Erschienen in: Computational Statistics & Data Analysis ; 76 (2014). - S. 158-171
https://dx.doi.org/10.1016/j.csda.2013.07.028

Estimating GARCH-type models with symmetric stable innovations: Indirect inference versus maximum likelihood

Giorgio Calzolari [a], Roxana Halbleib [b,*], Alessandro Parrini [c]

[a] *Department of Statistics, University of Firenze, Italy*
[b] *Department of Economics, University of Konstanz, P.O. Box 124, Universitaetstrasse 10, 78464, Konstanz, Germany*
[c] *Department of Econometrics, VU University Amsterdam, Netherlands*

A B S T R A C T

Financial returns exhibit conditional heteroscedasticity, asymmetric responses of their volatility to negative and positive returns (leverage effects) and fat tails. The α-stable distribution is a natural candidate for capturing the tail-thickness of the conditional distribution of financial returns, while the GARCH-type models are very popular in depicting the conditional heteroscedasticity and leverage effects. However, practical implementation of α-stable distribution in finance applications has been limited by its estimation difficulties. The performance of the indirect inference approach using GARCH models with Student's t distributed errors as auxiliary models is compared to the maximum likelihood approach for estimating GARCH-type models with symmetric α-stable innovations. It is shown that the expected efficiency gains of the maximum likelihood approach come at high computational costs compared to the indirect inference method.

Keywords:
Symmetric α-stable distribution
GARCH-type models
Indirect inference
Maximum likelihood
Leverage effects
Student's t distribution

1. Introduction

Most of the financial returns exhibit conditional heteroscedasticity and heavy-tailedness. While conditional heteroscedasticity is standardly captured by means of GARCH or stochastic volatility (SV) models (e.g. Bollerslev (1986) and Ghysels et al. (1996)), depicting the empirically observed fat-thickness 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 conditional heteroscedasticity 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. Proposals of likelihood-free inference are only recently available in the Bayesian context: e.g., Peters et al. (2012).

* Corresponding author. Tel.: +49 0 753188 5373; fax: +49 0 7531 88 4450.
E-mail addresses: calzolar@ds.unifi.it (G. Calzolari), roxana.chiriac@uni-konstanz.de, roxana.halbleib@uni-konstanz.de (R. Halbleib), a.parrini@student.vu.nl (A. Parrini).

In this paper we focus on estimating GARCH-type models with symmetric α-stable innovations by means of the indirect inference (IndInf) method proposed by Gouriéroux et al. (1993) and maximum likelihood (ML) as described in Nolan (1997). The indirect inference estimation approach has already proved its adequacy in estimating the parameters of the stable distribution in Lombardi and Calzolari (2008) and Lombardi and Veredas (2009) and Garcia et al. (2011). Lombardi and Calzolari (2009) use the indirect inference approach to estimate a SV model with α-stable innovations. Differently from their approach, we focus on comparing the estimation results stemming from IndInf and ML when estimating GARCH-type models with symmetric α-stable innovations. We focus on estimating GARCH-type specifications due to their popularity among practitioners and academics. The popularity of GARCH models over SV models originates in their ability of straightforwardly capturing empirical features of financial volatilities, such as: asymmetric responses to negative and positive returns, known in literature as leverage effects, high persistence or long memory as well as causality and correlation effects with further economic variables such as: Volatility Index (VIX), inflation, etc.. To illustrate this, we estimate, besides simple GARCH specifications, also Threshold GARCH (TGARCH) models as introduced by Glosten et al. (1993) that capture leverage effects, which are highly relevant in financial applications.

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 (1995b) 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, 2002). In what regards the estimation, Liu and Brorsen (1995a) propose the ML approach, however for very specific values of the parameters and GARCH specifications.

The aim of our paper is to alleviate the estimation problems in implementing GARCH-type models with symmetric α-stable innovations under a very general parameter setting. Our implementation does not impose any parameter or model specification constraints. The IndInf estimation method uses GARCH-type specifications with Student's innovations as auxiliary models. 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 common GARCH-type specifications for the conditional heteroscedasticity, 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-thickness of the distribution. In what regards the ML implementation, we apply the method described by Nolan (1997) and Matsui and Takemura (2006) based on the numerical evaluation of the symmetric stable density and its derivatives for a wide range of parameter constellations. Furthermore, we adapt their procedure to estimate both the parameters of the symmetric α-stable distribution and of the GARCH specifications. As an alternative to Nolan's (1997) approach, one can consider the approach of Chenyao et al. (1999) that uses fast Fourier transforms to approximate the stable density functions.

Thus, our paper, besides contributing to the existing literature for implementing further types of GARCH models, such as TGARCH models with symmetric α-stable innovations, it also directly compares the performance of standard ML and IndInf when estimating a wide range of GARCH-type models with symmetric α-stable innovations.

Within a thorough Monte Carlo experiment and an empirical application to twelve time series of financial returns of DJIA, SP500, IBM and GE, sampled at different frequencies (daily, weekly, monthly), we provide valuable empirical evidence in favor of applying IndInf over ML under a very general model specification and parameter settings. We show that, although both methods provide accurate estimates, the expected efficiency gain of ML comes at high computational costs: besides being easier to implement, IndInf reports estimation results up to ten times faster than ML.

The rest of the paper is organized as follows: Section 2 gives a short introduction to the symmetric α-stable distributions, Section 3 focuses on describing the models of interest, namely GARCH and TGARCH with symmetric α-stable innovations, Section 4 shortly introduces the estimation methods and describes their practical implementation for estimating the models of interest. Section 5 presents the results of a Monte Carlo experiment, while Section 6 shows results from estimating the models on real data. Section 7 concludes.

2. Symmetric 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. In this paper, we focus on the symmetric stable distribution, which is a subclass of the stable family of distributions with no asymmetry. 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.

In general 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 k and a real number d such that

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

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

$X \sim N(\mu, s^2)$. The sum of n independent copies of X is $N(n\mu, ns^2)$ distributed and $[X_1 + X_2 + \cdots + X_n]/k - d \stackrel{d}{=} X$, where $k = \sqrt{n}$ and $d = (\sqrt{n} - 1)\mu$.

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

$$\phi(t) := \exp\left\{\mathrm{i}\delta t - \sigma^{\alpha}|t|^{\alpha}\right\}, \tag{2}$$

where $\alpha \in]0, 2]$ is the index of stability or characteristic exponent that describes the tail-thickness of the distribution (small values correspond to thick tails), $\sigma \in \mathbb{R}^+$ is the scale parameter and $\delta \in \mathbb{R}$ is the location parameter. This representation is a slight variation of parameterization (M) in Zolotarev (1986), adapted to the symmetric stable distributions. The symmetric stable distribution is, thus, characterized by three parameters (α, σ, δ) and is denoted as $S(\alpha, \sigma, \delta)$.

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

$$X = \sigma Z + \delta \tag{3}$$

is $S(\alpha, \sigma, \delta)$ distributed. Z is thus the "standardized" version of X. The characteristic function of a standardized symmetric α-stable distribution reduces to

$$\phi(t) = \exp\{-|t|^{\alpha}\}.$$

The symmetric α-stable density functions admit closed form only in a very few special cases: if $\alpha = 2$, then the symmetric stable distribution coincides to a normal distribution with mean parameter δ and variance parameter $2\sigma^2$; if $\alpha = 1$, then the stable distribution coincides to a Cauchy distribution with location parameter δ and scale parameter σ.

