t) = \mathrm{e}^{-|t|^\alpha}$$

for both $k = 0, 1$.

The α-stable density functions admit closed form only in a very few special cases: if $\alpha = 2$, then the stable distribution coincides to a normal distribution with mean parameter δ and variance parameter $2\gamma^2$. Since $\tan \frac{\pi\alpha}{2} = 0$, the characteristic function is real and, hence, the distribution is always symmetric, regardless of the values of β, which becomes unidentified; if $\alpha = 1$ and $\beta = 0$, then the stable distribution coincides to a Cauchy distribution with location parameter δ and scale parameter γ; and, if $\alpha = 1/2$ and $\beta = \pm 1$, then the stable distribution coincides to a Lévy distribution with location parameter δ and scale parameter γ.

A further nice property of the stable distribution is that one can simulate pseudo-random numbers. Chambers et al. (1976) develop an algorithm by starting from two independent variables V and W, with V uniformly distributed on $(-\frac{\pi}{2}, \frac{\pi}{2})$ and W exponentially distributed with mean 1, and $0 < \alpha \leq 2$. Thus, symmetric stable pseudo-random numbers can be obtained as follows

$$Z = \begin{cases} \frac{\sin \alpha V}{(\cos V)^{1/\alpha}} \left[\frac{\cos((\alpha-1)V)}{W}\right]^{(1-\alpha)/\alpha} & \textit{if } \alpha \neq 1 \\ \tan W & \textit{if } \alpha = 1. \end{cases} \tag{7}$$

Z has a $\mathrm{S}_0(\alpha, 0) = \mathrm{S}_1(\alpha, 0)$ distribution. Non-symmetric stable pseudo-random numbers can be obtained for any $-1 \leq \beta \leq 1$ by defining $\zeta = \arctan(\beta \tan \frac{\pi\alpha}{2})/\alpha$ and constructing

$$Z = \begin{cases} \frac{\sin \alpha(\zeta+V)}{(\cos \alpha\zeta \cos V)^{1/\alpha}} \left[\frac{\cos(\alpha\zeta+(\alpha-1)V)}{W}\right]^{(1-\alpha)/\alpha} & \textit{if } \alpha \neq 1 \\ \frac{2}{\pi}\left[(\frac{\pi}{2} + \beta V)\tan V - \beta \ln\left(\frac{\frac{\pi}{2} W \cos V}{\frac{\pi}{2}+\beta V}\right)\right] & \textit{if } \alpha = 1, \end{cases} \tag{8}$$

which has a $\mathrm{S}_0(\alpha, \beta)$ distribution. Pseudo-random numbers containing also the location and the scale parameters δ_0 and γ may be straightforwardly obtained using the standardization given in Equation (5). Similarly, pseudo-random numbers with $\mathrm{S}_1(\alpha, \beta, \gamma, \delta_1)$ distribution can be obtained by exploiting Equation (6).

3 α-Stable GARCH Models

Several studies have highlighted the fact that heavy - tailedness of asset returns can be the consequence of conditional heteroskedasticity. The GARCH models of Bollerslev (1986) have become very popular for their ability to account for volatility clustering and heavy tails. However, some empirical studies (e.g., Yang and Brorsen (1993)) indicate that the tail behavior of GARCH models remains too short even with Student-t distributed error terms. Furthermore, the Student-t distribution lacks the stability-under-addition property. Stability is desirable because stable distributions provide a very good approximation for large classes of distributions. To overcome these weaknesses, one can apply GARCH models with α-stable innovations.

Although asymmetrical distributions and leverage effects are of particular importance in financial applications, we will focus in this paper mainly on the symmetric stable distributions, more precisely on GARCH models with symmetric stable innovations, which were first proposed by McCulloch (1985)[1]. However, the model introduced by McCulloch (1985) is restricted to absolute values and to an integrated conditional standard deviation model. Here we adopt the model introduced by Liu and Brorsen (1995*b*), which is more general, and adapt it to symmetric stable innovations.

