Economic Modelling 53 (2016) 278–288

Contents lists available at ScienceDirect

Economic Modelling

journal homepage: www.elsevier.com/locate/ecmod

# A discussion on the innovation distribution of the Markov regime-switching GARCH model

Yanlin Shi [a,c,*], Lingbing Feng [b]

[a] Research School of Finance, Actuarial Studies and Statistics, The Australian National University, ACT 2601, Australia
[b] International Institute for Financial Studies, Jiangxi University of Finance and Economics, Nanchang, Jiangxi, China
[c] School of Demography, The Australian National University, ACT 2601, Australia

ARTICLE INFO

*Article history:*
Accepted 27 November 2015
Available online 5 January 2016

*Keywords:*
GARCH model
Regime-switching
Fat-tailed distribution
Tempered stable distribution

ABSTRACT

The Markov Regime-Switching Generalized autoregressive conditional heteroskedastic (MRS-GARCH) model is a widely used approach to model the financial volatility with potential structural breaks. The original innovation of the MRS-GARCH model is assumed to follow the Normal distribution, which cannot accommodate fat-tailed properties commonly existing in financial time series. Many existing studies point out that this problem can lead to inconsistent estimates. To overcome it, the Student's t-distribution and General Error Distribution (GED) are the two most popular alternatives. However, a recent study points out that the Student's t-distribution lacks stability. Also, it incorporates the α-stable distribution in the GARCH-type model. The issue of the α-stable distribution is that its second moment does not exist. To solve this problem, the tempered stable distribution, which retains most characteristics of the α-stable distribution and has defined moments, is a natural candidate. In this paper, we conduct a series of simulation studies to demonstrate that MRS-GARCH model with tempered stable distribution consistently outperform that with Student's t-distribution and GED. Our empirical study on the S&P 500 daily return volatility also generates robust results. Therefore, we argue that the tempered stable distribution could be a widely useful tool for modeling the financial volatility in general contexts with a MRS-GARCH-type specification.



## 1. Introduction

The Generalized autoregressive conditional heteroskedastic (GARCH) model proposed by Bollerslev (1986) can successfully capture many attractive characteristics of financial time series like time-varying heteroskedasticity and volatility persistence (Ho et al., 2013). Therefore, it has become a standard approach to study the financial volatility in the past few decades. Volatility persistence describes how fast the effect of a shock on the volatility will die away. It has been extensively observed and studied in various fields of economics and finance in the past few decades (Narayan and Narayan, 2007, 2011). The analysis of volatility persistence can help researchers understand how the financial series evolves and improve the forecasting quality (Franses and van Dijk, 1996; Narayan and Sharma, 2014; Shi and Ho, 2015b; Westerlund and Narayan, 2012). However, the major problem of the original GARCH model is that it assumes that the conditional volatility has only one regime over the entire period. In practice, due to the potential presence of policy changes and other shocks in the real economy, structural breaks may exist for the financial data collected over a long period (Ang and Timmermann, 2012; Shi and Ho, 2015a). In the presence of structural breaks, Lamoureux and Lastrapes (1990), Hamilton and Susmel (1994) and Klaassen (2002) argue that GARCH model always implies a high volatility persistence.

Hamilton (1988, 1989) proposes the Markov Regime-Switching (MRS) model to allow parameters to transit between state spaces. Based on his works, many studies have been conducted to incorporate the MRS specification into the GARCH framework, which leads to MRS-GARCH models with various structures (Cai, 1994; Dueker, 1997; Gray, 1996; Haas et al., 2004; Hamilton and Susmel, 1994; Klaassen, 2002; Lin, 1998). A main difference between those models is the specific algorithm to solve the path dependency problem of the conditional volatility.

Despite this difference, all the available MRS-GARCH models are constructed based on the assumption that the innovation follows a Normal (Gaussian) distribution. However, significant evidence suggests that the financial time series is rarely Gaussian but typically leptokurtic and exhibits fat-tail behavior (Bollerslev, 1987; Ho et al., 2013; Stanley et al., 2008; Susmel and Engle, 1994). For the original GARCH model, (Bollerslev and Wooldridge, 1992) propose the Quasi Maximum Likelihood Estimation (QMLE), which can accommodate for fat-tailedness through their specification and still generates consistent estimates. For the MRS model, however, if regimes are not Gaussian but leptokurtic, the use of within-regime normality can seriously affect the identification of the regime process (Ardia, 2009; Haas, 2009; Klaassen, 2002).

* Corresponding author. Tel.: +61 2 6125 7318.
*E-mail address:* yanlin.shi@anu.edu.au (Y. Shi).

http://dx.doi.org/10.1016/j.econmod.2015.11.018
0264-9993/

Additionally, Haas and Paolella (2012) further argue that QMLE of the MRS-GARCH model is not consistent. Consequently, the sought of an appropriate distribution to accommodate the excess kurtosis of the financial time series becomes an essential issue for the MRS-GARCH model.

To solve this problem, a common solution is to employ a fat-tailed distribution such as the Student's t-distribution or General Error Distribution (GED) (Haas et al., 2004; Klaassen, 2002; Marcucci, 2005). However, a recent study by Calzolari et al. (2014) argues that the widely used Student's t-distribution is problematic. Its most serious drawback is that this distribution lacks in stability under aggregation, which is of particular importance in portfolio applications and risk management in finance study. As a replacement of the Student's t, the $\alpha$-stable distribution is recommended by Calzolari et al. (2014) to model the fat-tailed behavior of financial time series for the original GARCH model. Additionally, they argue that similar to the Student's t, $\alpha$-stable distribution can be easily adapted to account for many characteristics of volatility such as asymmetry in the underlying financial time series. Unfortunately, since the second moment of the $\alpha$-stable distribution does not exist in most cases, GARCH-type model with this distribution will lead to problematic interpretation. Hence, the sought of an alternative distribution would be of particular interest for the application of GARCH-type model.

The tempered stable distribution is a natural substitution of the $\alpha$-stable one. Firstly introduced[1] in Koponen (1995), tempered stable distribution covers several well-known subclasses like Variance Gamma, bilateral Gamma and CGMY distributions (Küchler and Tappe, 2013). The advantage of it is that this distribution retains most of the attractive properties of the $\alpha$-stable distribution and has defined moments.

In this paper, we employ the tempered stable distribution in the MRS-GARCH model and argue that it outperforms the commonly used Student's t and GED. As to the specification of MRS-GARCH model, we use the one described in Haas et al. (2004).[2] To demonstrate that, we conduct a series of simulation studies to compare the performance of MRS-GARCH model with those three distributions, where all the data generation processes are MRS-GARCH(1,1). First, we set the true distribution as Student's t and GED, respectively. Via twelve combinations of different transition probabilities and GARCH parameters, MRS-GARCH models with three distinct fat-tailed distributions are systematically analyzed. It is demonstrated that when the true distribution is Student's t or GED, the MRS-GARCH model with tempered stable distribution generates very similar results as the one with the true distribution. More importantly, it outperforms the other competitor in terms of consistency, efficiency and overall performance. Second, we let the tempered stable be the true distribution. Twelve sets of simulations are further constructed, including different choices of transition probabilities and tempered stable distribution parameters. In this scenario, both the Student's t or GED distributions cannot perform as well as the tempered stable. Consequently, we argue that the tempered stable distribution could be a widely useful tool for modeling the financial volatility in general contexts with a MRS-GARCH-type specification.

To empirically compare the MRS-GARCH model with different distributions, we apply them to the daily return of the S&P 500 index. The results suggest that the MRS-GARCH model with tempered stable distribution outperforms that with all the other distributions. Besides, the estimated parameters across different distributions are relatively close to each other. The three fitted conditional volatility series are also quite similar. Additionally, the fitted smoothing probability series of the MRS-GARCH model with tempered stable and GED are roughly the same. The fitted series of the MRS-GARCH model with the Student's t-distribution differs from them. Finally, the density of the fitted tempered stable distribution is closer to that of the standardized data, compared with those of the other two fitted distributions.

The remainder of this paper proceeds as follows. Section 2 describes the specification of the MRS-GARCH model employed in this study. Section 3 explains how the Student's t, GED and tempered stable distributions can be applied to the MRS-GARCH model. We conduct three independent simulation studies in Section 4. The empirical results are discussed in Section 5. Section 6 concludes the paper.

## 2. MRS-GARCH model

### 2.1. GARCH model

GARCH model is proposed by Bollerslev (1986), which is generalized from the seminal work on ARCH model by Engle (1982). Because of its capabilities to capture some important characteristics of financial time series like time varying heteroskedasticity and volatility clustering, GARCH model has become a standard way to study financial volatility (Ho et al., 2013). The original GARCH(1,1) model has the following specification.

$$\begin{aligned} r_t &= \mu + \varepsilon_t \text{ and } \varepsilon_t = \eta_t \sqrt{h_t} \\ h_t &= \omega + \alpha\varepsilon_{t-1}^2 + \beta h_{t-1} \end{aligned} \tag{1}$$

where $r_t$ is the interested financial time series, $\varepsilon_t$ is the residual series, $h_t$ is its conditional volatility and $\eta_t$ is an identical and independent innovation sequence. Therefore, the parameter $\alpha + \beta$ measures the volatility persistence, which means how fast the current shock to the volatility will die away (Ho et al., 2013). In order to ensure that $h_t$ is stationary and always positive, Bollerslev (1986) suggests to apply the constraints $\alpha + \beta < 1$ and $\omega, \alpha, \beta \geq 0$.

### 2.2. MRS-GARCH model

The main weakness of the GARCH model is that it assumes that the conditional volatility has only one regime over the entire period. Therefore, the GARCH model is not suitable when the data exhibit structure breaks over the sample period, since it implies a high volatility persistence in such case (Hamilton and Susmel, 1994; Klaassen, 2002; Lamoureux and Lastrapes, 1990).

To overcome this problem, various MRS-GARCH models have been proposed by extending the idea of MRS (Hamilton, 1988, 1989) to the GARCH-type framework. In this paper, we employ the two-state MRS-GARCH(1,1) model investigated in Haas et al. (2004), and its specification is as follows:[3]

$$\begin{aligned} &r_t = \mu + \varepsilon_t \text{ and } \varepsilon_t = \eta_t \sqrt{h_{s_t,t}} \\ &h_{s_t,t} = \begin{cases} \omega_1 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{1,t-1} \text{ when } s_t = 1 \\ \omega_2 + \alpha_2 \varepsilon_{t-1}^2 + \beta_2 h_{2,t-1} \text{ when } s_t = 2 \end{cases} \end{aligned} \tag{2}$$

where $s_t$ is the state that $r_t$ lies in at time $t$ and $h_{s_t,t}$ is the conditional volatility in state $s_t$ at time $t$. In addition, the sequence $\{s_t\}$ is assumed to be a stationary, irreducible and latent Markov process with discrete state space {1,2} and transition matrix $P_i = [p_{jk}]$ where $p_{jk} = P(s_{t+1} = k | s_t = j)$ is the one-step transition probability of moving from state $j$ to state $k$ ($j, k \in \{1, 2\}$). The benefit of this specification is that the volatility persistence in state $j$ is measured by $P_j = \alpha_j + \beta_j$, as in the GARCH model (Haas et al., 2004).

[1] In this case, the associated Levy processes are called “truncated Levy flights”, the appropriateness of which to be applied in the GARCH-type model is also discussed in Constantinides and Savel'ev (2013).