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

$$Z = \begin{cases} \dfrac{\sin \alpha V}{(\cos V)^{1/\alpha}} \left[\dfrac{\cos((\alpha - 1)V)}{W}\right]^{(1-\alpha)/\alpha} & \text{if } \alpha \neq 1 \\ \tan V & \text{if } \alpha = 1. \end{cases} \tag{4}$$

Z has a $S(\alpha, 1, 0)$ distribution. Pseudo-random numbers containing the location and the scale parameters δ and σ may be straightforwardly obtained using the standardization given in Eq. (3). One should notice that most of computer packages using Eq. (4) do not generate "standardized" stable random numbers in the classical sense. For instance, when $\alpha = 2$, the generated normal random variables have variance 2. Thus, in order to get variables "standardized" in the ordinary sense, we divide them by $2^{1/\alpha}$.[1]

3. Symmetric α-stable GARCH-type models

Several studies have highlighted the fact that heavy—tailedness of asset returns can be the consequence of conditional heteroscedasticity. 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's t distributed error terms. Furthermore, the Student's 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.

GARCH models with symmetric stable innovations were first proposed by McCulloch (1985). 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 (1995b), which is more general, and adapt it to a standard GARCH specification with symmetric stable innovations.

In the context of stable distributions, the GARCH specifications model the squared of conditional scale of the distribution. Thus, due to the fact that, for these distributions, the second moments do not exist, using the term "conditional heteroscedasticity" is not entirely correct. However, for convenience, we still use it in what it follows.

We define the variable Y_t to follow a symmetric α-stable GARCH(1, 1) if:

$$Y_t = c + \epsilon_t, \quad \epsilon_t = z_t \sigma_t, \tag{5}$$

$$\sigma_t^2 = \omega + a\epsilon_{t-1}^2 + b\sigma_{t-1}^2 \tag{6}$$

with $\omega > 0$, $a \geqslant 0$, $b \geqslant 0$ and z_t being identically and independently distributed as a standard symmetric α-stable variable. The model from above could be easily generalized to a GARCH(p, q) model by including additional lags. For $\alpha = 2$, it

[1] Here we would like to thank one of the referees for suggesting to us this standardization. This will always be done implicitly in what follows. Thus, for $\alpha = 2$, the "standardization" is given by $\sqrt{2}$.

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: α, ω, a, b.

As already mentioned by Liu and Brorsen (1995a), the stationarity conditions for a symmetric α-stable GARCH model are stricter than the conditions for the normal GARCH. However, by applying Lyapunov type exponents, one can obtain the necessary and sufficient conditions for assuring the strictly stationarity of the stable GARCH process. Thus, the top-Lyapunov condition $E\log(az_t^2 + b) < 0$ is the necessary and sufficient condition for the existence of a strictly stationary solution to Eqs. (5)–(6) (Mittnik et al., 2002). This result was first established by Nelson (1990) under the assumption that $\max(E\log z_t^2, 0) < \infty$ and $\omega, \alpha > 0$ (see Francq and Zakoian (2010), for a more detailed discussion on the stationarity of GARCH). The top-Lyapunov condition given above is numerically verified within the empirical applications described in Sections 5 and 6.

The standard GARCH model described so far ignores the information on the direction of returns and how they affect volatility. In practice the volatility responds asymmetrically to positive and negative returns. More precisely the reaction to negative returns is greater than the reaction to positive returns. This effect, first identified by Black (1976) and known in literature as the leverage effect, can be captured within the GARCH framework by so called Threshold GARCH model, which was proposed by Glosten et al. (1993).

We define here the variable Y_t to follow a symmetric α-stable TGARCH(1, 1) if:

$$Y_t = c + \epsilon_t, \quad \epsilon_t = z_t\sigma_t, \tag{7}$$

$$\sigma_t^2 = \omega + a\epsilon_{t-1}^2 + \gamma\epsilon_{t-1}^2\mathbf{1}_{[\epsilon_{t-1}<0]} + b\sigma_{t-1}^2 \tag{8}$$

where $\mathbf{1}_{[\cdot]}$ is the indicator function, ω, a, b follow the assumptions of the GARCH model described above and z_t is identically and independently distributed as a standard symmetric α-stable variable. By introducing an interaction term of the lagged squared shocks with a dummy for the sign of the shock, TGARCH manages to account for the leverage effect. In particular, if $\gamma > 0$, then the impact of a negative shock in $t-1$ on the "conditional variance" in t is larger compared to the impact of a positive shock. Clearly, the slope from a positive to a negative shock is not smooth, but discrete. Similar to the standard GARCH model, TGARCH can be easily generalized to TGARCH(p, q) by including additional lags. Moreover, we set $c = 0$. Thus the unknown parameters of the model are: $\alpha, \omega, a, b, \gamma$.

The necessary and sufficient conditions for assuring the strict stationarity of the stable TGARCH process can be derived similar to the GARCH process from the top-Lyapunov condition. Thus, the necessary and sufficient condition for the existence of a strictly stationary solution to Eqs. (7)–(8) is given by $E\log(az_t^2 + \gamma z_t^2\mathbf{1}_{[z_t<0]} + b) < 0$ under the same assumptions as for the GARCH model (see Francq and Zakoian (2012)). Similar to the GARCH case, we verify this condition numerically in Sections 5 and 6.

The specification given in Eqs. (5) and (6) is so far implemented and estimated in Liu and Brorsen (1995a) by means of ML for very specific values of the parameters. Although very appealing, applying 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. Nolan (1997) and Matsui and Takemura (2006) implement the ML approach by numerically evaluating the stable density function and its derivatives in order to estimate the parameters specific to the stable distribution. In this paper we adapt their approach and integrate the GARCH and TGARCH models in the estimation procedure. As an alternative, we also apply the IndInf method, which proves to be a valuable alternative to estimate the stable parameters (Lombardi and Calzolari, 2008; Garcia et al., 2011). Section 4 gives a thorough description of the two estimation methods and of their practical implementation when estimating the parameters of symmetric stable (T)GARCH models as specified above.

4. Estimation methods

Let $y_t, t = 1, \ldots, T$ be a series of observed values of the random variable Y_t defined in Eqs. (5)–(6) or (7)–(8) and characterized by the density function $f_0(y_t; \alpha, \sigma_t, 0)$. Given the symmetry of the distribution, we have that: $f_0(y_t; \alpha, \sigma_t, 0) = f_0(-y_t; \alpha, \sigma_t, 0)$. The link to the standard symmetric stable distribution is given by: $\frac{1}{\sigma_t}f_0(\frac{y_t}{\sigma_t}; \alpha, 1, 0) = f_0(y_t; \alpha, \sigma_t, 0)$. Let's denote the unknown parameter vector $\theta = (\alpha, \omega, a, b)'$ if Y_t has the representation given in Eqs. (5) and (6) and $\theta = (\alpha, \omega, a, b, \gamma)'$ if Y_t has the representation given in Eqs. (7) and (8). Thus, θ is in the interior of the parameter set $\boldsymbol{\Theta} \in \mathbb{R}^r$, where $r = 4$, if Y_t is given in Eqs. (5) and (6) and $r = 5$ if Y_t is given in Eqs. (7) and (8). Denote $\theta_\mathbf{0}$ to be the true value of the parameter vector θ, which is also in the interior of $\boldsymbol{\Theta}$.