The variable Y_t is defined to follow a symmetric α-stable GARCH(1,1) if:

$$Y_t = c + \epsilon_t, \qquad \epsilon_t = z_t \sigma_t \tag{9}$$

$$\sigma_t^\lambda = \omega + \alpha_1 |\epsilon_{t-1}|^\lambda + \beta_1 \sigma_{t-1}^\lambda \tag{10}$$

with $\omega \geqslant 0, \alpha_1 \geqslant 0, \beta_1 \geqslant 0$ and z_t being identically and independently distributed as a standard symmetric α-stable variable, $z_t \sim S_0(\alpha, 0)$. The model from above could be easily generalized to a GARCH(p, q) model by including additional lags. Traditionally, the GARCH model corresponds to $\lambda = 2$, which is also the focus of the present paper. If $\alpha = 2$, then the model from above collapses to the GARCH-normal model of Bollerslev (1986). Without loss of generality, we assume $c = 0$. Thus the unknown parameters of the model are: $\alpha, \omega, \alpha_1, \beta_1$.

As already mentioned by Liu and Brorsen (1995*a*), the stationarity conditions for such a model are stricter than the conditions for the normal GARCH. While for $\lambda < \alpha$ and $1 < \alpha \leqslant 2$, Mittnik et al. (2000) show that the process has a unique strictly stationary solution when $\omega > 0, \alpha_1 \geqslant 0, \beta_1 \geqslant 0$ and $E|z_t|^\lambda \alpha_1 + \beta_1 \leqslant 1$, there are no analytical stationarity conditions for the case $1 < \alpha \leqslant 2$ and $\lambda = 2$. As in Liu and Brorsen (1995*a*), we will verify the stationarity conditions of the model given in equations 9-10 for $\lambda = 2$ by means of Monte Carlo experiments (see Section (6)).

[1]Studying non-symmetric α-stable GARCH models as well as symmetric α-stable Threshold GARCH models is left for further research.

The specification given in equations 9 and 10 is so far implemented and estimated in Liu and Brorsen (1995a) by means of ML method for very specific values of λ. Although very appealing, applying the ML to estimate the model from above found so far little application in the existing literature. This might be due to the difficulty of implementing the ML approach to estimate the parameters of the stable distribution, given that the distribution has a closed-form density function only for very specific values of α. As an alternative, our paper applies the IndInf approach, which proves to be a valuable alternative to the ML method to estimate the stable parameters (Lombardi and Calzolari (2008) and Garcia et al. (2011)). Sections 4 and 5 give a through description of the IndInf method and of its application to estimate the parameters of the symmetric stable GARCH model specified above.

4 Indirect Inference Estimation

The indirect inference estimation method introduced by Gouriéroux et al. (1993) is a simulation-based technique, which is suitable to solve difficult or intractable estimation problems. The absence of the closed-form density for the stable-distribution as well as of the moments of order greater than two makes this method a valuable candidate for the estimation of its parameters (see Lombardi and Calzolari (2008) and Garcia et al. (2011)). The idea behind the IndInf estimation method is to replace the model of interest (true model) with an approximated model, which is easier to handle and estimate (auxiliary model). One important requirement of this technique is that one can easily simulate random values from the true model.

Let $y_t,\ t = 1, ..., T$ be a series of observed values of the random variable Y_t, characterized by a probability density function $f_0(\theta, y_t)$, which is intractable or difficult to handle. The unknown parameter vector θ is in the interior of the parameter set $\mathbf{\Theta} \in \mathbb{R}^r$, hence is of dimension $r \times 1$. Denote θ_0 to be the true value of the parameter vector

θ, which is also in the interior of $\mathbf{\Theta}$. Denote by $\mathcal{L}_0(y_1, y_2, \ldots, y_T; \theta)$ the log-likelihood function of the true model based on the observed values. Thus, given the intractability of $\mathcal{L}_0(y_1, y_2, \ldots, y_T; \theta)$, the maximum likelihood estimator of $\theta_{\mathbf{0}}$, which is given by:

$$\widehat{\theta_{\mathbf{ML}}} = \arg\max_{\theta} \mathcal{L}_0(y_1, y_2, \ldots, y_T; \theta) \tag{11}$$

is not available. However, as already mentioned above, it is assumed that it is possible to generate independent random draws of y_t and to obtain artificial values $y_1^s, ..., y_T^s$ for a given value of the parameter vector θ.