[2] The reason for choosing this specification is that the path dependency problem is solved by proposing independent conditional volatility series for each regime. Hence, for this MRS-GARCH model, each of those independent series has exactly the same properties as the conditional volatility of the original GARCH model. Therefore, the interpretation of interested properties like volatility persistence can be made for each regime.

[3] In this study, we only allow regime-switching in the conditional volatility equation. It is possible to allow regime-switching only in mean or in both mean and conditional volatility. We have tested all the cases, and the results are consistent and available upon request.

As argued by Mullen et al. (2011), Eq. (2) can capture the volatility clustering as in the GARCH model, as well as allowing the structure breaks in unconditional variance. In the $j$th regime, the unconditional variance of $r_t$ is:

$$\overline{\sigma}_j^2 = \frac{\omega_j}{1-\alpha_j-\beta_j} \tag{3}$$

as long as $\alpha_j + \beta_j < 1$; that is, the process is covariance stationary (Bollerslev, 1986; Haas et al., 2004). In this paper, we indicate state 1 as the calm state and state 2 as the turbulent state, and it is restricted that $\overline{\sigma}_1^2 < \overline{\sigma}_2^2$.

In order to estimate parameters of the MRS-GARCH model, MLE is employed. Therefore, the innovation series $\eta_t$ needs to follow a specific distribution. Originally, GARCH model is developed based on standard Normal (Gaussian) distribution. In other words, $\eta_t = \varepsilon_t/\sqrt{h_t} \sim N(0,1)$[4]. Hence, the conditional density of $\varepsilon_t$ can be constructed as follows.

$$\begin{aligned} \Omega_{t-1} &= \{\varepsilon_{t-1}, \varepsilon_{t-2}, \ldots, \varepsilon_1\} \\ \theta &= (\mu, \omega_1, \omega_2, \alpha_1, \alpha_2, \beta_1, \beta_2, p_{11}, p_{22})' \\ f(\varepsilon_t | s_t = j, \theta, \Omega_{t-1}) &= \frac{1}{\sqrt{2\pi h_{j,t}}} e^{-\frac{\varepsilon_t^2}{2h_{j,t}}} \end{aligned} \tag{4}$$

where $\Omega_{t-1}$ is the information set at time $t-1$, $\theta$ is the vector of parameters, and $f(\varepsilon_t|s_t = j, \theta_i, \Omega_{t-1})$ is the conditional density of $\varepsilon_t$. This stems from the fact that in state $j$, $\varepsilon_t$ follows a Normal distribution with mean 0, variance $h_{j,t}$ given time $t-1$.

Plugging the filtered probability in state $j$ at time $t-1$, $\rho_{j,t-1} = P(s_{t-1} = j|\theta_i, \Omega_{t-1})$, into Eq. (4) and integrating out the state variable $s_{t-1}$, the density function in Eq. (4) becomes:

$$f(\varepsilon_t|\theta\Omega_{t-1}) = \sum_{j=1}^{2}\sum_{k=1}^{2} p_{jk}\rho_{j,t-1} f(\varepsilon_t|s_t = j, \theta, \Omega_{t-1}) \tag{5}$$

$\rho_{j,t-1}$ can be obtained by an integrative algorithm given in Hamilton (1989). The log-likelihood function corresponding to Eq. (2) is as follows:

$$L(\theta|\varepsilon) = \sum_{t=2}^{T} \ln f(\varepsilon_t|\theta, \Omega_{t-1}) \text{ where } \varepsilon = (\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_T)' \tag{6}$$

and the maximum likelihood estimator $\hat{\theta}$ is obtained by maximizing Eq. (6).

Since $\{s_t\}$ is a latent process, a certain statistic is needed to identify the state structure of $r_t$. Therefore, we estimate the smoothing probability of the calm state as follows (Hamilton, 1989):

$$P(s_t = 1|\theta, \Omega_T) = \rho_{1,t}\left[\frac{p_{11}P(s_{t+1} = 1|\theta, \Omega_T)}{P(s_{t+1} = 1|\theta, \Omega_t)} + \frac{p_{12}P(s_{t+1} = 2|\theta, \Omega_T)}{P(s_{t+1} = 2|\theta, \Omega_t)}\right]. \tag{7}$$

Using the fact that $P(s_T = 1|\theta, \Omega_T) = \rho_{1,T}$, the smoothing probability series $P(s_t = 1|\theta, \Omega_T)$ can be generated by iterating Eq. (7) backwards from $T$ to 1. Further, we apply the widely recognized rule (Hamilton, 1989) that $r_t$ lies in the turbulent state if $P(s_t = 1|\theta, \Omega_T) < 0.5$ and otherwise in the calm state. Hence, the state structure of $r_t$ can be estimated from the fitted MRS-GARCH model.

## 3. Alternative distributions for the innovation sequence

Although the original GARCH model is based on the Gaussian distribution, significant evidence suggests that the financial time series is rarely Gaussian but typically leptokurtic and exhibits fat-tail behavior (Bollerslev, 1987; Ho et al., 2013; Stanley et al., 2008; Susmel and Engle, 1994). Bollerslev and Wooldridge (1992) propose the QMLE, which can accommodate for fat-tailedness through their specification and still generates consistent estimates. However, for the MRS-GARCH model, if regimes are not Gaussian but leptokurtic, the use of within-regime normality can seriously affect the identification of the regime process (Klaassen, 2002; Ardia, 2009; Haas, 2009). Additionally, Haas and Paolella (2012) further argue that QMLE of the MRS-GARCH model is not consistent. As a result, the MRS-GARCH model with Gaussian distribution will be inappropriate in practice. Hence, the sought of an alternative distribution to accommodate the excess kurtosis of the financial time series becomes an essential issue for the MRS-GARCH model.

### 3.1. Student's t and general error distribution

Among the existing literature, Student's t-distribution and GED are two widely used alternatives in finance study using GARCH-type models (Chkili et al., 2012; Fan et al., 2008; Mabrouk and Saadi, 2012; Zhu and Galbraith, 2011). Both of those two distributions can capture leptokurtic and heavy-tail behaviors. When they are applied to the MRS-GARCH model, the corresponding density functions of are changed as follows.

$$\begin{aligned} &\text{Student's t}: f(\varepsilon_t|s_t = j, \theta, \Omega_{t-1}) = \frac{\Gamma\left(\frac{v+1}{2}\right)}{\Gamma\left(\frac{v}{2}\right)\sqrt{\pi(v-2)h_{j,t}}}\left[1 + \frac{\varepsilon_t^2}{(v-2)h_{j,t}}\right]^{\frac{v+1}{2}} \\ &\text{GED}: f(\varepsilon_t|s_t = j, \theta, \Omega_{t-1}) = \frac{ve^{-\frac{1}{2}\left|\frac{\varepsilon_t}{\lambda\sqrt{h_{j,t}}}\right|^{v}}}{\lambda 2^{(v+1)/v}\Gamma(1/v)} \\ &\text{where } \lambda = \left[\frac{2^{(-2/v)}\Gamma(1/v)}{\Gamma(3/v)}\right]^{\frac{1}{2}} \end{aligned} \tag{8}$$

and $v$ is the degree of freedom. Then, the MLE estimator $\hat{\theta}$ can be obtained in the same way as described in Section 2.2.

### 3.2. Tempered stable distribution

Despite their attractive properties to capture excess kurtosis and fat-tails, existing literature argues that the Student's t and GED still have unsolved problems. For example, Yang and Brorsen (1993) indicate that the tail behavior of GARCH-type model remains too short even with Student's t distributed error terms. Furthermore, Calzolari et al. (2014) suggests that the Student's t-distribution lacks the stability-under-addition property. Since stability is desirable in portfolio applications and risk management, a distribution could overcome this issue is of particular importance.

#### 3.2.1. Symmetric **α**-stable distribution

As suggested by Calzolari et al. (2014), the $\alpha$-stable distribution (also known as stable family of distributions) is a replacement of the traditionally used fat-tail distribution. Its most outstanding characteristic is that $\alpha$-stable distribution can further overcome the stability problem of the Student's t. The $\alpha$-stable distribution constitutes a generalization of the Gaussian distribution by allowing for asymmetry and heavy tails. In general, a random variable $x$ is said to be stably distributed 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

$$kx + d \stackrel{d}{=} c_1x_1 + c_2x_2 \tag{9}$$

where $x_1$ and $x_2$ are independent variables and have the same distribution as $x$. The notation $\stackrel{d}{=}$ indicates equality in distribution. In particular, if $d = 0$, $x$ is said to be strictly stable. According to Calzolari et al. (2014),

[4] Since the mean of $\varepsilon_t$ is 0, $\eta_t$ is sometimes named as standardized residual.

theoretical foundations of $\alpha$-stable distribution lay on the generalized 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.

Since $\alpha$-stable distribution does not have a close form of density function, the best way to describe it is by means of its characteristic function. If we only consider the symmetric $\alpha$-stable distribution case, then its characteristic function is of the form

$$\phi(t) = \exp\{i\delta t - \sigma^{\alpha}|t|^{\alpha}\} \tag{10}$$

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), $\sigma \in \mathrm{R}^{+}$ is the scale parameter and $\delta \in \mathrm{R}$ is the location parameter (Calzolari et al., 2014). The symmetric $\alpha$-stable distribution is then characterized by $(\alpha, \sigma, \delta)$ and is denoted as $S(\alpha, \sigma, \delta)$. Therefore, the standardized symmetric version is $S(\alpha, 1, 0)$ with the following characteristic function

$$\phi(t) = \exp\{-|t|^{\alpha}\}. \tag{11}$$

Despite its attractive properties, the second moment of the $\alpha$-stable distribution does not exist in most cases. Consequently, the application of this distribution to GARCH-type model will cause serious problems. For instance, the interpretation of conditional volatility would fail. Therefore, the sought of a substitute of the $\alpha$-stable distribution, which has similar attractive properties and defined moments, would be of particular interest.

### 3.2.2. Tempered stable distribution

A natural candidate is the tempered stable distribution. A general case of the tempered stable distribution is characterized by six parameters and denoted as $TS(\alpha^{+}, C^{+}, \lambda^{+}; \alpha^{-}, C^{-}, \lambda^{-})$. The levy measure of such random variable $x$ is

$$\nu(x) = \frac{C^{-}}{|x|^{1+\alpha^{-}}} e^{-\lambda^{-}|x|} 1_{x<0} + \frac{C^{+}}{|x|^{1+\alpha^{+}}} e^{-\lambda^{+}|x|} 1_{x>0}. \tag{12}$$

Therefore, a tempered stable distribution with zero mean has the following characteristic function (Cont and Tankov, 2004).

$$\begin{aligned}\phi(t) = \exp\Bigg\{&\Gamma(-\alpha^{+})(\lambda^{+})^{\alpha^{+}} C^{+}\left[\left(1-\frac{it}{\lambda^{+}}\right)^{\alpha^{+}} - 1 + \frac{it\alpha^{+}}{\lambda^{+}}\right] \\ &+ \Gamma(-\alpha^{-})(\lambda^{-})^{\alpha^{-}} C^{-}\left[\left(1-\frac{it}{\lambda^{-}}\right)^{\alpha^{-}} - 1 + \frac{it\alpha^{-}}{\lambda^{-}}\right]\Bigg\}\end{aligned} \tag{13}$$

where $\alpha^{+}, \alpha^{-} < 2$ and $C^{+}, C^{-}, \lambda^{+}, \lambda^{-} > 0$. Its first four cumulants are hereby defined as:

$$\begin{aligned}\kappa_1 &= 0 \\ \kappa_2 &= \Gamma(2-\alpha^{+})C^{+}(\lambda^{+})^{\alpha^{+}-2} + \Gamma(2-\alpha^{-})C^{-}(\lambda^{-})^{\alpha^{-}-2} \\ \kappa_3 &= \Gamma(3-\alpha^{+})C^{+}(\lambda^{+})^{\alpha^{+}-3} + \Gamma(3-\alpha^{-})C^{-}(\lambda^{-})^{\alpha^{-}-3} \\ \kappa_4 &= \Gamma(4-\alpha^{+})C^{+}(\lambda^{+})^{\alpha^{+}-4} + \Gamma(4-\alpha^{-})C^{-}(\lambda^{-})^{\alpha^{-}-4}.\end{aligned} \tag{14}$$

Hence, the first four moments of $x$ can be found via the following relations:

$$\begin{aligned}m_1 &= \kappa_1 \\ m_2 &= \kappa_2 + \kappa_1^2 \\ m_3 &= \kappa_3 + 3\kappa_2\kappa_1 + \kappa_1^3 \\ m_4 &= \kappa_4 + 4\kappa_3\kappa_1 + 3\kappa_2^2 + 6\kappa_2\kappa_1^2 + \kappa_1^4.\end{aligned} \tag{15}$$

Therefore, it is clear that the tempered stable distribution has defined moments, so that it can be further applied to describe the innovation of GARCH-type model.