Maximum likelihood. The absence of the closed-form density for the stable distribution makes the estimation of the parameters of the stable distribution by ML a very difficult task. Nolan (1997) overcomes this difficulty, by applying the (M) parameterization of Zolotarev (1986) and derives numerical formulas for the computation of the stable density and its derivatives. Thus, he derives the density of a standardized symmetric stable distributed variable $z_t = \frac{Y_t}{\sigma_t}$, for $z_t > 0$ and $\alpha \neq 1$ to be given by:

$$f(z_t; \alpha, 1, 0) = \frac{\alpha}{\pi|\alpha - 1|z_t}\int_0^{\frac{\pi}{2}} g(x; \alpha, z_t)\exp\{-g(x; \alpha, z_t)\}dx, \tag{9}$$

where

$$g(x;\alpha,z_t)=\left(\frac{z_t\cos x}{\sin\alpha x}\right)^{\frac{\alpha}{\alpha-1}}\frac{\cos(\alpha-1)x}{\cos x}. \tag{10}$$

For $\alpha = 1, f(z_t; 1, 1, 0) = \frac{1}{\pi(1+z_t^2)}$. As mentioned in Nolan (1997), properties of the function given in Eq. (9) make the numerical integration feasible, as the function g is continuous and positive, strictly increasing from 0 to ∞ for $\alpha < 1$ and strictly decreasing from ∞ to 0 for $\alpha < 1$. Thus $g(\cdot)\exp\{-g(\cdot)\}$ has a unique maximum attained at x_1 satisfying $g(x_1; \alpha, z_t) = 1$. Nolan kindly provides on his webpage a useful Fortran package, called STABLE to compute ML estimates of the parameters of a general stable distribution. Although we did not directly use this package, it extensively inspired us in writing our procedures. Matsui and Takemura (2006) provides improvements to Nolan's (1997) approach that help to estimate the parameters of the symmetric stable distribution at the boundary cases, i.e., when the underlying random variable approaches zero or ∞ and α is near the value 1 or 2. Thus, when $x \to 0$ or $x \to \infty$ and $\alpha \neq 1$, they derive specific expressions of $f(z_t; \alpha, 1, 0)$ based on asymptotic expansions as stated in Sections 2.4 and 2.5 of Zolotarev (1986), while for the case $\alpha = 1$ and $\alpha = 2$, they use Taylor expansions of $f(z_t; \alpha, 1, 0)$ around these values by giving specific expressions for the partial derivatives of the function with respect to α.

Both procedures of Nolan (1997) and Matsui and Takemura (2006) are derived for "homoscedastic" random variables. For our purposes, we adapt these procedures to incorporate the conditional scale by means of the (T)GARCH specifications, as described in Section 3. In particular, given the range of the α-parameter values of practical interest in financial applications (between 1.7 and 2, e.g. Mandelbrot (1967)), we choose appropriately the numerical methods to compute the density f; for "small" and "large" values of the random variable, we adopt appropriate series expansions, while for "intermediate" values, we perform numerical integration of Eq. (9) with Gaussian quadrature, with 64 points. The variance–covariance matrix of the ML estimates is obtained by numerical computation of the second order derivatives of the log-likelihood. The advantage of ML is that it provides efficient estimates, however at the cost of some difficulties in the computational implementation.

Indirect inference. As an alternative to ML, we apply the indirect inference estimation method introduced by Gouriéroux et al. (1993), which is a simulation-based technique suitable to solve difficult or intractable estimation problems. This method has already proved to be a valuable candidate for the estimation of the parameters of the stable distribution in 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. Moreover, for identification purposes, the dimension of the parameter vector of the auxiliary model should be equal or larger than the dimension of the parameter vector of the true model.

Thus IndInf uses 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 $\Psi \in \mathbb{R}^q$, with $q \geqslant r$. 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:

$$\hat{\psi} = \arg\max_{\psi} \mathcal{L}^a(y_1, y_2, \ldots, y_T; \psi). \tag{11}$$

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

$$\widehat{\psi_{ST}}(\theta) = \arg\max_{\psi} \frac{1}{S}\sum_{s=1}^{S} \mathcal{L}^a(y_1^s(\theta), y_2^s(\theta), \ldots, y_T^s(\theta); \psi). \tag{12}$$

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

$$\hat{\theta}(\Omega) = \arg\min_{\theta} [\hat{\psi} - \widehat{\psi_{ST}}(\theta)]' \, \Omega \, [\hat{\psi} - \widehat{\psi_{ST}}(\theta)], \tag{13}$$

where Ω is a weighting matrix, which is symmetric nonnegative definite and defines the metric. Denote $p(\theta) = \lim_{T\to\infty} \widehat{\psi_{ST}}(\theta)$ 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 $p(\theta) = \psi$. 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 $T \to \infty$.

When the problem is just identified, i.e. the dimension of the two parameter vectors is equal, $r = q$, the results are independent of the choice of the matrices that define the metric, Ω. On the contrary, when $q > r$, it would be necessary to choose a metric Ω to measure the distance between $\hat{\psi}$ and $\widehat{\psi_{ST}}(\theta)$. The optimal choice of Ω is

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

where $J(\psi_0)$ 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. The corresponding IndInf estimator is denoted by $\hat{\theta}^*$.

Alternatively, following Gallant and Tauchen (1996), one can consider directly the score of the auxiliary model and 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:

$$\tilde{\theta}(\Sigma) = \arg\min_{\theta} \frac{\partial \mathcal{L}_T^{a,s}}{\partial \psi'}(\theta, \hat{\psi}) \Sigma \frac{\partial \mathcal{L}_T^{a,s}}{\partial \psi}(\theta, \hat{\psi}), \tag{14}$$

where $\mathcal{L}_T^{a,s}(\theta, \psi) \equiv \frac{1}{S}\sum_{s=1}^{S} \frac{1}{T}\mathcal{L}^a(y_1^s(\theta), y_2^s(\theta), \ldots, y_T^s(\theta); \psi)$ and Σ is a weighting matrix, which is symmetric nonnegative definite and defines the metric. Gouriéroux et al. (1993) show that the two family of estimators, $\hat{\theta}(\Omega)$ and $\tilde{\theta}(\Sigma)$ are asymptotically equivalent. The optimal value of Σ, namely Σ^* is $I(\psi_0)^{-1}$. Provided that a closed form for the gradient of the auxiliary model is available, this approach has an important computational advantage: it avoids the need of repeating the numerical optimization in Eq. (12). This is the reason why we chose to implement it in all our empirical exercises.

As derived in Gouriéroux et al. (1993), the variance–covariance matrix of the IndInf estimator from Eq. (13) is given in the optimal case by

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

and of the estimator from Eq. (14) is given in the optimal case by

$$W(S, \Sigma^*) = \left(1 + \frac{1}{S}\right)\left[\frac{\partial^2 \mathcal{L}_\infty^{a,s}}{\partial \psi \partial \theta'}(\theta_{\mathbf{0}}, \psi_{\mathbf{0}}) I(\psi_0)^{-1} \frac{\partial^2 \mathcal{L}_\infty^{a,s}}{\partial \theta \partial \psi'}(\theta_{\mathbf{0}}, \psi_{\mathbf{0}})\right]^{-1}, \tag{16}$$

where $\mathcal{L}_\infty^{a,s}(\theta, \psi) = \lim_{T\to\infty} \mathcal{L}_T^{a,s}(\theta, \psi)$. Gouriéroux et al. (1993) show that the two estimators are equivalent and denote them generically by W_S^*.

Consistent estimates of W_S^* can be obtained by numerical derivation of $p(\theta)$ with respect to θ and evaluated at $\hat{\theta}^*$ or by numerical derivation of $\mathcal{L}_\infty^{a,s}(\theta, \psi)$ with respect to θ and ψ and evaluated at $\hat{\theta}^*$ and $\hat{\psi}$, where $p(\theta)$ and $\mathcal{L}_\infty^{a,s}$ are numerically computed from simulated data and by replacing $I(\psi_0)$ and $J(\psi_0)$ by their empirical counterparts evaluated at $\hat{\psi}$.

As already mentioned above, for our purposes, we implement the IndInf method by considering as auxiliary models (T)GARCH approaches with Student's t innovations. The choice of Student's 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's t distribution in terms of $\eta = \nu^{-1}$, which is the reciprocal of the degrees of freedom ν. Thus, the auxiliary model for estimating the model given in Eqs. (5) and (6) is a GARCH(1, 1) model with Student's t distributed innovations:

$$Y_t = c_a + \xi_t, \quad \xi_t = u_t\sqrt{h_t}, \tag{17}$$

$$h_t = \omega_a + a_a\xi_{t-1}^2 + b_a h_{t-1} \tag{18}$$

with $\omega_a > 0$, $a_a \geqslant 0$, $b_a \geqslant 0$ and u_t is identically and independently distributed as a symmetric Student's t variable, $u_t \sim t_{1/\eta}$. Similar to Eq. (5), we set c_a to 0. Thus the parameter vector ψ is given by $\psi = (\eta, \omega_a, a_a, b_a)'$ and it has the same dimension 4×1 ($q = 4$) as the parameter vector of the model given in Eqs. (5) and (6), namely $\theta = (\alpha, \omega, a, b)'$. 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 Σ, respectively Σ^* by the identity matrix, I_4.