The idea behind IndInf estimation method is to find an auxiliary density function $f^a(y_t, \psi)$, which is easier to handle and which is characterized by the parameter vector ψ in the set $\Psi \in \mathbb{R}^q$. The corresponding log-likelihood function of the auxiliary model is given by $\mathcal{L}^a(y_1, y_2, \ldots, y_T; \psi)$, which is available analytically.

The IndInf estimation method implies the following steps: firstly, compute the pseudo-ML (PML) estimator of the pseudo-true ψ_0 from:

$$\hat{\psi} = \arg\max_{\psi} \mathcal{L}^a(y_1, y_2, \ldots, y_T; \psi) \tag{12}$$

Under standard regularity conditions of the PML estimation technique, its distribution is given by:

$$\sqrt{T}(\hat{\psi} - \psi_0) \xrightarrow{d} N\big(0_q, (J(\psi_0) I^{-1}(\psi_0) J(\psi_0))^{-1}\big),$$

where 0_q is a $q \times 1$ vector of zeros, $J(\psi_0)$ is minus the expectation of the Hessian of the log-likelihood of the auxiliary model and $I(\psi_0)$ is the Fisher information matrix of the auxiliary model.

Secondly, for a given value of θ, simulate S paths of length T from the initial model:

$y_1^s(\theta), \ldots, y_T^s(\theta)$, with $s = 1, \ldots, S$ and estimate

$$\widehat{\psi_{ST}}(\theta) = \arg\max_{\psi} \frac{1}{S} \sum_{s=1}^{S} \mathcal{L}^a(y_1^s(\theta), y_2^s(\theta), \ldots, y_T^s(\theta); \psi) \tag{13}$$

Thirdly, find the indirect inference estimator $\hat{\theta}$ such that $\hat{\psi}$ and $\widehat{\psi_{ST}}(\theta)$ are as close as possible:

$$\hat{\theta}(\Omega) = \arg\min_{\theta} [\hat{\psi} - \widehat{\psi_{ST}}(\theta)]' \, \Omega \, [\hat{\psi} - \widehat{\psi_{ST}}(\theta)] \tag{14}$$

where Ω is a weighting matrix, which is symmetric nonnegative definite and defines the metric. Denote $p(\theta)$ to be the link between θ and ψ as a binding function, such that $p(\theta_{\mathbf{0}}) = \psi_0$. The third step involves, in general, numerical optimization, since, in most cases, there is no analytical correspondence between ψ and θ, i.e., there is no analytical solution to $p(\theta) = \psi$.

An alternative approach, introduced by Gallant and Tauchen (1996), considers directly the score of the auxiliary model. The idea is to find the optimal θ such that the score, computed on the simulated observations and at the value $\hat{\psi}$ is as close as possible to zero. Provided that a closed form for the gradient of the auxiliary model is available, this approach has an important computational advantage: it avoids the numerical optimization in Equation (13).