Küchler and Tappe (2013) suggest that the behavior of the sample paths of $x$ depends on the values of $\alpha^{+}$ and $\alpha^{-}$. In particular, if $\alpha^{+} = \alpha^{-}$, $x$ follows a classical tempered stable distribution. In addition, if we further require $C^{+} = C^{-}$, then $x$ follows a CGMY distribution (Carr et al., 2002).

- For $\alpha^{+}, \alpha^{-} < 0$ we have $\nu(\mathrm{R}) < \infty$, and thus, $x$ is a compound Poisson process.
- For $\alpha^{+}, \alpha^{-} \in [0, 1)$ we have $\nu(\mathrm{R}) = \infty$, but $\int_{-1}^{1}|x|\nu(dx) < \infty$. Therefore, $x$ is a finite-variation process making infinitely many jumps in each interval of positive length, which we can express as $x_t = \sum_{s \le t}\Delta x_s$.
- For $\alpha^{+}, \alpha^{-} \in (1, 2)$ we have $\int_{-1}^{1}|x|\nu(dx) = \infty$. Thus, $x$ has sample paths of infinite variation.

To apply this distribution into the MRS-GARCH model, we require $x$ to be standardized. Bianchi et al. (2010)argues that one way to achieve the standardization is letting

$$C^{+} = \frac{p(\lambda^{+})^{2-\alpha^{+}}}{\Gamma(2-\alpha^{+})} \text{ and } C^{-} = \frac{(1-p)(\lambda^{-})^{2-\alpha^{-}}}{\Gamma(2-\alpha^{-})} \tag{16}$$

where $p \in (0, 1)$, then $x \sim TS(\alpha^{+}, \lambda^{+}, \alpha^{-}, \lambda^{-}, p)$ has zero mean and unit variance.

Combining Eqs. (13) and (16), we now have a standardized tempered stable distribution. This distribution is expected to retain all the attractive properties similar to those of $\alpha$-stable distribution and can be further employed to describe the innovation of the MRS-GARCH model.

### 3.2.3. MRS-GARCH model with tempered stable distribution

Since the tempered stable distribution has defined fourth moment, the sufficient condition to ensure the asymptotic properties of the MRS-GARCH model will not be affected.[5] Hence, we expect the MLE of MRS-GARCH model with tempered stable distribution would provide consistent estimators.

In terms of estimation, the procedures of MLE as discussed in Section 2.2 can still be used. Since the tempered stable distribution also does not have a closed form of density function, a certain numerical algorithm is needed to compute it (Kim et al., 2008). As argued by Mittnik et al. (1999), compared with other approximation methods, discrete Fourier transform is accurate and efficient to estimate parameters of stable family distributions, especially when $N = 2^{13}$ or above. Therefore, we employ the discrete Fourier transform method to obtain the estimated $f(\varepsilon_t|s_t, \theta, \Omega_{t-1})$ via the following iteration steps:

1. Acquiring the minimum and maximum of $\eta_t = \varepsilon_t/\sqrt{h_t}$ as $\eta_1$ and $\eta_2$, respectively;
2. Calculating the values of $\phi(t)$ for the tempered stable distribution determined by the estimates of $\alpha^{+}, C^{+}, \lambda^{+}, \alpha^{-}, C^{-}$ and $\lambda^{-}$ via Eq. (13), where $t$ evenly ranges from $\eta_1 - 0.1$ to $\eta_2 + 0.1$ with size $N = 2^{15}$;[6]
3. Using discrete Fourier transform to find the values of the corresponding density function for $t$; and
4. Employing the linear interpolation to $eta_t$ that fall between the prespecified equally-spaced density values of $t$.

Hence, the interpolated values will be the estimates of $f(\varepsilon_t|s_t, \theta, \Omega_{t-1})$. By further applying them to Eqs. (5) and (6), the required log-likelihood values can be obtained.

## 4. Comparisons between distributions: simulation studies

In this Section, we will conduct three simulation studies to compare the performance of Student's t, GED and tempered stable distributions.

[5] Specific conditions can be found in Haas et al. (2004).

[6] As discussed in Mittnik et al. (1999), discrete Fourier transform works most efficiently for $N$ being expressed in terms of a power of 2.

**Table 1**
Simulation results: Student's t-distribution.

| $p_{11}$ | $p_{22}$ | $P_1$ | $P_2$ | $Mean_{ll}$ | $Bias_{p_{11}}$ | $SE_{p_{11}}$ | $RMSE_{p_{11}}$ | $Bias_{p_{22}}$ | $SE_{p_{22}}$ | $RMSE_{p_{22}}$ | $Bias_{P_1}$ | $SE_{P_1}$ | $RMSE_{P_1}$ | $Bias_{P_2}$ | $SE_{P_2}$ | $RMSE_{P_2}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *Panel A: Student's t-distribution* | | | | | | | | | | | | | | | | |
| 0.99 | 0.999 | 0.75 | 0.95 | −9411 | −0.0034 | 0.0094 | 0.0100 | −0.0001 | 0.0006 | 0.0006 | −0.0767 | 0.2096 | 0.2232 | −0.0075 | 0.0192 | 0.0206 |
| | | 0.85 | 0.85 | −7955 | −0.0093 | 0.0509 | 0.0517 | −0.0048 | 0.0439 | 0.0442 | −0.0383 | 0.1634 | 0.1679 | 0.0961 | 0.0333 | 0.1017 |
| | | 0.95 | 0.75 | −7491 | −0.0180 | 0.0582 | 0.0609 | −0.0027 | 0.0098 | 0.0102 | −0.0608 | 0.1288 | 0.1424 | −0.1098 | 0.0687 | 0.1295 |
| 0.999 | 0.99 | 0.75 | 0.95 | −4208 | −0.0002 | 0.0007 | 0.0007 | −0.0047 | 0.0114 | 0.0123 | −0.0269 | 0.0662 | 0.0715 | −0.0430 | 0.1336 | 0.1403 |
| | | 0.85 | 0.85 | −4676 | −0.0006 | 0.0020 | 0.0020 | −0.0056 | 0.0157 | 0.0167 | −0.0104 | 0.0412 | 0.0425 | 0.0669 | 0.1458 | 0.1605 |
| | | 0.95 | 0.75 | −6105 | −0.0074 | 0.0429 | 0.0436 | −0.0183 | 0.0649 | 0.0674 | −0.0179 | 0.0343 | 0.0387 | −0.0828 | 0.1784 | 0.1967 |
| 0.99 | 0.99 | 0.75 | 0.95 | −6984 | −0.0011 | 0.0029 | 0.0031 | −0.0007 | 0.0027 | 0.0027 | −0.1580 | 0.1573 | 0.2230 | −0.0016 | 0.0263 | 0.0263 |
| | | 0.85 | 0.85 | −6420 | −0.0009 | 0.0040 | 0.0041 | −0.0009 | 0.0036 | 0.0037 | −0.0179 | 0.0687 | 0.0710 | 0.0936 | 0.0491 | 0.1057 |
| | | 0.95 | 0.75 | −6829 | −0.0046 | 0.0122 | 0.0130 | −0.0042 | 0.0127 | 0.0134 | −0.0248 | 0.0416 | 0.0484 | −0.1097 | 0.1005 | 0.1488 |
| 0.999 | 0.999 | 0.75 | 0.95 | −6682 | −0.0002 | 0.0008 | 0.0008 | −0.0005 | 0.0018 | 0.0019 | −0.0274 | 0.1100 | 0.1134 | −0.0040 | 0.0310 | 0.0312 |
| | | 0.85 | 0.85 | −6262 | −0.0004 | 0.0020 | 0.0020 | −0.0004 | 0.0011 | 0.0012 | −0.0093 | 0.0521 | 0.0529 | 0.1000 | 0.0475 | 0.1107 |
| | | 0.95 | 0.75 | −6858 | −0.0010 | 0.0031 | 0.0033 | −0.0007 | 0.0030 | 0.0031 | −0.0146 | 0.0324 | 0.0356 | −0.1142 | 0.0952 | 0.1487 |
| *Panel B: GED distribution* | | | | | | | | | | | | | | | | |
| 0.99 | 0.999 | 0.75 | 0.95 | −9486 | −0.0129 | 0.0742 | 0.0753 | −0.0186 | 0.1167 | 0.1181 | −0.0675 | 0.2152 | 0.2255 | −0.0390 | 0.0199 | 0.0438 |
| | | 0.85 | 0.85 | −8024 | −0.0195 | 0.0765 | 0.0790 | −0.1232 | 0.2924 | 0.3172 | −0.0543 | 0.1510 | 0.1604 | 0.0722 | 0.0732 | 0.1028 |
| | | 0.95 | 0.75 | −7553 | −0.1368 | 0.2271 | 0.2651 | −0.2124 | 0.3132 | 0.3785 | −0.1013 | 0.1237 | 0.1598 | −0.1388 | 0.1478 | 0.2027 |
| 0.999 | 0.99 | 0.75 | 0.95 | −4259 | −0.0159 | 0.0239 | 0.0287 | −0.2001 | 0.3099 | 0.3689 | −0.0581 | 0.1290 | 0.1414 | −0.0779 | 0.1539 | 0.1725 |
| | | 0.85 | 0.85 | −4709 | −0.0382 | 0.0250 | 0.0457 | −0.5789 | 0.3392 | 0.6710 | −0.0496 | 0.0609 | 0.0786 | 0.1301 | 0.2104 | 0.2473 |
| | | 0.95 | 0.75 | −6126 | −0.0441 | 0.0254 | 0.0509 | −0.7052 | 0.2752 | 0.7570 | −0.0779 | 0.0199 | 0.0804 | 0.1089 | 0.1002 | 0.1479 |
| 0.99 | 0.99 | 0.75 | 0.95 | −7048 | −0.0030 | 0.0031 | 0.0043 | −0.0024 | 0.0036 | 0.0044 | −0.2010 | 0.1745 | 0.2662 | −0.0496 | 0.0279 | 0.0569 |
| | | 0.85 | 0.85 | −6482 | −0.0086 | 0.0121 | 0.0148 | −0.0713 | 0.2238 | 0.2349 | −0.0712 | 0.0748 | 0.1032 | 0.0263 | 0.0985 | 0.1019 |
| | | 0.95 | 0.75 | −6868 | −0.0412 | 0.0396 | 0.0571 | −0.5714 | 0.3801 | 0.6863 | −0.0804 | 0.0425 | 0.0910 | 0.0093 | 0.1793 | 0.1796 |
| 0.999 | 0.999 | 0.75 | 0.95 | −6757 | −0.0010 | 0.0032 | 0.0034 | −0.0061 | 0.0593 | 0.0596 | −0.0697 | 0.1284 | 0.1461 | −0.0475 | 0.0425 | 0.0637 |
| | | 0.85 | 0.85 | −6335 | −0.0031 | 0.0095 | 0.0100 | −0.0381 | 0.1718 | 0.1760 | −0.0666 | 0.0675 | 0.0948 | 0.0537 | 0.0796 | 0.0960 |
| | | 0.95 | 0.75 | −6925 | −0.0247 | 0.0857 | 0.0892 | −0.1974 | 0.3560 | 0.4071 | −0.0566 | 0.0598 | 0.0823 | −0.0863 | 0.1429 | 0.1669 |
| *Panel C: Tempered stable distribution* | | | | | | | | | | | | | | | | |
| 0.99 | 0.999 | 0.75 | 0.95 | −9410 | −0.0043 | 0.0114 | 0.0122 | −0.0002 | 0.0009 | 0.0009 | −0.0414 | 0.1741 | 0.1790 | −0.0191 | 0.0199 | 0.0276 |
| | | 0.85 | 0.85 | −7953 | −0.0052 | 0.0190 | 0.0197 | −0.0007 | 0.0022 | 0.0023 | −0.0131 | 0.1482 | 0.1488 | 0.0813 | 0.0317 | 0.0873 |
| | | 0.95 | 0.75 | −7489 | −0.0174 | 0.0545 | 0.0572 | −0.0067 | 0.0321 | 0.0328 | −0.0538 | 0.1033 | 0.1164 | −0.1309 | 0.0924 | 0.1603 |
| 0.999 | 0.99 | 0.75 | 0.95 | −4204 | −0.0002 | 0.0007 | 0.0007 | −0.0052 | 0.0208 | 0.0214 | −0.0282 | 0.0659 | 0.0716 | −0.0449 | 0.1132 | 0.1218 |
| | | 0.85 | 0.85 | −4672 | −0.0003 | 0.0015 | 0.0015 | −0.0170 | 0.0999 | 0.1013 | −0.0167 | 0.0386 | 0.0421 | 0.0516 | 0.1594 | 0.1675 |
| | | 0.95 | 0.75 | −6101 | −0.0061 | 0.0345 | 0.0350 | −0.0433 | 0.1561 | 0.1620 | −0.0235 | 0.0274 | 0.0361 | −0.0906 | 0.1979 | 0.2176 |
| 0.99 | 0.99 | 0.75 | 0.95 | −6983 | −0.0011 | 0.0029 | 0.0031 | −0.0008 | 0.0027 | 0.0028 | −0.1491 | 0.1420 | 0.2059 | −0.0153 | 0.0268 | 0.0308 |
| | | 0.85 | 0.85 | −6417 | −0.0005 | 0.0040 | 0.0040 | −0.0009 | 0.0036 | 0.0037 | −0.0192 | 0.0711 | 0.0736 | 0.0787 | 0.0542 | 0.0956 |
| | | 0.95 | 0.75 | −6826 | −0.0038 | 0.0109 | 0.0116 | −0.0044 | 0.0126 | 0.0133 | −0.0278 | 0.0408 | 0.0494 | −0.1329 | 0.1223 | 0.1806 |
| 0.999 | 0.999 | 0.75 | 0.95 | −6681 | −0.0002 | 0.0008 | 0.0008 | −0.0005 | 0.0017 | 0.0018 | −0.0221 | 0.0934 | 0.0960 | −0.0160 | 0.0320 | 0.0357 |
| | | 0.85 | 0.85 | −6260 | −0.0004 | 0.0018 | 0.0018 | −0.0004 | 0.0018 | 0.0019 | −0.0149 | 0.0554 | 0.0574 | 0.0872 | 0.0474 | 0.0993 |
| | | 0.95 | 0.75 | −6855 | −0.0012 | 0.0048 | 0.0050 | −0.0007 | 0.0023 | 0.0024 | −0.0171 | 0.0258 | 0.0310 | −0.1265 | 0.1090 | 0.1670 |