The auxiliary model for estimating the model given in Eqs. (7) and (8) is a TGARCH(1, 1) model with Student's t distributed innovations:

$$Y_t = c_a + \xi_t, \quad \xi_t = u_t\sqrt{h_t}, \tag{19}$$

$$h_t = \omega_a + a_a\xi_{t-1}^2 + \gamma_a\xi_{t-1}^2\mathbf{1}_{[\xi_{t-1}<0]} + b_a h_{t-1} \tag{20}$$

with $\omega_a > 0$, $a_a \geqslant 0$, $b_a \geqslant 0$ and u_t is identically and independently distributed as a symmetric Student's t variable, $u_t \sim t_{1/\eta}$. Similar to Eq. (17), we set c_a to 0. Thus the parameter vector ψ is given by $\psi = (\eta, \omega_a, a_a, b_a, \gamma_a)'$ and it has the same dimension 5×1 ($q = 5$) as the parameter vector of the model given in Eqs. (7) and (8), namely $\theta = (\alpha, \omega, a, b, \gamma)'$. Thus, similar to the GARCH case, in the IndInf optimization routine we replace the metric Σ, respectively Σ^* by the identity matrix, I_5.

As a result, between the true and auxiliary model there is a rather "natural" correspondence between the parameters (same number of parameters; just identified case): besides the correspondence between the tail-thickness parameters mentioned above, there is a direct correspondence between the parameters of the conditional heteroscedasticity models: ω, a and b (γ) are the (T)GARCH parameters of the true model, while ω_a, a_a and b_a (γ_a) are the (T)GARCH parameters of the auxiliary model.

For both models, we simply minimize the quadratic form in Eq. (14). In the just-identified case, when the minimum is in the interior of the parameter space, the value of the function at the minimum should be zero. Thus, a great computational benefit (at least in terms of speed) is obtained if we directly compute the estimator of θ as the solution of the equations

Table 1
Monte Carlo results for estimating the GARCH(1, 1) model with symmetric stable innovations as given in Eqs. (5) *and* (6) *by indirect inference*: average estimates and standard errors (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			
ω	a	b	α	ω	a	b	α	ω_a	a_a	b_a	η
			1.80	.010	.200	.780	1.798	.007	.145	.779	.235
				(.0012)	(.0101)	(.0083)	(.0168)	(.0009)	(.0074)	(.0079)	(.0122)
			1.85	.010	.200	.780	1.849	.008	.153	.779	.196
				(.0012)	(.0102)	(.0087)	(.0161)	(.0009)	(.0078)	(.0083)	(.0123)
.01	.20	.78	1.90	.010	.200	.779	1.899	.008	.161	.779	.153
				(.0011)	(.0103)	(.0091)	(.0149)	(.0009)	(.0084)	(.0088)	(.0127)
			1.95	.010	.200	.779	1.949	.009	.172	.779	.102
				(.0011)	(.0104)	(.0097)	(.0126)	(.0010)	(.0091)	(.0093)	(.0137)
			1.98	.010	.200	.779	1.980	.009	.182	.779	.060
				(.0011)	(.0104)	(.0100)	(.0094)	(.0010)	(.0100)	(.0095)	(.0155)
			1.80	.010	.100	.880	1.798	.008	.073	.879	.235
				(.0017)	(.0059)	(.0056)	(.0167)	(.0012)	(.0042)	(.0054)	(.0122)
			1.85	.010	.100	.880	1.849	.008	.076	.879	.196
				(.0015)	(.0061)	(.0061)	(.0161)	(.0011)	(.0046)	(.0058)	(.0123)
.01	.10	.88	1.90	.010	.100	.880	1.899	.008	.080	.879	.153
				(.0014)	(.0064)	(.0066)	(.0149)	(.0011)	(.0051)	(.0063)	(.0127)
			1.95	.010	.100	.879	1.949	.009	.086	.879	.102
				(.0014)	(.0067)	(.0072)	(.0126)	(.0012)	(.0057)	(.0070)	(.0137)
			1.98	.010	.100	.879	1.980	.009	.091	.879	.060
				(.0014)	(.0069)	(.0077)	(.0094)	(.0013)	(.0064)	(.0074)	(.0156)
			1.80	.010	.050	.930	1.798	.008	.036	.930	.235
				(.0015)	(.0035)	(.0040)	(.0167)	(.0011)	(.0024)	(.0038)	(.0121)
			1.85	.010	.050	.930	1.849	.008	.038	.930	.196
				(.0014)	(.0037)	(.0045)	(.0161)	(.0010)	(.0028)	(.0043)	(.0123)
.01	.05	.93	1.90	.010	.050	.930	1.899	.008	.040	.929	.153
				(.0014)	(.0041)	(.0051)	(.0149)	(.0011)	(.0032)	(.0049)	(.0127)
			1.95	.010	.050	.930	1.949	.009	.043	.929	.102
				(.0016)	(.0045)	(.0061)	(.0126)	(.0013)	(.0038)	(.0058)	(.0137)
			1.98	.010	.050	.929	1.980	.009	.045	.929	.060
				(.0018)	(.0049)	(.0069)	(.0094)	(.0016)	(.0044)	(.0066)	(.0156)

system $\frac{\partial \mathcal{L}_T^{a,s}}{\partial \psi}(\theta, \hat{\psi}) = 0$ (q equations), which is the empirical counterpart of $\frac{\partial \mathcal{L}_\infty^{a,s}}{\partial \psi}(\theta, \hat{\psi})$. The numerical solution of such a system can be obtained using the Newton–Raphson method, with numerical computation of the Jacobian matrix $\frac{\partial^2 \mathcal{L}_T^{a,s}}{\partial \psi \partial \theta'}$. Non-singularity of the Jacobian matrix ensures the one-to-one correspondence between θ and ψ parameters, at least in some neighborhood of the solution. The same "numerical" Jacobian will be used in the estimation of the variance–covariance matrix, as requested by Eq. (16). Although the IndInf method involves extensive simulation exercises, it is easier to implement than ML, which involves heavy numerical integrations.

5. Monte Carlo study

A detailed set of Monte Carlo experiments is performed to check the reliability of the ML and IndInf method when applied to (T)GARCH(1, 1) models with symmetric α-stable noise.

We adopt a moderately large length of the time series in all experiments ($T = 10000$, roughly comparable with the length of the daily series in the empirical application described in Section 6). As a multiplicative length-factor to produce simulated series, we take $S = 10$: thus 100,000 is in all experiments the length of the simulated series to be handled by the auxiliary model. Each set of simulation results is obtained with $R = 1000$ Monte Carlo replications. In all simulations we use the same random numbers, however, with different (T)GARCH parameterizations.

We chose the values of the parameters to mimic real-case values (the only exception being the ω parameter, which is chosen to be larger, namely 0.01). Thus to generate symmetric stable GARCH(1, 1) processes, we have three cases with b ranging from 0.78 to 0.93 and a ranging from 0.05 to 0.2. To generate symmetric stable TGARCH(1, 1) processes, we have also three cases with b ranging from 0.78 to 0.93 and a and γ ranging from 0.025 to 0.1. As far as the tail-thickness parameter α is concerned, we experiment with five different values, ranging from a "close to Gaussian" value (1.98) to a moderate "fat-tail" value (1.80).

Tables 1–4 report the results of the ML and IndInf estimation based on simulated data. More precisely, Tables 1 and 2 report the results from estimating the GARCH(1, 1) model with symmetric stable innovations by IndInf and, respectively, ML, while Tables 3 and 4 present the results for the TGARCH(1, 1) model with symmetric stable innovations.