Under certain regularity conditions (see, Gouriéroux et al. (1993)), the indirect inference estimator $\hat{\theta}(\Omega)$ is consistent and asymptotically normal for S fixed and $T \to \infty$:

$$\sqrt{T}(\hat{\theta}(\Omega) - \theta_{\mathbf{0}}) \xrightarrow{d} N\big(0, W(S, \Omega)\big)$$

where

$$\begin{aligned} W(S, \Omega) = \Big(1 + \frac{1}{S}\Big) \Big[\frac{\partial p'}{\partial \theta}(\theta_{\mathbf{0}}) \, \Omega \, \frac{\partial p}{\partial \theta'}(\theta_{\mathbf{0}})\Big]^{-1} \frac{\partial p'}{\partial \theta}(\theta_{\mathbf{0}}) \\ \times \Omega \, J(\psi_0)^{-1} I(\psi_0) J(\psi_0)^{-1} \, \Omega \, \frac{\partial p}{\partial \theta'}(\theta_{\mathbf{0}}) \Big[\frac{\partial p'}{\partial \theta}(\theta_{\mathbf{0}}) \, \Omega \, \frac{\partial p}{\partial \theta'}(\theta_{\mathbf{0}})\Big]^{-1}. \end{aligned}$$

First, one should note that the dimension of the auxiliary parameter ψ, namely q must be greater than or equal to the dimension of the parameter of interest θ, namely r in order to get a unique solution for $\hat{\theta}$. Second, when the problem is just identified, i.e. the dimension of the two parameter vectors is equal, the results are independent of the choice of the matrices that define the metrics, Ω. On the contrary, when $q > r$, it is necessary to choose a metric Ω to measure the distance between $\hat{\psi}$ and $\widehat{\psi_{ST}}(\theta)$. The optimal choice of Ω is

$$\Omega^* = J(\psi_0)I(\psi_0)^{-1}J(\psi_0)$$

and the corresponding IndInf estimator is denoted by $\hat{\theta}^*$. Based on Ω^*, the variance-covariance matrix of the IndInf estimator reduces to

$$W_S^* \equiv W(S, \Omega^*) = \left(1 + \frac{1}{S}\right)\left[\frac{\partial p'}{\partial \theta}(\theta_{\mathbf{0}})\, J(\psi_0)I(\psi_0)^{-1}J(\psi_0)\, \frac{\partial p}{\partial \theta'}(\theta_{\mathbf{0}})\right]^{-1} \quad (15)$$

5 Indirect Estimation of Stable GARCH Processes

Since simulated values from α-stable distributions can be straightforwardly obtained as described in Section 2, the indirect inference approach is particularly suited to estimate the parameters of the α-stable GARCH models presented in Section 3. We implement this method by considering as an auxiliary model a GARCH approach with Student's t innovations[2]. The choice of the Student-t distribution is motivated by the fact that its parameters have a clear and interpretable matching to those of the α-stable distribution: the degrees of freedom parameter ν is naturally linked to the tail parameter α, as both describe the thickness of the tail. Here we implement the Student-t distribution in terms of $\eta = \nu^{-1}$, which is the reciprocal of the degrees of freedom ν. Thus, the auxiliary model is given by

$$Y_t = c_a + \xi_t, \qquad \xi_t = u_t\sqrt{h_t} \quad (16)$$

[2]A similar approach is implemented by Calzolari et al. (2004)

$$h_t = \omega_a + \alpha_{1,a}\xi_{t-1}^2 + \beta_{1,a}h_{t-1} \tag{17}$$

with $\omega_a \geqslant 0, \alpha_{1,a} \geqslant 0, \beta_{1,a} \geqslant 0$ and u_t is identically and independently distributed as a symmetric Student-t variable, $u_t \sim t_{1/\eta}$. Similar to Equation 9, we set c_a to 0.

In the IndInf framework presented in Section 4, the parameter vector θ is given by $\theta = (\omega, \alpha_1, \beta_1, \alpha)'$ and is of dimension 4×1 ($r = 4$) and the parameter vector ψ is given by $\psi = (\omega_a, \alpha_{1,a}, \beta_{1,a}, \eta)'$ and is also of dimension 4×1 ($q = 4$). Thus the dimension of the true parameter vector and the auxiliary parameter vector is the same and, therefore, in the IndInf optimization routine we replace the metric Ω, respectively Ω^* by the identity matrix, I_4.