The data generation process is MRS-GARCH(1,1) as described in Eq. (2) in all cases. True distributions of the three studies are therefore Student's t, GED and tempered stable, respectively.

### 4.1. Student's t-distribution

First, we set the true distribution as Student's t with 3 degrees of freedom.[7] Altogether, twelve sets of simulations of the MRS-GARCH(1,1) process with different $p_{11}$, $p_{22}$,[8] $P_1$ and $P_2$ [9] are generated, where $\mu=0, \omega_1=0.1, \omega_2=0.5$, sample size is 5000 and number of replicates for each set is 300. To avoid the starting bias, 10,000 points are generated for each simulation, and then only the last 5000 are kept. Moreover, to avoid simulation bias, 500 such replicates are produced for each combination, while the first 200 are discarded.

The simulated data are fitted into MRS-GARCH model with Student's t (MRS-GARCH-t), GED (MRS-GARCH-G) and tempered stable (MRS-GARCH-S) distributions, respectively. In Table 1, the mean log-likelihood (LL), bias, standard error (SE) and root-mean-square-error (RMSE) of $p_{11}$, $p_{22}$, $P_1$ and $P_2$ are reported. Bias is the mean difference between the true parameter and its estimate, SE is the standard error of the estimates, and RMSE is the square root of the mean of squared difference between the true parameter and its estimate.

Although tempered stable distribution has four more parameters than Student's t and GED, the log-likelihood is still a preliminary indicator of the model performance. Not surprisingly, MRS-GARCH-G has smaller log-likelihood than the true model MRS-GARCH-t in all cases. Nevertheless, it is worth noticing that GARCH-S can yield slightly greater log-likelihood compared with MRS-GARCH-t. Therefore, it seems that MRS-GARCH-S model might lead to quite satisfied results, even when the true distribution is Student's t.

#### 4.1.1. Estimates of transition probabilities

In terms of bias comparison, most absolute values of MRS-GARCH-G are over 0.01 for $p_{11}$. one of them even exceeds 0.1. Those of MRS-GARCH-t and MRS-GARCH-S are almost all smaller than 0.01. As to $p_{22}$, the difference is much greater. MRS-GARCH-G cannot generate consistent estimates in most cases, where some absolute bias are as large as 0.57. All the bias of MRS-GARCH-t and MRS-GARCH-S are much smaller and quite similar.

SE stands for the estimation efficiency. For $p_{11}$, it is observed that SEs of MRS-GARCH-t and MRS-GARCH-S are basically on the same scale and are overall smaller than those of MRS-GARCH-G. Again, SEs of MRS-GARCH-G become much worse for $p_{22}$, where almost all values are greater than 0.1. On the other hand, MRS-GARCH-S still has similar performance as MRS-GARCH-t.

[7] We also consider the cases of 4, 5 and 6 degrees of freedom. The results are robust and available upon request.

[8] We also conduct studies with smaller transition probabilities of 0.25, 0.5, 0.75 and 0.9. The results are consistent and available upon request.

[9] $\alpha$ and $\beta$ are respectively set to 0.1 and 0.65, 0.2 and 0.65, and 0.2 and 0.75 for volatility persistence of 0.75, 0.85 and 0.95.