Table 2
Monte Carlo results for estimating the GARCH(1, 1) model with symmetric stable innovations as given in Eqs. (5) *and* (6) *by maximum likelihood*: average estimates and standard errors (in parentheses) over $R = 1000$ Monte Carlo replications, based on $T = 10000$ number of observations and $S = 10$ number of simulation paths.

True parameters				Estimated parameters			
ω	a	b	α	ω	a	b	α
.01	.20	.78	1.80	.009 (.0011)	.185 (.0087)	.780 (.0078)	1.801 (.0140)
			1.85	.010 (.0010)	.189 (.0089)	.780 (.0082)	1.851 (.0130)
			1.90	.010 (.0010)	.193 (.0091)	.780 (.0086)	1.900 (.0114)
			1.95	.010 (.0010)	.196 (.0094)	.780 (.0091)	1.950 (.0091)
			1.98	.010 (.0010)	.198 (.0096)	.780 (.0095)	1.982 (.0062)
.01	.10	.88	1.80	.010 (.0015)	.093 (.0049)	.880 (.0051)	1.801 (.0140)
			1.85	.010 (.0013)	.095 (.0052)	.880 (.0055)	1.851 (.0130)
			1.90	.010 (.0013)	.096 (.0055)	.880 (.0060)	1.900 (.0114)
			1.95	.010 (.0013)	.098 (.0059)	.880 (.0066)	1.950 (.0090)
			1.98	.010 (.0014)	.099 (.0062)	.880 (.0072)	1.982 (.0062)
.01	.05	.93	1.80	.010 (.0013)	.046 (.0029)	.930 (.0037)	1.801 (.0140)
			1.85	.010 (.0013)	.047 (.0032)	.930 (.0041)	1.851 (.0129)
			1.90	.010 (.0013)	.048 (.0035)	.930 (.0047)	1.900 (.0114)
			1.95	.010 (.0015)	.049 (.0040)	.930 (.0055)	1.950 (.0090)
			1.98	.010 (.0017)	.050 (.0044)	.930 (.0063)	1.982 (.0062)

We verify numerically the strict stationarity conditions described in Section 3 and find that for all parameter combinations given in the tables, the top-Lyapunov conditions are negative. However, in a very few cases (5 of the 30 cases) the estimation results for the simulated data contain 1 up to 3 outliers that affect the mean and variances of the estimates of ω, but not of a, b and γ. These very few outliers (which we discard in the results presented in the tables) are due to a behavior of the simulated series similar to a "non-stationary" case. According to Francq and Zakoian (2012), in this case, the PML estimator of ω is inconsistent, however the estimators of a, b and γ remain consistent. Our results show that the IndInf estimators follow the same pattern. However, the medians of the estimates computed on the "non-discarded" results, which are better measures in the presence of outliers and which can be obtained from the authors upon request, are very close to the true values of the parameters. Moreover, the interquartile ranges exhibit the same behavior as the standard deviations reported in the tables.

Before commenting on the results reported in the tables, we need to point out the remarkable speed in convergence of the IndInf method compared to ML: IndInf provides estimation results around ten times faster than the ML. Thus, for instance, to estimate the model in Eqs. (5) and (6) on a computer with a processor Intel i7, 2.67 Ghz, for $\alpha = 1.8$, $a = 0.05$ and $b = 0.93$, for each Monte Carlo replication, IndInf converges in 0.43 s, while ML needs 4.5 s to converge, while, in order to estimate the model in Eqs. (7) and (8), for $\alpha = 1.8$, $a = 0.025$, $\gamma = 0.025$ and $b = 0.93$, IndInf converges in 0.53 s, while ML needs 5.5 s per replication. This is an additional computational advantage of IndInf over ML, besides the greater implementation easiness.

Regarding the statistical performance of the estimation methods, one may say that they are quite remarkable. With very few exceptions, estimates of the model of interest (true model) stemming from both ML and IndInf approaches "seem unbiased" (differences between the average estimates and the parameters used to generate the data are observable only after the third digit). Moreover, the expected gains in efficiency provided by ML compared to IndInf are minimal in most cases: the differences between the empirical standard deviations presented in parentheses are only observable after three digits. Concerning the stability parameter, one may notice that some entries of the tables are similar (or equal) across the GARCH parameterization. Differences would appear only if more significant digits were displayed.

The variance of ω, a, b and γ estimated by IndInf are always larger than their counterparts in the auxiliary model, but the difference is not very large. Moreover, the estimate of b is nearly unbiased also in the auxiliary model (see b_a); on the

Table 3
Monte Carlo results for estimating the TGARCH(1, 1) model with symmetric stable innovations as given in Eqs. (7) *and* (8) *by indirect inference*: average estimates and standard errors (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				
ω	a	b	γ	α	ω	a	b	γ	α	ω_a	a_a	b_a	γ_a	η
				1.80	.010	.100	.780	.101	1.799	.007	.073	.779	.072	.235
					(.0009)	(.0114)	(.0109)	(.0266)	(.0169)	(.0006)	(.0067)	(.0097)	(.0107)	(.0121)
				1.85	.010	.100	.780	.101	1.849	.008	.076	.779	.076	.196
					(.0009)	(.0116)	(.0116)	(.0264)	(.0162)	(.0007)	(.0072)	(.0105)	(.0113)	(.0123)
.01	.10	.78	.10	1.90	.010	.100	.779	.101	1.899	.008	.081	.779	.080	.153
					(.0010)	(.0118)	(.0126)	(.0260)	(.0150)	(.0007)	(.0079)	(.0114)	(.0120)	(.0127)
				1.95	.010	.100	.779	.101	1.950	.009	.086	.779	.086	.102
					(.0010)	(.0120)	(.0139)	(.0254)	(.0126)	(.0008)	(.0087)	(.0125)	(.0129)	(.0137)
				1.98	.010	.100	.779	.101	1.980	.009	.091	.779	.091	.060
					(.0011)	(.0121)	(.0147)	(.0249)	(.0094)	(.0010)	(.0094)	(.0132)	(.0138)	(.0156)
				1.80	.010	.050	.880	.051	1.799	.007	.036	.880	.036	.235
					(.0010)	(.0080)	(.0077)	(.0207)	(.0170)	(.0007)	(.0039)	(.0065)	(.0062)	(.0121)
				1.85	.010	.050	.880	.050	1.849	.008	.038	.879	.038	.196
					(.0011)	(.0084)	(.0084)	(.0211)	(.0162)	(.0008)	(.0043)	(.0072)	(.0067)	(.0123)
.01	.05	.88	.05	1.90	.010	.050	.880	.051	1.899	.008	.040	.879	.040	.153
					(.0012)	(.0089)	(.0094)	(.0214)	(.0150)	(.0009)	(.0049)	(.0081)	(.0074)	(.0127)
				1.95	.010	.050	.880	.051	1.950	.009	.043	.879	.043	.102
					(.0013)	(.0093)	(.0108)	(.0216)	(.0126)	(.0011)	(.0057)	(.0092)	(.0082)	(.0137)
				1.98	.010	.050	.879	.051	1.980	.009	.046	.879	.045	.060
					(.0015)	(.0096)	(.0120)	(.0215)	(.0094)	(.0012)	(.0063)	(.0101)	(.0089)	(.0156)
				1.80	.010	.025	.930	.025	1.799	.007	.018	.930	.018	.235
					(.0011)	(.0056)	(.0060)	(.0160)	(.0170)	(.0008)	(.0023)	(.0045)	(.0036)	(.0121)
				1.85	.010	.025	.930	.025	1.849	.008	.019	.930	.019	.196
					(.0012)	(.0061)	(.0068)	(.0168)	(.0163)	(.0009)	(.0026)	(.0052)	(.0040)	(.0123)
.01	.025	.93	.025	1.90	.010	.025	.930	.025	1.899	.008	.020	.929	.020	.153
					(.0015)	(.0067)	(.0079)	(.0178)	(.0150)	(.0010)	(.0031)	(.0061)	(.0046)	(.0127)
				1.95	.010	.025	.929	.024	1.950	.009	.022	.929	.021	.102
					(.0046)	(.0200)	(.0295)	(.0369)	(.0129)	(.0013)	(.0038)	(.0075)	(.0054)	(.0137)
				1.98	.010	.025	.929	.025	1.979	.009	.023	.929	.023	.060
					(.0032)	(.0257)	(.0329)	(.0478)	(.0185)	(.0016)	(.0045)	(.0088)	(.0062)	(.0156)

contrary, $\widehat{\omega}$, $\widehat{a}$ and $\widehat{\gamma}$ are always remarkably downward biased in the auxiliary model (see ω_a, a_a and γ_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: the larger the α is, the smaller the η is, which indicates a large value for the Student's t degrees of freedom ν.