Numerical results on the indirect inference estimation of the α-stable GARCH(1,1) model are displayed in sections 6 and 7.

6 Monte Carlo Study

A detailed set of Monte Carlo experiments has been performed to check the reliability of the indirect estimation method when applied to the GARCH(1,1) model with symmetric α-stable noise. Student's t GARCH(1,1) has been used as an auxiliary model. As already mentioned in Section 5, there is a rather "natural" correspondence between the parameters of the two models (same number of parameters; just identified case): ω, α_1 and β_1 are the GARCH parameters of the true model, while ω_a, $\alpha_{1,a}$ and $\beta_{1,a}$ are the GARCH parameters of the auxiliary model. The tail-thickness parameter in the true model is α, while in the auxiliary model is η, which is the reciprocal of the (symmetric) Student's-t degrees of freedom ν.

The values of the parameters have been chosen as to mimic real-case values (the only exception being the ω parameter, which has been chosen to be larger); a moderately large length of the time series has been adopted in all experiments (T=10000, roughly comparable with the length of the daily series in the empirical application described in Section 7). As a multiplicative length-factor to produce simulated series, we take $S = 10$: thus 100000 is in all experiments the length of the simulated series to be handled by the auxiliary model.

As a constant for the GARCH parameter, namely ω, the value 0.01 has been always adopted: it will therefore not be reported in Table 6, which gives the estimation results for the present Monte Carlo experiments. A value of the persistence parameter $\alpha_1+\beta_1 = 0.98$ has been obtained with different combinations of α_1 and β_1, varying α_1 from 0.20 to 0.05.

As far as the tail-thickness parameter α is concerned, four different values have been experimented with, ranging from a ”close to Gaussian” value (1.98) to a moderate ”fat-tail” value (1.80).

Other combinations of parameters have also been tried. For instance a larger value of the persistence parameter $\alpha_1 + \beta_1 = 0.99$, as well as smaller values of α (till 1.60), implying even thicker tails. Not all these additional experiment were successful, as the combination of GARCH and thick-tail noise in several cases led to ”exploding” values of the simulated series. This was a sort of experimental evaluation of the non-stationarity of the process for some combinations of parameter values. At the same time these parameter values were not realistic, when compared with the values estimated from real series, thus they are not reported here for the sake of brevity. On the contrary, there were no problems when simulating series of data using the ”realistic” combinations of parameters included in the table. This has been considered a sort of empirical or computational assessment of the stationarity conditions, not otherwise analytically analyzed.

Each set of simulation results presented in Table 6 has been obtained with R=1000 Monte Carlo replications. In each replication, T=10000 ”pseudo observations” have been generated, and indirect estimation is obtained by setting to zero (or ”minimizing”) the score of the auxiliary model on $S \times T$=100000 simulated data (as in Gallant and Tauchen (1996)). Performances of the estimation method are quite remarkable. With very few exceptions,

Table 1: Monte Carlo results: average estimates and standard deviations (in parentheses) over $R = 1000$ Monte Carlo replications, based on $T = 10000$ number of observations and $S = 10$ number of simulation paths