**Table 2**
Simulation results: GED.

| $p_{11}$ | $p_{22}$ | $P_1$ | $P_2$ | $Mean_{ll}$ | $Bias_{p_{11}}$ | $SE_{p_{11}}$ | $RMSE_{p_{11}}$ | $Bias_{p_{22}}$ | $SE_{p_{22}}$ | $RMSE_{p_{22}}$ | $Bias_{P_1}$ | $SE_{P_1}$ | $RMSE_{P_1}$ | $Bias_{P_2}$ | $SE_{P_2}$ | $RMSE_{P_2}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *Panel A: Student's t-distribution* | | | | | | | | | | | | | | | | |
| 0.99 | 0.999 | 0.75 | 0.95 | −10,743 | −0.0621 | 0.1826 | 0.1929 | −0.0181 | 0.0590 | 0.0617 | −0.1130 | 0.2150 | 0.2429 | 0.0247 | 0.0153 | 0.0290 |
| | | 0.85 | 0.85 | −8946 | −0.1107 | 0.2383 | 0.2627 | −0.0404 | 0.0963 | 0.1044 | −0.1022 | 0.2075 | 0.2313 | 0.1399 | 0.0394 | 0.1453 |
| | | 0.95 | 0.75 | −8311 | −0.2409 | 0.3242 | 0.4039 | −0.0799 | 0.1193 | 0.1436 | −0.1485 | 0.2233 | 0.2682 | −0.0704 | 0.0772 | 0.1045 |
| 0.999 | 0.99 | 0.75 | 0.95 | −5030 | −0.0100 | 0.0691 | 0.0698 | −0.0114 | 0.0410 | 0.0425 | −0.0436 | 0.1260 | 0.1333 | −0.0018 | 0.0922 | 0.0922 |
| | | 0.85 | 0.85 | −5687 | −0.1293 | 0.2652 | 0.2950 | −0.0553 | 0.1122 | 0.1251 | −0.0720 | 0.2188 | 0.2303 | 0.1183 | 0.1373 | 0.1812 |
| | | 0.95 | 0.75 | −7424 | −0.5175 | 0.3546 | 0.6274 | −0.1766 | 0.1548 | 0.2348 | −0.1460 | 0.1746 | 0.2276 | 0.0886 | 0.0896 | 0.1260 |
| 0.99 | 0.99 | 0.75 | 0.95 | −8036 | −0.0028 | 0.0039 | 0.0048 | −0.0025 | 0.0036 | 0.0044 | −0.2185 | 0.1907 | 0.2899 | 0.0261 | 0.0221 | 0.0341 |
| | | 0.85 | 0.85 | −7401 | −0.0422 | 0.1402 | 0.1464 | −0.0158 | 0.0410 | 0.0439 | −0.0785 | 0.1544 | 0.1732 | 0.1357 | 0.0677 | 0.1516 |
| | | 0.95 | 0.75 | −7877 | −0.2868 | 0.3223 | 0.4314 | −0.1084 | 0.1242 | 0.1648 | −0.1046 | 0.1754 | 0.2042 | 0.0141 | 0.1083 | 0.1092 |
| 0.999 | 0.999 | 0.75 | 0.95 | −7932 | −0.0034 | 0.0415 | 0.0416 | −0.0012 | 0.0148 | 0.0148 | −0.0458 | 0.1299 | 0.1377 | 0.0242 | 0.0243 | 0.0343 |
| | | 0.85 | 0.85 | −7351 | −0.0007 | 0.0036 | 0.0037 | −0.0005 | 0.0013 | 0.0014 | 0.0249 | 0.0580 | 0.0631 | 0.1310 | 0.0486 | 0.1397 |
| | | 0.95 | 0.75 | −7864 | −0.0582 | 0.1911 | 0.1997 | −0.0181 | 0.0555 | 0.0583 | −0.0217 | 0.1007 | 0.1030 | −0.0376 | 0.0855 | 0.0934 |
| *Panel B: GED distribution* | | | | | | | | | | | | | | | | |
| 0.99 | 0.999 | 0.75 | 0.95 | −10,700 | −0.0059 | 0.0171 | 0.0181 | −0.0003 | 0.0016 | 0.0016 | −0.1241 | 0.2228 | 0.2550 | −0.0031 | 0.0158 | 0.0161 |
| | | 0.85 | 0.85 | −8907 | −0.0056 | 0.0171 | 0.0180 | −0.0008 | 0.0034 | 0.0035 | −0.0510 | 0.1678 | 0.1754 | 0.0947 | 0.0277 | 0.0987 |
| | | 0.95 | 0.75 | −8278 | −0.0202 | 0.0785 | 0.0810 | −0.0043 | 0.0189 | 0.0194 | −0.0622 | 0.1312 | 0.1452 | −0.1171 | 0.0849 | 0.1446 |
| 0.999 | 0.99 | 0.75 | 0.95 | −4985 | −0.0003 | 0.0017 | 0.0017 | −0.0067 | 0.0246 | 0.0255 | −0.0463 | 0.0865 | 0.0981 | −0.0415 | 0.1160 | 0.1232 |
| | | 0.85 | 0.85 | −5649 | −0.0014 | 0.0142 | 0.0143 | −0.0061 | 0.0277 | 0.0283 | −0.0134 | 0.0414 | 0.0435 | 0.0595 | 0.1418 | 0.1538 |
| | | 0.95 | 0.75 | −7406 | −0.0078 | 0.0367 | 0.0375 | −0.0204 | 0.0701 | 0.0731 | −0.0106 | 0.0231 | 0.0254 | −0.0875 | 0.1437 | 0.1682 |
| 0.99 | 0.99 | 0.75 | 0.95 | −7995 | −0.0014 | 0.0030 | 0.0033 | −0.0010 | 0.0030 | 0.0031 | −0.2376 | 0.1726 | 0.2936 | −0.0095 | 0.0248 | 0.0265 |
| | | 0.85 | 0.85 | −7360 | −0.0011 | 0.0046 | 0.0047 | −0.0009 | 0.0039 | 0.0040 | −0.0443 | 0.0918 | 0.1020 | 0.0844 | 0.0578 | 0.1023 |
| | | 0.95 | 0.75 | −7847 | −0.0070 | 0.0214 | 0.0225 | −0.0064 | 0.0232 | 0.0241 | −0.0218 | 0.0546 | 0.0588 | −0.1233 | 0.1307 | 0.1797 |
| 0.999 | 0.999 | 0.75 | 0.95 | −7886 | −0.0008 | 0.0065 | 0.0065 | −0.0003 | 0.0011 | 0.0011 | −0.0644 | 0.1327 | 0.1475 | −0.0051 | 0.0246 | 0.0251 |
| | | 0.85 | 0.85 | −7305 | −0.0004 | 0.0013 | 0.0014 | −0.0004 | 0.0011 | 0.0011 | −0.0104 | 0.0532 | 0.0542 | 0.0951 | 0.0431 | 0.1044 |
| | | 0.95 | 0.75 | −7820 | −0.0011 | 0.0035 | 0.0037 | −0.0010 | 0.0037 | 0.0038 | −0.0126 | 0.0407 | 0.0426 | −0.1111 | 0.0994 | 0.1491 |
| *Panel C: Tempered stable distribution* | | | | | | | | | | | | | | | | |
| 0.99 | 0.999 | 0.75 | 0.95 | −10,700 | −0.0059 | 0.0159 | 0.0169 | −0.0004 | 0.0023 | 0.0023 | −0.0487 | 0.1292 | 0.1381 | −0.0039 | 0.0151 | 0.0156 |
| | | 0.85 | 0.85 | −8905 | −0.0066 | 0.0243 | 0.0252 | −0.0014 | 0.0078 | 0.0080 | −0.0267 | 0.1268 | 0.1296 | 0.0938 | 0.0278 | 0.0979 |
| | | 0.95 | 0.75 | −8277 | −0.0223 | 0.0865 | 0.0894 | −0.0039 | 0.0166 | 0.0171 | −0.0377 | 0.0783 | 0.0869 | −0.1179 | 0.0722 | 0.1382 |
| 0.999 | 0.99 | 0.75 | 0.95 | −4983 | −0.0002 | 0.0011 | 0.0011 | −0.0062 | 0.0190 | 0.0200 | −0.0318 | 0.0647 | 0.0721 | −0.0407 | 0.1108 | 0.1181 |
| | | 0.85 | 0.85 | −5646 | −0.0002 | 0.0011 | 0.0012 | −0.0035 | 0.0143 | 0.0148 | −0.0054 | 0.0369 | 0.0373 | 0.0528 | 0.1500 | 0.1590 |
| | | 0.95 | 0.75 | −7402 | −0.0034 | 0.0127 | 0.0131 | −0.0293 | 0.0986 | 0.1028 | −0.0083 | 0.0190 | 0.0207 | −0.0400 | 0.1469 | 0.1522 |
| 0.99 | 0.99 | 0.75 | 0.95 | −7997 | −0.0014 | 0.0032 | 0.0035 | −0.0012 | 0.0031 | 0.0033 | −0.1350 | 0.1267 | 0.1852 | −0.0084 | 0.0222 | 0.0237 |
| | | 0.85 | 0.85 | −7358 | −0.0010 | 0.0047 | 0.0048 | −0.0012 | 0.0058 | 0.0060 | −0.0332 | 0.0872 | 0.0933 | 0.0822 | 0.0581 | 0.1006 |
| | | 0.95 | 0.75 | −7845 | −0.0046 | 0.0167 | 0.0173 | −0.0057 | 0.0180 | 0.0189 | −0.0164 | 0.0342 | 0.0379 | −0.1243 | 0.1137 | 0.1685 |
| 0.999 | 0.999 | 0.75 | 0.95 | −7884 | −0.0010 | 0.0106 | 0.0107 | −0.0003 | 0.0011 | 0.0011 | −0.0394 | 0.0937 | 0.1016 | −0.0056 | 0.0242 | 0.0248 |
| | | 0.85 | 0.85 | −7303 | −0.0004 | 0.0014 | 0.0014 | −0.0004 | 0.0011 | 0.0011 | −0.0077 | 0.0492 | 0.0498 | 0.0946 | 0.0435 | 0.1041 |
| | | 0.95 | 0.75 | −7818 | −0.0011 | 0.0046 | 0.0047 | −0.0012 | 0.0047 | 0.0048 | −0.0122 | 0.0383 | 0.0402 | −0.1089 | 0.0957 | 0.1450 |

Finally, RMSE is a combination of bias and SE, which is employed as the overall performance indicator in the existing simulation studies. Consistent with the above results, MRS-GARCH-G is the least preferred model in all cases. RMSEs of MRS-GARCH-S are very close to those of MRS-GARCH-t for most sets.

#### 4.1.2. Estimates of volatility persistence

The difference between models in the estimates of volatility persistence is much smaller than that of transition probabilities. However, MRS-GARCH-G still cannot perform as well as MRS-GARCH-t or MRS-GARCH-S. More specifically, in most sets for both $P_1$ and $P_2$, it leads to the largest bias and SE. For instance, almost all absolute biases of $P_1$ obtained from MRS-GARCH-G are greater than 0.05. As to the other two models, most of the values are around 0.03. In terms of RMSE, MRS-GARCH-G has greater values than MRS-GARCH-S for all the twelve sets of $P_1$. Although MRS-GARCH-G performs relatively better for $P_2$, it is still not preferred in eight cases, compared with MRS-GARCH-S. The results of MRS-GARCH-S are overall quite close to those of MRS-GARCH-t. It is worth noticing that in quite a few cases, MRS-GARCH-S is even preferred to the true model. For instance, eight out of twelve RMSEs of $P_1$ obtained from MRS-GARCH-S are smaller than those of MRS-GARCH-t.

In conclusion, for estimates of both transition probabilities and volatility persistence, MRS-GARCH-S model outperforms MRS-GARCH-G in almost all cases. Further, results of MRS-GARCH-S model are overall very close to those of the true model.

### 4.2. GED

Next, we set the true distribution as GED with 1.1 degree of freedom.[10] Twelve sets of simulations with the same combinations of parameters as those in Section 4.1 are constructed. Replicates and each simulation are also truncated in the same manners to avoid simulation bias.

Simulation results are reported in Table 2. As a preliminary evidence of model performance, log-likelihood results are consistent with those presented in Section 4.1. More specifically, MRS-GARCH-t cannot perform as well as the true model in all cases. The results of MRS-GARCH-S are overall close to those of MRS-GARCH-G. In 10 out of the 12 sets, MRS-GARCH-S even generates higher log-likelihood than the true model.

#### 4.2.1. Estimates of transition probabilities

In the case of bias comparison, MRS-GARCH-t leads to a greater value than both MRS-GARCH-G and MRS-GARCH-S in all cases of $p_{11}$ and $p_{22}$. Relatively speaking, its performance in estimating $p_{11}$ is even worse, where one bias exceeds 0.5. As to MRS-GARCH-S, it consistently generates similar results as the true model in all sets of $p_{11}$ and $p_{22}$. The largest difference between those results is only 0.0044, where the absolute bias of MRS-GARCH-S is actually smaller than that of MRS-GARCH-G.

[10] We also consider the cases of 1.3, 1.5 and 1.8 degrees of freedom. The results are robust and available upon request.

The efficiency comparison has the same conclusion. MRS-GARCH-t is the least preferred model in all cases for both $p_{11}$ and $p_{22}$. MRS-GARCH-S and MRS-GARCH-G perform quite similarly. Therefore, according to RMSE, MRS-GARCH-S outperforms MRS-GARCH-t in all sets. Its results are consistently close to those of the true model. The largest difference of RMSE between them is only 0.0297, while the rest are mostly smaller than 0.01.

#### 4.2.2. Estimates of volatility persistence

The difference between biases of the estimated volatility persistence is mixed among the three models. For $P_1$, MRS-GARCH-t is not preferred to MRS-GARCH-S in all cases, though it has smaller absolute values than MRS-GARCH-G in four sets. In terms of $P_2$, in four out of twelve combinations, MRS-GARCH-t has smaller absolute biases than either MRS-GARCH-S or MRS-GARCH-G. The difference between MRS-GARCH-S and the true model is still small in most cases. It is worth noticing that in all sets of $P_1$ and eight sets of $P_2$, MRS-GARCH-S even outperforms MRS-GARCH-G.

The comparison of SE is similar to that of bias. In vast majority of the simulations, MRS-GARCH-t leads to the least efficient estimates for $P_1$. The difference between SES for $P_2$ is much smaller in most cases. MRS-GARCH-S generates more efficient estimates than the true model in all sets of $P_1$ and seven sets of $P_2$. Finally, the overall performance indicator RMSE still suggests that MRS-GARCH-t is the least preferred model to estimate either $P_1$ or $P_2$ in most cases. MRS-GARCH-S outperforms the true model in all sets of $P_1$ and eleven out of twelve sets of $P_2$. Overall, the difference between estimates of MRS-GARCH-S and MRS-GARCH-G is very small.

In conclusion, for estimates of both transition probabilities and volatility persistence, MRS-GARCH-S model outperforms MRS-GARCH-t in most cases. Further, results of MRS-GARCH-S model are overall close to those of the true model. It is also worth noticing that MRS-GARCH-S even outperforms MRS-GARCH-G in the estimation of volatility persistence in most combinations.