We also try other combinations of parameters: for instance, larger values of a and b as well as smaller values of α (till 1.70), implying even thicker tails. Not all these additional experiments are successful, as the combination of GARCH and thick-tail noise in several cases leads to "exploding" values of the simulated series. This is 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 are not realistic, when compared with the values estimated from real series, thus they are not reported here for the sake of brevity. However, they can be obtained on request from the authors. The main difference from the results presented above regards the consistency of ω and a: both IndInf and ML methods provide inconsistent estimators of ω, as the series mimic a non-stationary behavior. Additionally, the ML estimators of parameter a seem also not very accurate, regardless of the choice of $\alpha < 1.80$. This is not the case of IndInf that provides accurate results for all choices of α.

We also estimate series of lower length such as $T = 1000$. For the sake of brevity, we choose not to report the results here. However, they can be obtained on request from the authors. Besides a lower accuracy in the estimation, the results follow the same pattern as the one described above. The only exception is the estimation of ω, which seems to be sometimes inconsistent. This may be due to the fact that the results have some outliers stemming from simulated data that mimic the behavior of non-stationary series. This occurs mainly for $\alpha < 1.80$ with the GARCH parameter values we adopt in our Monte Carlo experiments presented above. Lowering the value of α increases the frequency of outliers: for example, for the GARCH parameters $a = 0.10$ and $b = 0.88$, the top Lyapunov coefficients are still negative, but very close to zero. However, when α approaches 1.7, the top Lyapunov coefficient remains very close to zero, but becomes positive. As mentioned above, for these "non-stationary" cases, Francq and Zakoian (2012) show that the PML estimators of a, b and γ are consistent, while the estimator of ω is not. Our results show further that the IndInf estimators follow the same pattern. However, the medians of the ω estimates are all very close to the true values and the interquantile ranges follow the same patterns as the ones of the standard deviations reported in the paper.

Table 4
Monte Carlo results for estimating the TGARCH(1, 1) model with symmetric stable innovations as given in Eqs. (7) *and* (8) *by maximum likelihood*: average estimates and standard errors (in parentheses) over $R = 1000$ Monte Carlo replications, based on $T = 10000$ number of observations and $S = 10$ number of simulation paths.

True parameters					Estimated parameters				
ω	a	b	γ	α	ω	a	b	γ	α
.01	.10	.78	.10	1.80	.009 (.0008)	.092 (.0083)	.780 (.0100)	.093 (.0131)	1.801 (.0140)
				1.85	.010 (.0008)	.094 (.0087)	.780 (.0107)	.095 (.0135)	1.851 (.0129)
				1.90	.010 (.0009)	.096 (.0092)	.780 (.0116)	.097 (.0138)	1.900 (.0114)
				1.95	.010 (.0009)	.098 (.0096)	.779 (.0128)	.099 (.0141)	1.950 (.0090)
				1.98	.010 (.0010)	.099 (.0100)	.779 (.0136)	.100 (.0140)	1.982 (.0062)
.01	.05	.88	.05	1.80	.009 (.0009)	.046 (.0048)	.880 (.0064)	.047 (.0077)	1.801 (.0140)
				1.85	.010 (.0010)	.047 (.0053)	.880 (.0070)	.047 (.0081)	1.851 (.0129)
				1.90	.010 (.0010)	.048 (.0058)	.880 (.0079)	.048 (.0086)	1.900 (.0114)
				1.95	.010 (.0012)	.049 (.0064)	.879 (.0092)	.049 (.0091)	1.950 (.0090)
				1.98	.010 (.0014)	.050 (.0068)	.879 (.0103)	.050 (.0094)	1.982 (.0062)
.01	.025	.93	.025	1.80	.009 (.0010)	.023 (.0028)	.930 (.0045)	.023 (.0046)	1.801 (.0140)
				1.85	.010 (.0011)	.024 (.0032)	.930 (.0051)	.024 (.0050)	1.851 (.0129)
				1.90	.010 (.0012)	.024 (.0037)	.930 (.0059)	.024 (.0055)	1.900 (.0114)
				1.95	.010 (.0015)	.024 (.0043)	.929 (.0074)	.025 (.0062)	1.950 (.0090)
				1.98	.010 (.0019)	.025 (.0049)	.929 (.0089)	.025 (.0067)	1.982 (.0062)

6. Empirical application

The (T)GARCH(1, 1) models with symmetric α-stable noise defined in Section 3 are 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 February 14th, 2013 (753, 3275, and 16,377 observations for the monthly, weekly and daily series, respectively);
2. Standard & Poor's 500 (SP500) index from January 1st, 1964 to February 14th, 2013 (589, 2563, and 12,817 observations, respectively):
3. IBM share's prices from January 1st, 1973 to February 14th, 2013 (481, 2093, and 10,467 observations, respectively) and
4. GE share's prices from January 1st, 1973 to February 14th, 2013 (481, 2093, and 10,467 observations, respectively).

Data have been obtained from Thomson Reuters Datastream.

Indirect estimation is performed using a large value of the multiplicative length-factor for the simulated series ($S = 100$). Convergence is always achieved inside the parameter space of the auxiliary model. If in the Monte Carlo exercise the choice of the initial values for the iterative procedure in both estimation methods was unproblematic, as we knew the true values of the parameters, in the real data application, this is no longer the case. However, the close correspondence between the (T)GARCH parameters of the true and auxiliary model within the IndInf method provides great help in this direction. Thus, we choose as initial values in both methods the parameter values of the auxiliary model estimated in the IndInf procedure. The only "additional" problem is the choice of a "good" initial value for the α parameter. Despite this problem, the convergence time of the two estimation methods follows the same pattern as in the simulation exercise: e.g., on the same computer with an Intel i7, 2.67 Ghz processor, ML applied on the series of daily returns of IBM converges in 4.60 s, while the IndInf approach performed on $S = 10$ simulated series, converges in 0.65 s. When we increase the multiplicative length factor for the simulated series to $S = 100$, the estimation time for the same series within IndInf becomes about 5.80 s. The additional time for computing the variance–covariance matrix is slightly higher for the ML estimates (about 1.80 s) than for IndInf estimates (about 1.50 s).