Parameters of the true model			Estimated parameters					
			True model			Auxiliary model		
α_1	β_1	α	α_1	β_1	α	$\alpha_{1,a}$	$\beta_{1,a}$	η
.200	.780	1.80	.200	.780	1.80	.157	.779	.235
			(.0097)	(.0077)	(.0168)	(.0076)	(.0074)	(.0122)
		1.85	.200	.780	1.85	.161	.779	.196
			(.0099)	(.0082)	(.0161)	(.0080)	(.0079)	(.0124)
		1.90	.200	.779	1.90	.167	.779	.153
			(.0100)	(.0088)	(.0149)	(.0085)	(.0085)	(.0127)
		1.95	.200	.779	1.95	.175	.779	.102
			(.0102)	(.0094)	(.0126)	(.0092)	(.0091)	(.0137)
		1.98	.201	.779	1.98	.184	.779	.059
			(.0109)	(.0100)	(.0089)	(.0109)	(.0097)	(.0154)
.100	.880	1.80	.100	.880	1.80	.083	.879	.153
			(.0060)	(.0055)	(.0193)	(.0044)	(.0051)	(.0127)
		1.85	.100	.880	1.85	.081	.879	.196
			(.0059)	(.0058)	(.0161)	(.0047)	(.0055)	(.0123)
		1.90	.100	.880	1.90	.083	.879	.153
			(.0062)	(.0064)	(.0149)	(.0051)	(.0061)	(.0127)
		1.95	.100	.879	1.95	.087	.879	.102
			(.0066)	(.0071)	(.0071)	(.0057)	(.0068)	(.0137)
		1.98	.100	.879	1.98	.092	.879	.060
			(.0069)	(.0076)	(.0094)	(.0064)	(.0073)	(.0156)
.050	.930	1.80	.050	.929	1.80	.039	.929	.234
			(.0033)	(.0037)	(.0168)	(.0026)	(.0037)	(.0126)
		1.85	.050	.930	1.85	.040	.930	.196
			(.0036)	(.0043)	(.0161)	(.0028)	(.0041)	(.0123)
		1.90	.050	.930	1.90	.042	.929	.153
			(.0062)	(.0040)	(.0050)	(.0149)	(.0149)	(.0127)
		1.95	.050	.930	1.95	.044	.929	.102
			(.0045)	(.0060)	(.0126)	(.0038)	(.0057)	(.0137)
		1.98	.050	.929	1.98	.046	.929	.060
			(.0049)	(.0069)	(.0094)	(.0044)	(.0066)	(.0156)

estimates of the model of interest (true model) seem unbiased (differences between the average estimates and the parameters used to generate the data are observable only after the third digit). Moreover, our computations show that the empirical standard deviations presented in parentheses are very close to the theoretical standard deviations as derived from Equation (15) (differences are only observable after five digits).[3]

Concerning α_1 and β_1, the variance of the parameters in the model of interest, as expected, are always larger than the corresponding auxiliary parameters, but the difference is not very large.

The estimate of β_1 is nearly unbiased also in the auxiliary model (see $\beta_{1,a}$); on the contrary, $\widehat{\alpha_1}$ is always remarkably downward biased in the auxiliary model (see $\alpha_{1,a}$). The bias is being adjusted by the indirect estimation procedure. Moreover, one can observe the direct correspondence between the estimates of α and the estimates of ν: larger the α is, smaller the η is, which indicates a large value for the Student's t degrees of freedom ν.

7 Empirical Application

The GARCH(1,1) model with symmetric α-stable noise has been estimated on the series of monthly, weekly and daily log-returns of:

1. Dow Jones Industrial Average (DJIA) stock index from May 4th, 1950 to June 25th, 2012 (744, 3241, and 16212 observations for the monthly, weekly and daily series, respectively);
2. Standard & Poor's 500 (SP500) index from January 1st, 1964 to June 25th, 2012 (581, 2529, and 12649 observations, respectively) and
3. IBM share's prices from January 1st, 1973 to June 25th, 2012 (473, 2059, and 10299 observations, respectively).

[3]Tables with theoretical standard deviations can be obtained from the authors upon request.

Data have been obtained from Thomson Reuters Datastream.

Indirect estimation has been performed using a large value of the multiplicative length-factor for the simulated series ($S = 100$). Convergence has always been achieved inside the parameter space of the auxiliary model. The only "additional" problem with respect to the simulation exercise is the choice of a "good" value for the α parameter, as an initial value for the iterative procedure.

Table 2: Empirical results. Standard deviations are reported in parentheses. M stands for monthly, W stands for weekly and D stands for daily.