### 4.3. Tempered stable distribution

Finally, we set the true distribution as tempered stable with three sets of different parameters (all the $p$ are set to 0.5), including one case of CGMY distribution (when $\alpha^+ = \alpha^- = 0.5$ and $\lambda^+ = \lambda^- = 1.0$) and two general cases. Altogether, twelve sets of simulations are constructed. $\omega_1$, $\omega_2$, $\alpha_1$, $\alpha_2$, $\beta_1$ and $\beta_2$ are set to 0.1, 0.5, 0.1,0.2, 0.65 and 0.75, respectively, in all cases. All the sample sizes are still 5000. Replicates and each simulation are further truncated in the same manners as in Sections 4.1 and 4.2 to avoid simulation bias.

Simulation results are reported in Table 3. Not surprisingly, log-likelihood values of both MRS-GARCH-t and MRS-GARCH-G are all smaller than those of the true model. Also, MRS-GARCH-G outperforms MRS-GARCH-t only in the CGMY distribution cases, while MRS-GARCH-t generates lager log-likelihood in the other general cases.

**Table 3**
Simulation results: tempered stable distribution.

| $p_{11}$ | $p_{22}$ | $a_1$ | $a_2$ | $\lambda_1$ | $\lambda_2$ | $Mean_{ll}$ | $Bias_{p_{11}}$ | $SE_{p_{11}}$ | $RMSE_{p_{11}}$ | $Bias_{p_{22}}$ | $SE_{p_{22}}$ | $RMSE_{p_{22}}$ | $Bias_{P_1}$ | $SE_{P_1}$ | $RMSE_{P_1}$ | $Bias_{P_2}$ | $SE_{P_2}$ | $RMSE_{P_2}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *Panel A: Student's t-distribution* | | | | | | | | | | | | | | | | | | |
| 0.99 | 0.999 | 0.5 | 0.5 | 1.0 | 1.0 | −10,652 | −0.0192 | 0.0931 | 0.0950 | −0.0050 | 0.0354 | 0.0358 | −0.1179 | 0.2285 | 0.2571 | 0.0217 | 0.0151 | 0.0265 |
| | | 0.2 | 1.2 | 1.2 | 0.2 | −9724 | −0.0204 | 0.0894 | 0.0917 | −0.0136 | 0.0746 | 0.0759 | −0.0607 | 0.2059 | 0.2146 | 0.0295 | 0.0131 | 0.0323 |
| | | 1.2 | 0.2 | 0.2 | 1.2 | −9756 | −0.0342 | 0.1300 | 0.1345 | −0.0212 | 0.0922 | 0.0946 | −0.0628 | 0.2170 | 0.2259 | 0.0315 | 0.0129 | 0.0340 |
| 0.999 | 0.99 | 0.5 | 0.5 | 1.0 | 1.0 | −5012 | −0.0025 | 0.0233 | 0.0234 | −0.0085 | 0.0395 | 0.0404 | −0.0357 | 0.1096 | 0.1152 | −0.0114 | 0.1118 | 0.1124 |
| | | 0.2 | 1.2 | 1.2 | 0.2 | −4419 | −0.0160 | 0.0866 | 0.0881 | −0.0171 | 0.0755 | 0.0774 | 0.0004 | 0.0970 | 0.0970 | 0.0108 | 0.0843 | 0.0850 |
| | | 1.2 | 0.2 | 0.2 | 1.2 | −4422 | −0.0207 | 0.0965 | 0.0987 | −0.0281 | 0.1117 | 0.1152 | 0.0017 | 0.0980 | 0.0980 | 0.0051 | 0.1029 | 0.1030 |
| 0.99 | 0.99 | 0.5 | 0.5 | 1.0 | 1.0 | −7928 | −0.0020 | 0.0033 | 0.0038 | −0.0018 | 0.0030 | 0.0035 | −0.2244 | 0.1775 | 0.2862 | 0.0244 | 0.0214 | 0.0324 |
| | | 0.2 | 1.2 | 1.2 | 0.2 | −7203 | −0.0019 | 0.0036 | 0.0041 | −0.0015 | 0.0030 | 0.0033 | −0.1407 | 0.1606 | 0.2135 | 0.0359 | 0.0170 | 0.0397 |
| | | 1.2 | 0.2 | 0.2 | 1.2 | −7210 | −0.0021 | 0.0034 | 0.0040 | −0.0018 | 0.0029 | 0.0035 | −0.1463 | 0.1616 | 0.2180 | 0.0365 | 0.0169 | 0.0402 |
| 0.999 | 0.999 | 0.5 | 0.5 | 1.0 | 1.0 | −7714 | −0.0018 | 0.0253 | 0.0253 | −0.0022 | 0.0321 | 0.0322 | −0.0327 | 0.1302 | 0.1342 | 0.0204 | 0.0263 | 0.0333 |
| | | 0.2 | 1.2 | 1.2 | 0.2 | −7131 | −0.0039 | 0.0416 | 0.0418 | −0.0034 | 0.0358 | 0.0360 | −0.0167 | 0.1247 | 0.1258 | 0.0306 | 0.0378 | 0.0487 |
| | | 1.2 | 0.2 | 0.2 | 1.2 | −6970 | −0.0028 | 0.0346 | 0.0348 | −0.0013 | 0.0147 | 0.0148 | −0.0022 | 0.1046 | 0.1046 | 0.0369 | 0.0179 | 0.0410 |
| *Panel B: GED distribution* | | | | | | | | | | | | | | | | | | |
| 0.99 | 0.999 | 0.5 | 0.5 | 1.0 | 1.0 | −10,639 | −0.0047 | 0.0149 | 0.0156 | −0.0025 | 0.0396 | 0.0397 | −0.1202 | 0.2199 | 0.2507 | −0.0070 | 0.0172 | 0.0186 |
| | | 0.2 | 1.2 | 1.2 | 0.2 | −9745 | −0.0104 | 0.0527 | 0.0538 | −0.0069 | 0.0545 | 0.0549 | −0.0934 | 0.2441 | 0.2614 | −0.0255 | 0.0205 | 0.0327 |
| | | 1.2 | 0.2 | 0.2 | 1.2 | −9779 | −0.0064 | 0.0314 | 0.0321 | −0.0022 | 0.0311 | 0.0312 | −0.1211 | 0.2421 | 0.2707 | −0.0233 | 0.0176 | 0.0291 |
| 0.999 | 0.99 | 0.5 | 0.5 | 1.0 | 1.0 | −4999 | −0.0029 | 0.0193 | 0.0195 | −0.0204 | 0.0720 | 0.0749 | −0.0549 | 0.0928 | 0.1078 | −0.0526 | 0.1380 | 0.1477 |
| | | 0.2 | 1.2 | 1.2 | 0.2 | −4435 | −0.0382 | 0.1299 | 0.1354 | −0.0569 | 0.1462 | 0.1569 | −0.0750 | 0.1130 | 0.1357 | −0.0673 | 0.1258 | 0.1426 |
| | | 1.2 | 0.2 | 0.2 | 1.2 | −4440 | −0.0300 | 0.1071 | 0.1112 | −0.0633 | 0.1578 | 0.1700 | −0.0748 | 0.1102 | 0.1332 | −0.0775 | 0.1500 | 0.1688 |
| 0.99 | 0.99 | 0.5 | 0.5 | 1.0 | 1.0 | −7917 | −0.0016 | 0.0029 | 0.0033 | −0.0013 | 0.0029 | 0.0032 | −0.2437 | 0.1742 | 0.2996 | −0.0128 | 0.0228 | 0.0261 |
| | | 0.2 | 1.2 | 1.2 | 0.2 | −7223 | −0.0023 | 0.0033 | 0.0040 | −0.0018 | 0.0031 | 0.0036 | −0.2117 | 0.1751 | 0.2747 | −0.0308 | 0.0271 | 0.0411 |
| | | 1.2 | 0.2 | 0.2 | 1.2 | −7228 | −0.0025 | 0.0032 | 0.0041 | −0.0021 | 0.0032 | 0.0038 | −0.1899 | 0.1658 | 0.2521 | −0.0313 | 0.0266 | 0.0411 |
| 0.999 | 0.999 | 0.5 | 0.5 | 1.0 | 1.0 | −7699 | −0.0003 | 0.0012 | 0.0012 | −0.0003 | 0.0012 | 0.0013 | −0.0531 | 0.1321 | 0.1423 | −0.0119 | 0.0426 | 0.0442 |
| | | 0.2 | 1.2 | 1.2 | 0.2 | −7153 | −0.0024 | 0.0320 | 0.0320 | −0.0018 | 0.0163 | 0.0164 | −0.0626 | 0.1237 | 0.1386 | −0.0311 | 0.0480 | 0.0571 |
| | | 1.2 | 0.2 | 0.2 | 1.2 | −6992 | −0.0021 | 0.0255 | 0.0256 | −0.0023 | 0.0306 | 0.0307 | −0.0608 | 0.1232 | 0.1374 | −0.0262 | 0.0328 | 0.0419 |
| *Panel C: Tempered stable distribution* | | | | | | | | | | | | | | | | | | |
| 0.99 | 0.999 | 0.5 | 0.5 | 1.0 | 1.0 | −10,630 | −0.0049 | 0.0145 | 0.0153 | −0.0002 | 0.0008 | 0.0008 | −0.0844 | 0.1789 | 0.1978 | −0.0033 | 0.0155 | 0.0158 |
| | | 0.2 | 1.2 | 1.2 | 0.2 | −9663 | −0.0049 | 0.0149 | 0.0157 | −0.0003 | 0.0009 | 0.0009 | −0.0687 | 0.1817 | 0.1943 | −0.0092 | 0.0178 | 0.0200 |
| | | 1.2 | 0.2 | 0.2 | 1.2 | −9696 | −0.0055 | 0.0154 | 0.0164 | −0.0004 | 0.0013 | 0.0013 | −0.0652 | 0.1912 | 0.2020 | −0.0065 | 0.0193 | 0.0204 |
| 0.999 | 0.99 | 0.5 | 0.5 | 1.0 | 1.0 | −4990 | −0.0002 | 0.0009 | 0.0009 | −0.0059 | 0.0169 | 0.0179 | −0.0416 | 0.0877 | 0.0971 | −0.0405 | 0.1197 | 0.1264 |
| | | 0.2 | 1.2 | 1.2 | 0.2 | −4358 | −0.0002 | 0.0008 | 0.0008 | −0.0042 | 0.0113 | 0.0121 | −0.0266 | 0.0719 | 0.0766 | −0.0328 | 0.1002 | 0.1054 |
| | | 1.2 | 0.2 | 0.2 | 1.2 | −4363 | −0.0003 | 0.0011 | 0.0012 | −0.0057 | 0.0157 | 0.0167 | −0.0227 | 0.0694 | 0.0730 | −0.0332 | 0.1084 | 0.1134 |
| 0.99 | 0.99 | 0.5 | 0.5 | 1.0 | 1.0 | −7909 | −0.0014 | 0.0030 | 0.0033 | −0.0012 | 0.0030 | 0.0032 | −0.1957 | 0.1429 | 0.2424 | −0.0067 | 0.0224 | 0.0233 |
| | | 0.2 | 1.2 | 1.2 | 0.2 | −7146 | −0.0012 | 0.0032 | 0.0034 | −0.0009 | 0.0026 | 0.0028 | −0.1789 | 0.1514 | 0.2344 | −0.0076 | 0.0249 | 0.0260 |
| | | 1.2 | 0.2 | 0.2 | 1.2 | −7153 | −0.0015 | 0.0031 | 0.0034 | −0.0012 | 0.0027 | 0.0029 | −0.1577 | 0.1328 | 0.2062 | −0.0067 | 0.0267 | 0.0276 |
| 0.999 | 0.999 | 0.5 | 0.5 | 1.0 | 1.0 | −7690 | −0.0002 | 0.0011 | 0.0012 | −0.0004 | 0.0013 | 0.0014 | −0.0409 | 0.1180 | 0.1249 | −0.0078 | 0.0444 | 0.0451 |
| | | 0.2 | 1.2 | 1.2 | 0.2 | −7070 | −0.0005 | 0.0026 | 0.0027 | −0.0003 | 0.0013 | 0.0013 | −0.0416 | 0.1108 | 0.1183 | −0.0095 | 0.0260 | 0.0277 |
| | | 1.2 | 0.2 | 0.2 | 1.2 | −6908 | −0.0004 | 0.0034 | 0.0034 | −0.0004 | 0.0013 | 0.0014 | −0.0313 | 0.1017 | 0.1065 | −0.0083 | 0.0287 | 0.0299 |

*4.3.1. Estimates of transition probabilities*

In the case of bias comparison, both MRS-GARCH-t and MRS-GARCH-G cannot perform as well as the true model in almost all sets of both $p_{11}$ and $p_{22}$. The difference between those absolute biases is relatively small, and the largest value is around 0.05. The performance of MRS-GARCH-t and MRS-GARCH-G are roughly the same, where the former is preferred in five out of twelve cases for both transition probabilities.