Tables 5–8 report the estimated values of the parameters of our models, as well as the standard deviations (in parentheses) of the estimated parameters of the model of interest, as given by Eq. (16) for IndInf and by the inverted Hessian

Table 5
Empirical results for estimating the GARCH(1, 1) model with symmetric stable innovations as given in Eqs. (5) *and* (6) *by indirect inference*: 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	a	b	α	ω_a	a_a	b_a	η
DJIA	*M*	.193e−03	.110	.763	1.906	.172e−03	.093	.748	.144
		(.754e−04)	(.0344)	(.0675)	(.0488)				
	W	.103e−04	.073	.894	1.941	.901e−05	.061	.894	.112
		(.214e−05)	(.0095)	(.0125)	(.0199)				
	D	.558e−06	.053	.932	1.897	.448e−06	.042	.932	.155
		(.599e−07)	(.0029)	(.0034)	(.0097)				
SP500	*M*	.836e−04	.102	.842	1.918	.631e−04	.083	.842	.146
		(.359e−04)	(.0272)	(.0354)	(.0465)				
	W	.927e−05	.097	.877	1.947	.742e−05	.081	.879	.110
		(.235e−05)	(.0140)	(.0155)	(.0238)				
	D	.308e−06	.054	.937	1.896	.226e−06	.043	.937	.156
		(.490e−07)	(.0028)	(.0029)	(.0103)				
IBM	*M*	.382e−03	.097	.816	1.945	.350e−03	.084	.806	.111
		(.189e−03)	(.0332)	(.0608)	(.0527)				
	W	.152e−04	.041	.936	1.870	.126e−04	.031	.937	.178
		(.422e−05)	(.0075)	(.0100)	(.0292)				
	D	.125e−05	.034	.952	1.849	.981e−06	.026	.953	.195
		(.176e−06)	(.0023)	(.0029)	(.0131)				
GE	*M*	.182e−03	.128	.830	1.961	.178e−03	.116	.821	.088
		(.890e−04)	(.0341)	(.0426)	(.0484)				
	W	.150e−04	.060	.921	1.937	.135e−04	.050	.921	.116
		(.470e−05)	(.0107)	(.0128)	(.0244)				
	D	.126e−05	.042	.947	1.904	.106e−05	.034	.946	.147
		(.194e−06)	(.0029)	(.0034)	(.0120)				

Table 6
Empirical results for estimating the GARCH(1, 1) model with symmetric stable innovations as given in Eqs. (5) *and* (6) *by maximum likelihood*: standard deviations are reported in parentheses. *M* stands for monthly, *W* stands for weekly and *D* stands for daily.

Data		Estimated parameters			
		ω	a	b	α
DJIA	*M*	.228e−3	.112	.731	1.892
		(.932e−4)	(.0369)	(.0809)	(.0444)
	W	.105e−4	.070	.893	1.925
		(.249e−5)	(.0102)	(.0142)	(.0195)
	D	.515e−6	.049	.934	1.878
		(.739e−7)	(.0033)	(.0041)	(.0111)
SP500	*M*	.914e−4	.098	.830	1.886
		(.475e−4)	(.0283)	(.0457)	(.0542)
	W	.105e−4	.096	.871	1.938
		(.281e−5)	(.0137)	(.0170)	(.0194)
	D	.277e−6	.052	.936	1.877
		(.638e−7)	(.0036)	(.0041)	(.0131)
IBM	*M*	.361e−3	.093	.814	1.888
		(.201e−3)	(.0368)	(.0726)	(.0567)
	W	.143e−4	.038	.936	1.836
		(.426e−5)	(.0066)	(.0101)	(.0298)
	D	.116e−5	.032	.952	1.812
		(.243e−6)	(.0035)	(.0051)	(.0144)
GE	*M*	.199e−3	.133	.816	1.944
		(.935e−4)	(.0346)	(.0389)	(.0436)
	W	.153e−4	.055	.924	1.918
		(.484e−5)	(.0090)	(.0116)	(.0247)
	D	.123e−5	.041	.946	1.869
		(.276e−6)	(.0056)	(.0069)	(.0135)

of the log-likelihood for ML. 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. Thus, Tables 5 and 6 report the results from estimating the GARCH(1, 1) model with symmetric stable

Table 7
Empirical results for estimating the TGARCH(1, 1) model with symmetric stable innovations as given in Eqs. (7) *and* (8) *by indirect inference*: 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	a	b	γ	α	ω_a	a_a	b_a	γ_a	η
DJIA	*M*	.299e−03	.019	.686	.211	1.923	.256e−03	.021	.685	.150	.129
		(.101e−03)	(.0275)	(.0868)	(.0824)	(.0455)					
	W	.162e−04	.013	.862	.162	1.955	.124e−04	.021	.879	.088	.103
		(.285e−05)	(.0124)	(.0155)	(.0358)	(.0186)					
	D	.761e−06	.015	.922	.092	1.912	.575e−06	.017	.929	.054	.145
		(.718e−07)	(.0038)	(.0042)	(.0122)	(.0095)					
SP500	*M*	.132e−03	.002	.817	.181	1.921	.109e−03	.002	.817	.144	.138
		(.552e−04)	(.0069)	(.0568)	(.0709)	(.0460)					
	W	.120e−04	.013	.866	.187	1.962	.889e−05	.020	.883	.111	.096
		(.238e−05)	(.0140)	(.0162)	(.0444)	(.0223)					
	D	.421e−06	.016	.933	.083	1.911	.319e−06	.015	.936	.057	.145
		(.574e−07)	(.0042)	(.0044)	(.0157)	(.0103)					
IBM	*M*	.697e−03	.029	.711	.247	1.960	.568e−03	.039	.732	.142	.097
		(.264e−03)	(.0396)	(.0784)	(.1003)	(.0478)					
	W	.157e−04	.035	.937	.008	1.871	.138e−04	.024	.934	.016	.176
		(.493e−05)	(.0154)	(.0139)	(.0414)	(.0291)					
	D	.153e−05	.014	.944	.057	1.860	.115e−05	.018	.948	.020	.191
		(.206e−06)	(.0068)	(.0034)	(.0165)	(.0140)					
GE	*M*	.298e−03	.041	.804	.178	1.962	.293e−03	.046	.797	.127	.088
		(.136e−03)	(.0378)	(.0558)	(.0939)	(.0482)					
	W	.175e−04	.040	.922	.033	1.941	.172e−04	.030	.917	.042	.110
		(.629e−05)	(.0150)	(.0183)	(.0441)	(.0240)					
	D	.154e−05	.014	.942	.067	1.915	.122e−05	.021	.945	.028	.143
		(.217e−06)	(.0081)	(.0036)	(.0179)	(.0123)					

Table 8
Empirical results for estimating the TGARCH(1, 1) model with symmetric stable innovations as given in Eqs. (7) *and* (8) *by maximum likelihood*: standard deviations are reported in parentheses. *M* stands for monthly, *W* stands for weekly and *D* stands for daily.

Data		Estimated parameters				
		ω	a	b	γ	α
DJIA	*M*	.297e−3	.021	.694	.171	1.917
		(.885e−4)	(.0365)	(.0722)	(.0677)	(.0428)
	W	.144e−4	.023	.878	.102	1.940
		(.302e−5)	(.0094)	(.0153)	(.0184)	(.0178)
	D	.663e−6	.019	.930	.064	1.891
		(.786e−7)	(.0029)	(.0041)	(.0054)	(.0107)
SP500	*M*	.130e−3	.000	.815	.181	1.914
		(.504e−4)	(.0000)	(.0451)	(.0485)	(.0473)
	W	.107e−4	.024	.880	.124	1.948
		(.272e−5)	(.0114)	(.0160)	(.0202)	(.0179)
	D	.381e−6	.018	.936	.069	1.897
		(.635e−7)	(.0031)	(.0041)	(.0059)	(.0123)
IBM	*M*	.621e−3	.044	.733	.161	1.907
		(.333e−3)	(.0357)	(.0978)	(.0995)	(.0568)
	W	.157e−4	.031	.934	.016	1.838
		(.453e−5)	(.0080)	(.0105)	(.0126)	(.0298)
	D	.138e−5	.023	.948	.023	1.816
		(.278e−6)	(.0033)	(.0056)	(.0050)	(.0143)
GE	*M*	.314e−3	.056	.793	.156	1.963
		(.136e−3)	(.0417)	(.0453)	(.0714)	(.0437)
	W	.199e−4	.034	.919	.042	1.926
		(.537e−5)	(.0105)	(.0117)	(.0165)	(.0242)
	D	.142e−5	.026	.944	.031	1.873
		(.267e−6)	(.0038)	(.0049)	(.0058)	(.0134)

innovations by IndInf and, respectively, ML, while Tables 7 and 8 presents the results for the TGARCH(1, 1) model with symmetric stable innovations.