Data		Estimated parameters							
		True model				Auxiliary model			
		ω	α_1	β_1	α	ω_a	$\alpha_{1,a}$	$\beta_{1,a}$	η
DJIA	M	.111e-3	.109	.820	1.951	.109e-3	.100	.808	.094
		(.4368e-4)	(.0300)	(.0444)	(.0394)				
	W	.987e-5	.068	.899	1.946	.887e-5	.059	.899	.106
		(.2127e-5)	(.0094)	(.0128)	(.0188)				
	D	.567e-6	.049	.933	1.899	.474e-6	.041	.933	.152
		(.6233e-7)	(.0028)	(.0033)	(.0101)				
SP500	M	.890e-4	.094	.848	1.938	.721e-4	.080	.848	.124
		(.4346e-4)	(.0282)	(.0407)	(.0501)				
	W	.914e-5	.092	.882	1.950	.753e-5	.078	.884	.105
		(.2364e-5)	(.0125)	(.0146)	(.0227)				
	D	.312e-6	.052	.936	1.895	.240e-6	.043	.936	.158
		(.4756e-7)	(.0030)	(.0032)	(.0115)				
IBM	M	.375e-3	.090	.821	1.940	.348e-3	.080	.812	.116
		(.1993e-3)	(.0318)	(.0628)	(.0529)				
	W	.145e-4	.039	.936	1.871	.127e-4	.032	.936	.177
		(.4180e-5)	(.0068)	(.0096)	(.0293)				
	D	.124e-5	.032	.952	1.851	.103e-5	.026	.952	.193
		(.1716e-6)	(.0023)	(.0027)	(.0131)				

Table 2 reports the estimated values of the parameters for both true and auxiliary models, as well as the standard deviations (in parentheses) of the estimated indirect inference parameters of the model of interest, computed based on Equation (15).[4] For a financial analyst, it can be interesting to observe that the largest value of the persistence parameter

[4]We refrain from reporting the standard deviations of the estimated parameters of the auxiliary model, as they are of no interest for our purposes. However, they can be easily obtained in the standard way typical to the PML procedure.

$(\alpha_1 + \beta_1)$ is for the daily returns of the SP500 index (0.988), and the smallest value of the α parameter (thickest tails of the standardized GARCH residuals) is for the daily returns of IBM (1.85). This is an expected result, since the indexes mirror an aggregate behavior of the composing stocks, which are differently affected by extreme financial events (e.g., Black Monday in 1987 or the previous financial crisis). This leads to thinner tails for the index return series than for a certain composing stock, especially for large-cap stocks with high liquidity, such as IBM.

Moreover, within each group of the considered financial assets, we observe that the returns sampled at the highest frequency (daily) exhibit the largest persistence as well as the fattest tail, as indicate it by the largest value of $\alpha_1 + \beta_1$ and the smallest value of α. This result is not surprising, since it is a well known fact that increasing the frequency in sampling stock returns increases the clustering and the persistence degree as well as the degree of fat-thickeness. Moreover, as the values of $\frac{\hat{\omega}}{1-\hat{\alpha}_1-\hat{\beta}_1}$ show, returns sampled at higher frequency exhibit larger (unconditional) variance than the ones sampled at lower frequency.

For a computational econometrician, it can be interesting to observe that β_1 in the model of interest is usually almost equal to its correspondent $\beta_{1,a}$ in the auxiliary model, but α_1 is always larger in the model of interest than its correspondent $\alpha_{1,a}$ in the auxiliary model. The close behavior between the true and the auxiliary model is motivated by the monotonic relationship between the α parameter in the model of interest and the "degrees of freedom" of the Student's-t in the auxiliary model, ν: the smaller α, the larger $\eta = \nu^{-1}$ (reciprocal of the degrees of freedom).