The story of SE is slightly different. When $p_{11}$ and $p_{22}$ are both equal to 0.99, MRS-GARCH-t and MRS-GARCH-G can generate small SEs (at around 0.003) which are close to those of the true model. It is true for both $p_{11}$ and $p_{22}$. However, in the rest cases, those models lead to much larger values, most of which are greater than 0.03. For the true model, most SEs are less than 0.03, and the largest value is only 0.0169. Therefore, although the consistency performance of MRS-GARCH-t and MRS-GARCH-G is close to the true model, they are less efficient in most cases. Also, MRS-GARCH-t is less efficient than MRS-GARCH-G in the estimation of $p_{11}$ in most combinations. For $p_{22}$, results of both models are quite similar. Finally, RMSE suggests that the overall performance of both MRS-GARCH-t and MRS-GARCH-G in the estimation of the transition probabilities is not as well as MRS-GARCH-S in almost all cases. Also, the overall performance of MRS-GARCH-t is slightly worse than MRS-GARCH-G in the estimation of $p_{11}$ and is quite similar as MRS-GARCH-G in the estimation of $p_{22}$.

*4.3.2. Estimates of volatility persistence*

For $P_2$, MRS-GARCH-t and MRS-GARCH-G generally produce larger biases than the true model. Additionally, the absolute biases generated by the three models are mostly smaller than 0.04. However, MRS-GARCH-t can lead to smallest biases for most sets of $P_1$, and MRS-GARCH-G is the least preferred model in all cases. It is also worth noticing that the generated absolute biases of $P_1$ are overall greater than those of $P_2$. Some of them even exceed 0.2. Despite that, the consistency performance in the estimation of volatility persistence of the three models are quite close to each other.

In terms of the efficiency, the stories of $P_1$ and $P_2$ are also different. For $P_1$, SEs of MRS-GARCH-S are smaller than those of MRS-GARCH-t and MRS-GARCH-G in all cases. MRS-GARCH-G is more efficient than MRS-GARCH-t for the CGMY distribution, but is less efficient for the

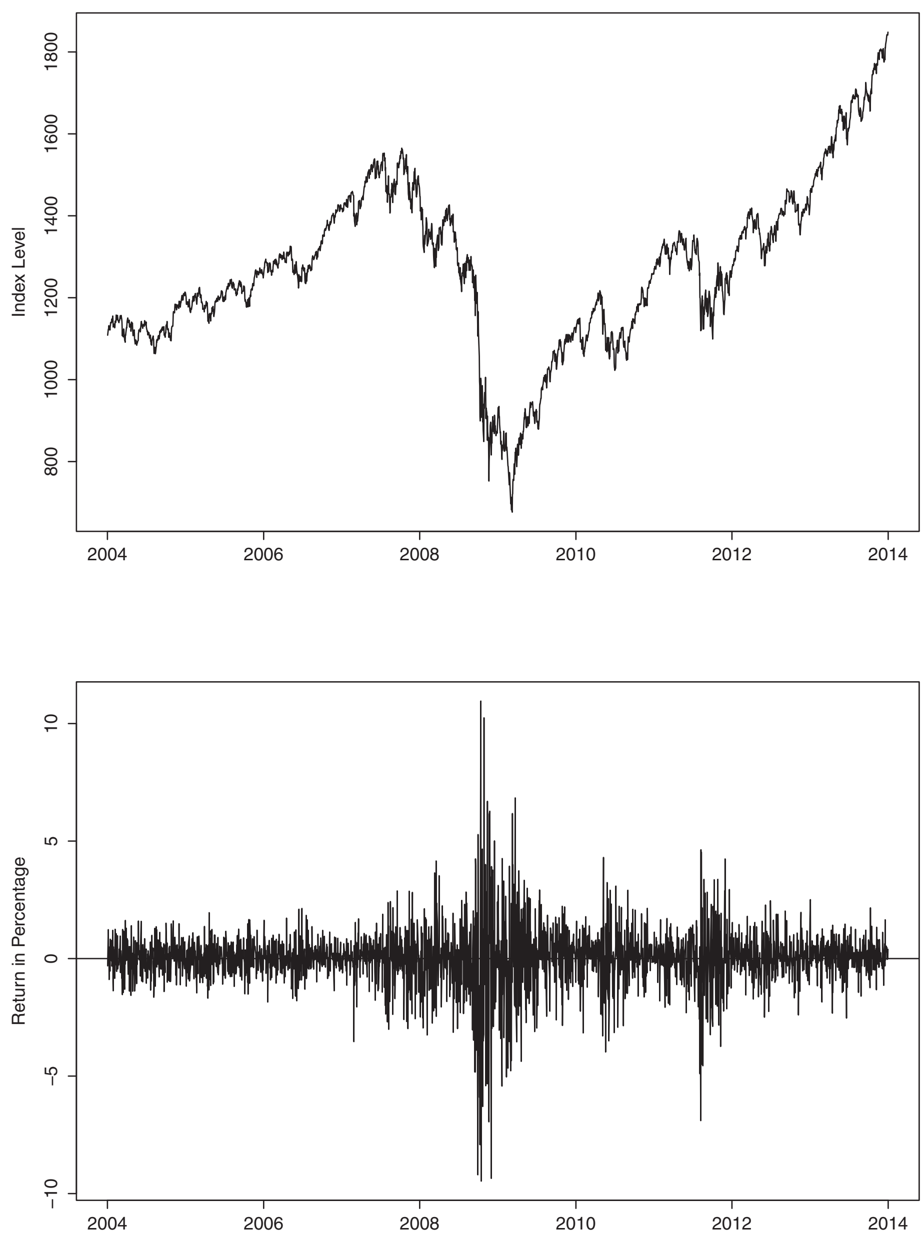


**Fig. 1.** S&P 500 daily index and return.

other general distributions. As to $P_2$, however, MRS-GARCH-t and generates smallest SEs, while MRS-GARCH-G is still the least efficient one in most cases. Besides, SEs of $P_1$ obtained from the three models are also generally greater than those of $P_2$. Most of SEs of $P_1$ are at around 0.15, while those of $P_2$ are smaller than 0.1. Finally, the overall performance indicator suggests that MRS-GARCH-S is still the preferred model in most combinations of $P_1$ and $P_2$. MRS-GARCH-t generally outperforms MRS-GARCH-G.

To summarize, when the true distribution is Student's t or GED, MRS-GARCH-S model consistently outperforms (in terms of RMSE) the competing model (MRS-GARCH-G or MRS-GARCH-t). Also, the results of MRS-GARCH-S and those of the true model are very close in most situations. Nevertheless, MRS-GARCH-S can even generate larger values of log-likelihood than the true model. When the true distribution is tempered stable, both the MRS-GARCH-t and MRS-GARCH-P cannot perform as well as the MRS-GARCH-S model. All the above observations are robust across different combinations of transition probabilities and volatility persistence. Therefore, we hereby argue that for a given financial time series with an unknown fat-tailed distribution of innovation, MRS-GARCH-S model should be employed to study the regime-switching conditional volatility.

## 5. Empirical results

To empirically compare MRS-GARCH models with Student's t, GED and tempered stable distributions, we work on the dataset of the daily S&P 500 Index (SP500). The daily closing prices for SP500 over the period from 1 January 2004 to 31 December 2013 are obtained from the Thomson Reuters Tick History (TRTH) database, which contains microsecond-time-stamped tick data dating back to January 1996. The database covers 35 million OTC and exchange-traded instruments worldwide, which are provided by the Securities Industry Research Centre of Australasia (SIRCA). The corresponding return in the percentage series is defined as the logarithm of the daily closing price differences times 100; that is, $r_t = 100 \times \log(S_t/S_{t-1})$.

The level and return of SP500 are plotted in Fig. 1. In the level plot, SP500 level clearly decreases between 2008 and 2009, which is the global financial crisis (GFC) period. Demonstrated in the return plot, from 2008 to 2010, the return becomes much more volatile. After 2010, the volatility tends to be smaller, with some turbulence around the end of 2010 and the beginning of 2012. Hence, with the presence of GFC and preliminary visual evidence, structural breaks may exist in the SP500 return. Consequently, as discussed in Section 2.2, MRS-GARCH model should be employed to study its volatility behavior instead of the GARCH model. In addition, the mean and standard error of the SP500 return are 0.0202 and 1.2876, respectively. The skewness is −0.3259, indicating that the SP500 return is slightly negatively skewed. The kurtosis is 11.0235, suggesting a non-Gaussian distribution. Thus, we perform the Kolmogorov–Smirnov and Jarque–Bera normality tests (not presented), where the null hypotheses indicating normality are rejected in both cases (p-values are 0.0000). As a result, as discussed in Section 3, the innovation of MRS-GARCH model should follow a fat-tailed distribution, rather than the Normal.

We fit the MRS-GARCH model with Student's t, GED and tempered stable distributions for the SP500 return. The results are presented in Table 4. Overall, all estimates are significant at 5% level in all models, and estimates of individual parameter from different models are quite close. Specifically, both $p_{11}$ and $p_{22}$ are greater than 0.99 in all models, suggesting that the expectations of staying at both the calm and turbulent states are quite long. $\alpha_1 + \beta_1$ are at around 0.92, and $\alpha_2 + \beta_2$ are at around 0.99. This indicates that volatility of SP500 return is remarkably persistent over the entire period in both states. Also, a shock will affect the volatility longer in the turbulent state than in the calm state. To compare the model performance, both log-likelihood and Akaike Information Criterion (AIC) are presented. It can be seen that MRS-GARCH-S is preferred to the other two specifications and MRS-GARCH-G outperforms MRS-GARCH-t.