One can notice that both methods provide similar results in what regards the estimated parameters. Moreover, we verify numerically the strict stationarity conditions stated in Section 3 and find that all estimated parameters satisfy the negativity condition of the top-Lyapunov exponent. Moreover, regardless of the estimation method, the parameters of the models of interest follow the same behavior as the parameters of the auxiliary model estimated within the IndInf approach: ω, a, γ decrease and b increases when increasing the sampling frequency of data. In the standard (T)GARCH framework, this can be interpreted as empirical evidence of the fact that the frequency of sampling stock returns increases the clustering and the persistence degree of their conditional volatilities. The nonexistence of the second moment within the stable distribution makes such interpretation difficult. However, the fact that the parameters behave similar to the ones of standard (T)GARCH models confirms the validity of our estimation results.

For a financial analyst, it can be interesting to observe that the value of the α parameter increases as a result of aggregation. Time aggregation increases the estimated value of α: we observe within each group of the considered financial assets that the estimated parameter α is higher for weekly and monthly data than for daily data. This result is not surprising, since it is a well known fact that increasing the frequency of sampling stock returns also increases the degree of fat-thickness of their distribution: daily returns exhibit higher and more extreme values than the returns sampled at lower frequency. Regarding the aggregation across stocks, we observe that the value of α estimated for IBM and GE are in general lower than their counterparts estimated on indexes series. This is also 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 and GE.

For a computational econometrician, it can be interesting to observe that b in the model of interest is usually almost equal to its correspondent b_a in the auxiliary model. The close behavior between the true and the auxiliary model is due to 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).

7. Conclusions

In this paper we apply the indirect inference and maximum likelihood methods to estimate the parameters of (T)GARCH models with symmetric stable innovations. As most of the financial returns exhibit conditional heteroscedasticity, leverage effects and fat-tails, there is a common practice among scholars and practitioners to capture these features by means of (T)GARCH models with fat-tailed distributed innovations. A standard choice is to consider Student's t distributed innovations, however, at the cost 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 a large cost 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 maximum likelihood approach developed by Nolan (1997) and Matsui and Takemura (2006) and the indirect inference method introduced by Gouriéroux et al. (1993).

The simulation study reveals very good results for a large pallet of parameter choices: in general, the estimates of the model of interest "seem unbiased" regardless of the estimation method. However, the efficiency gains provided by the maximum likelihood approach comes at some higher computational costs. The indirect inference, which is much easier to implement, is also faster than the maximum likelihood in estimating the parameters of interest. The empirical results from applying the method to twelve series of financial returns sampled at three different frequencies provide further empirical evidence on the performance of the two methods when estimating the parameters of (T)GARCH models with symmetric stable innovations.

Acknowledgments

We are grateful to the editors and the anonymous referees for insightful remarks that have significantly improved the paper. For further helpful comments we would like to thank Marco J. Lombardi, Alessandra Amendola and participants at the CFE ERCIM Conference 2012, Oviendo, Spain. All remaining errors are ours. Giorgio Calzolari gratefully acknowledges financial support of the project MIUR PRIN MISURA - Multivariate models for risk assessment. Roxana Halbleib gratefully acknowledges financial support from the European Social Fund, Ministry of Science, Research and Arts of Baden-Württemberg, Germany and from the Baden-Württemberg Stiftung, Germany through the Eliteprogramm für Postdoktorandinnen und Postdoktoranden.

References

Black, F., 1976. Studies of stock price volatility changes. Proceedings of the American Statistical Association, Business and Economic Statistics Section 177–181.
Bollerslev, T., 1986. Generalized autoregressive conditional heteroskedasticity. Journal of Econometrics 31, 307–327.
Calzolari, G., Fiorentini, G., 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, 532–546.
Chambers, J., Mallows, C., Stuck, B.W., 1976. A method for simulating stable random variables. Journal of the American Statistical Association 71, 340–344.
Chenyao, D., Doganoglu, T., Mittnik, S., Rachev, S., 1999. Maximum likelihood estimation of stable paretian models. Mathematical and Computer Modelling 29 (10–12), 275–293.

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.
Francq, C., Zakoian, J.-M., 2010. GARCH Models, Structure, Statistical Inference and Financial Applications. John Wiley and Sons, New York.
Francq, C., Zakoian, J.-M., 2012. Strict stationarity testing and estimation of explosive and stationary garch models. Econometrica 80, 821–861.
Gallant, A.R., Tauchen, G., 1996. Which moments to match? Econometric Theory 12, 657–681.
Garcia, R., Renault, E., Veredas, D., 2011. Estimation of stable distributions with indirect inference. Journal of Econometrics 161, 325–337.
Ghose, D., 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., Renault, E., 1996. Stochastic volatility. In: Maddala, G., Rao, C.R. (Eds.), Handbook of Statistics, Vol. 14. Elsevier Science Publishers, Amsterdam, pp. 119–192.
Glosten, L.R., Jagannathan, R., Runkle, D.E., 1993. On the relation between the expected value and the volatility of the nominal excess return of stocks. Journal of Finance 48, 1179–1801.
Gouriéroux, C., Monfort, A., Renault, E., 1993. Indirect inference. Journal of Applied Econometrics 8, S85–S118.
Liu, S., Brorsen, B., 1995a. Garch-stable as a model of futures price movements. Review of Quantitative Finance and Accounting 5, 155–167.
Liu, S., Brorsen, B., 1995b. Maximum likelihood estimation of a garch-stable model. Journal of Applied Econometrics 10 (3), 273–285.
Lombardi, M., Calzolari, G., 2008. Indirect estimation of α-stable distributions and processes. Econometrics Journal 11, 193–208.
Lombardi, M., Calzolari, G., 2009. Indirect estimation of α-stable stochastic volatility models. Computational Statistics & Data Analysis 53 (6), 2298–2308.
Lombardi, M., Veredas, D., 2009. Indirect estimation of elliptical stable distributions. Computational Statistics & Data Analysis 53, 2309–2324.
Mandelbrot, B.B., 1967. The variation of some other speculative prices. Journal of Business 40, 393–414.
Matsui, M., Takemura, A., 2006. Some improvements in numerical evaluation of symmetric stable density and its derivatives. Communications in Statistics, Theory Methods 35 (1–3), 149–172.
McCulloch, J., 1985. Interest-risk sensitive deposit insurance premia: stable ach estimates. Journal of Banking and Finance 9, 137–156.
Mittnik, S., Paolella, M., Rachev, S., 2000. Diagnosing and treating the fat tails in financial returns data. Journal of Empirical Finance 7, 389–416.
Mittnik, S., Paolella, M., Rachev, S., 2002. Stationarity of stable power-garch processes. Journal of Econometrics 106, 97–107.
Nelson, D.B., 1990. Stationarity and persistence in the garch(1, 1) model. Econometric Theory 6, 318–334.
Nolan, J.P., 1997. Numerical calculation of stable densities and distribution functions. Communications in Statistics, Stochastic Models 13, 759–774.
Panorska, A., Mittnik, S., Rachev, S., 1995. Stable garch models for financial time series. Applied Mathematics Letters 8, 33–37.
Peters, G.W., Sisson, S.A., Fan, Y., 2012. Likelihood-free Bayesian inference for α-stable models. Computational Statistics & Data Analysis 11, 3743–3756.
Yang, S., Brorsen, B., 1993. Nonlinear dynamics of daily futures prices: conditional heteroskedasticity or chaos? Journal of Futures Markets 13 (2), 175–191.
Zolotarev, V.M., 1986. One-dimensional stable distributions. In: Translation of Mathematical Monographs, vol. 65. American Mathematical Society, Providence (Translation of the Original 1983 Russian).