8 Conclusions

In this paper we apply the indirect inference method to estimate the parameters of a GARCH model with stable innovations. As most of the financial returns exhibit conditional heteroskedasticity and fat-tails, there is a common practice among scholars and practitioners to capture these features by means of GARCH models with fat-tailed distributed innovations. A standard choice is to consider Student's t distributed innovations, however, at the costs of lack of stability under aggregation. An alternative is to consider α-stable distributions that remain stable under aggregation and combination, which is particularly appealing in portfolio theory. However, this alternative comes at the costs of estimation, due to the absence of closed form density function and of moments for most of the parameter values. As a solution to this problem we apply the indirect inference method introduced by Gouriéroux et al. (1993) with a GARCH conditional variance and Student's t innovations as auxiliary model.

The simulation study reveals very good results for a large pallet of parameter choices: with very few exceptions, the estimates of the model of interest seem unbiased at a minimal variance cost. The empirical results from applying the method to nine series of financial returns sampled at three different frequencies provide further empirical evidence in favor of using the indirect inference method for estimating the parameters of a GARCH model with stable innovations.

References

Bollerslev, T. (1986), 'Generalized autoregressive conditional heteroskedasticity', *Journal of Econometrics* **31**, 307–327.

Calzolari, G., Fiorentini, G. and Sentana, E. (2003), 'Maximum likelihood estimation and inference in multivariate conditionally heteroscedastic dynamic regression models with student t innovations', *Journal of Business & Economic Statistics* **21**, 532–446.

Calzolari, G., Fiorentini, G. and Sentana, E. (2004), 'Constrained indirect estimation', *Review of Economic Studies* **71**, 945–973.

Chambers, J., Mallows, C. and Stuck, B. W. (1976), 'A method for simulating stable random variables', *Journal of the American Statistical Association* **71**, 340–344.

de Vries, C. (1991), 'On the relation between garch and stable processes', *Journal of Econometrics* **48**, 313–324.

Feller, W. (1966), *An Introduction to Probability Theory and its Applications*, John Wiley and Sons, New York.

Gallant, A. R. and Tauchen, G. (1996), 'Which moments to match?', *Econometric Theory* **12**, 657–681.

Garcia, R., Renault, E. and Veredas, D. (2011), 'Estimation of stable distributions with indirect inference', *Journal of Econometrics* **161**, 325–337.

Ghose, D. and Kroner, K. (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.

Ghysels, E., Harvey, A. and Renault, E. (1996), Stochastic volatility, *in* G. Maddala and C. R. Rao, eds, 'Handbook of Statistics', vol. 14, pp. 119-192. Elsevier Science Publishers, Amsterdam.

Gouriéroux, C., Monfort, A. and Renault, E. (1993), 'Indirect inference', *Journal of Applied Econometrics* **8**, S85–S118.

Liu, S. and Brorsen, B. (1995*a*), 'Garch-stable as a model of futures price movements', *Review of Quantitative Finance and Accounting* **5**, 155–167.

Liu, S. and Brorsen, B. (1995*b*), 'Maximum likelihood estimation of a garch-stable model', *Journal of Applied Econometrics* **10(3)**, 273–285.

Lombardi, M. and Calzolari, G. (2008), 'Indirect estimation of α-stable distributions and processes', *Econometrics Journal* **11**, 193–208.

Lombardi, M. and Calzolari, G. (2009), 'Indirect estimation of α-stable stochastic volatility models', *Computational Statistics and Data Analysis* **53(6)**, 2298–2308.

McCulloch, J. (1985), 'Interest-risk sensitive deposit insurance premia: Stable ach estimates', *Journal of Banking and Finance* **9**, 137–156.

Mittnik, S., Paolella, M. and Rachev, S. (2000), 'Diagnosing and treating the fat tails in financial returns data', *Journal of Empirical Finance* **7**, 389–416.

Panorska, A., Mittnik, S. and Rachev, S. (1995), 'Stable garch models for financial time series', *Applied Mathematics Letters* **8**, 33–37.

Yang, S. and Brorsen, B. (1993), 'Nonlinear dynamics of daily futures prices: Conditional heteroskedasticity or chaos?', *Journal of Futures Markets* **13(2)**, 175–191.