All the three estimated conditional volatility series[11] and the smoothing probability of the calm state are plotted in Fig. 2. Despite their different model performance, the three estimated conditional volatility series have similar shapes. It could be due to the similarity of estimates. The trends of those series are consistent with our observation of the return series, since the conditional volatility is comparatively larger from 2008 to 2010. According to the smoothing probability, MRS-GARCH-t suggests SP500 return starts with the calm state and switches to the turbulent state from 2008. It stays there until the end of 2010, with a short jump back to the calm state. After that, SP500 return switches back to the calm state and stays there for most of the time, except for two short periods (the end of 2011 and mid-2012) back to the turbulent state. The results of both MRS-GARCH-G and MRS-GARCH-S are very similar and different from that of MRS-GARCH-t to some extent. Basically, they suggest SP500 return lies in the turbulent state from 2008 to 2011 and lies in the calm state during the other periods. Despite this difference, the three state structure are largely concordant with the real macro-economic situation: the 2008 GFC with an effect lasting for around 3 years.

To further explore the difference between models, we report the smoothed density plot of the standardized SP500 return in Fig. 3. To compare the fitness of the distributions, we also plot the densities of Student's t and GED distributions with 7.2742 and 1.3301 degrees of freedom, respectively. Those are estimates from the MRS-GARCH-t and MRS-GARCH-G models presented in Table 4. The density of tempered stable distribution with parameters as estimated in Table 4 is also reported. It is clear that the shape of the tempered stable density is closer to that of the standardized SP500 return than the others. This provides further evidence that the MRS-GARCH-S outperforms the other models for our dataset.

**Table 4**
Empirical results: S&P 500 index.

| | Std-t | GED | T.Stb |
|---|---|---|---|
| $p_{11}$ | 0.9967 | 0.9993 | 0.9992 |
| | (0.0000) | (0.0000) | (0.0000) |
| $p_{22}$ | 0.9936 | 0.9987 | 0.9989 |
| | (0.0000) | (0.0000) | (0.0000) |
| $\alpha_1$ | 0.0611 | 0.0560 | 0.0613 |
| | (0.0020) | (0.0620) | (0.0310) |
| $\alpha_2$ | 0.0877 | 0.1064 | 0.1072 |
| | (0.0000) | (0.0000) | (0.0000) |
| $\beta_1$ | 0.8464 | 0.8564 | 0.8985 |
| | (0.0000) | (0.0000) | (0.0000) |
| $\beta_2$ | 0.9035 | 0.8879 | 0.8877 |
| | (0.0000) | (0.0000) | (0.0000) |
| $v$ | 7.2742 | 1.3301 | |
| | (0.0000) | (0.0000) | |
| $a_1$ | | | −0.1477 |
| | | | (0.0000) |
| $a_2$ | | | −1.8164 |
| | | | (0.0000) |
| $\lambda_1$ | | | 2.7876 |
| | | | (0.0000) |
| $\lambda_2$ | | | 3.1898 |
| | | | (0.0000) |
| $p$ | | | 0.4407 |
| | | | (0.0000) |
| *log. lik* | −3409 | −3395 | −3382 |
| *AIC* | 6839 | 6813 | 6793 |

## 6. Conclusion

MRS-GARCH model has enjoyed particular popularity in the financial study with the presence of structural breaks. However, the original

[11] We report conditional volatility here as the square root of $h_t$, so that it has the same scale as $r_t$.

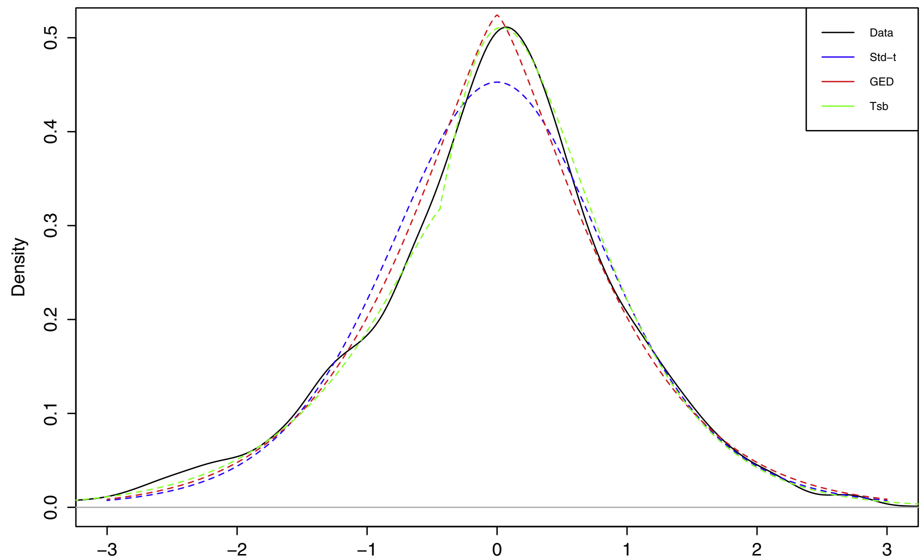


**Fig. 2.** Density plots of S&P 500 return and choices of distributions.

MRS-GARCH model is based on the Normal distribution, and its estimates will be inconsistent for fat-tail-distributed data. As the financial data in practice are rarely Gaussian, the seek of an appropriate distribution to accommodate their excess kurtosis becomes an essential issue for the application of the MRS-GARCH model. As widely used alternatives of the Normal distribution, the Student's t and GED are the common choices in most existing literature. A recent study by Calzolari et al. (2014) points out that due to the instability under aggregation, the Student's t is not an optimal choice. To overcome that, the $\alpha$-stable distribution is introduced. However, the undefined second moment of it brings in even more serious problems to be applied in the GARCH-type model.

To solve this issue, this paper proposes that the tempered stable distribution should be used instead. The reason is that this distribution retains all the attractive properties of the $\alpha$-stable distribution and has defined moments. Using three different simulation studies of the MRS-GARCH(1,1) process, we systematically demonstrate the appropriateness of the tempered stable distribution applied in the MRS-GARCH model. The first two studies assume that the true distributions are the Student's t and GED, respectively. In these two studies, results of MRS-GARCH-S are close to those of the true models. Additionally, MRS-GARCH-S generally outperforms its competitor (the other model except the true model) in terms of consistency, efficiency and overall performance. We construct different combinations of the underlying tempered stable distribution in the third study. Our results suggest that either MRS-GARCH-t or MRS-GARCH-G cannot perform as well as the MRS-GARCH-S model.

Finally, empirical evidence is further provided as the robustness check of our simulation results. We fit the daily return of the S&P 500 index into the three MRS-GARCH models, respectively. Our results suggest that MRS-GARCH-S is still preferred to the others. The density of fitted tempered stable distribution is closer to that of the standardized

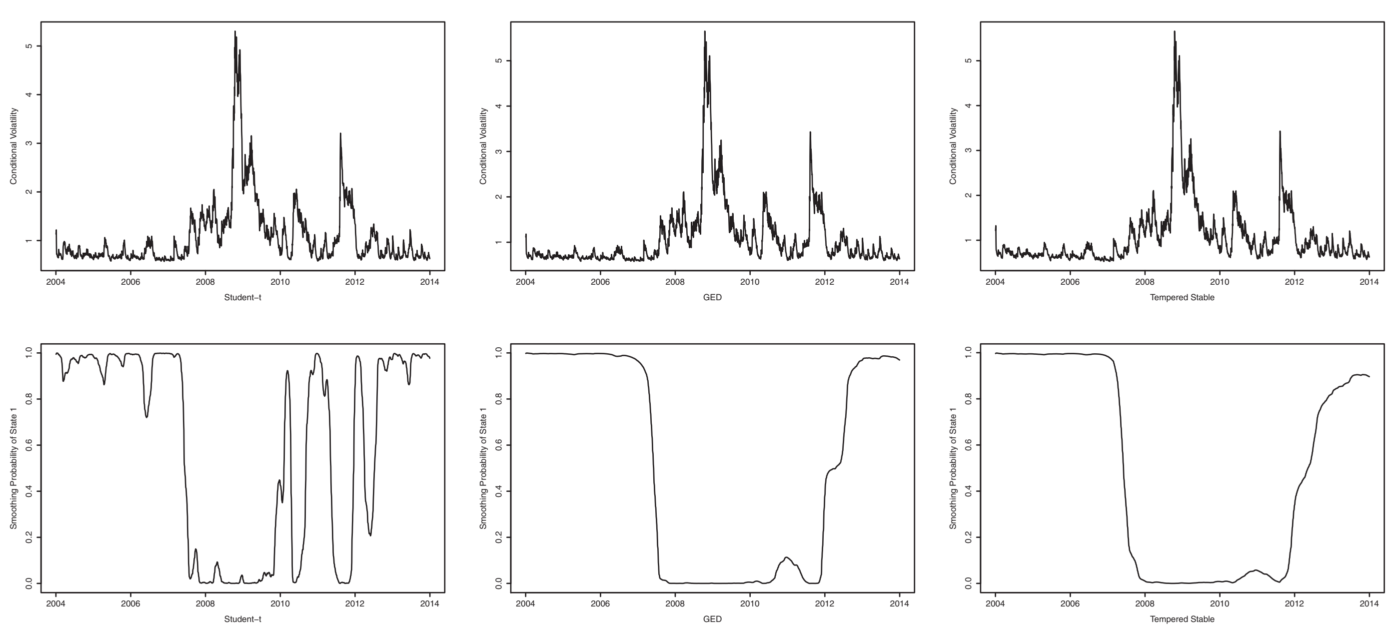


**Fig. 3.** Fitted conditional volatility and smoothing probability.

data than the other two. Hence, we argue that the tempered stable distribution could be a widely useful tool for modeling the financial volatility in general contexts with a MRS-GARCH-type specification.

## Acknowledgments

We are grateful to the ANU College of Business and Economics and Research School of Finance, Actuarial Studies and Statistics and Jiangxi University of Finance and Economics for their financial support. The authors would also like to thank Dave Allen, Felix Chan, Michael McAleer, Paresh Narayan, Morten Nielson, Albert Tsui, Zhaoyong Zhang, and participants at the 1st Conference on Recent Developments in Financial Econometrics and Applications, the 8th China R Conference (Nanchang), ANU Research School Brown Bag Seminar, Central University of Finance and Economics Seminar, China Meeting of Econometric Society, Chinese Economists Society China Annual Conference, Econometric Society Australasian Meeting, Ewha Woman's University Department of Economics Seminar, International Congress of the Modeling and Simulation Society of Australia and New Zealand and Shandong University Seminar. The usual disclaimer applies.

## References

Ang, A., Timmermann, A., 2012. Regime Changes and Financial Markets. Annu. Rev. Financ. Econ. 4 (1), 313–337.
Ardia, D., 2009. Bayesian estimation of a Markov-switching threshold asymmetric garch model with Student-t innovations. Econ. J. 12 (1), 105–126.
Bianchi, M.L., Rachev, S.T., Kim, Y.S., Fabozzi, F.J., 2010. Tempered stable distributions and processes in finance: numerical analysis. Mathematical and Statistical Methods for Actuarial Sciences and Finance. Springer, Italy, pp. 33–42.
Bollerslev, T., 1986. Generalized autoregressive conditional heteroskedasticity. J. Econ. 31 (3), 307–327.
Bollerslev, T., 1987. A conditional heteroskedastic time series model for speculative prices and rates of return. Rev. Econ. Stat. 69 (3), 542–547.
Bollerslev, T., Wooldridge, J.M., 1992. Quasi-maximum likelihood estimation and inference in dynamic models with time-varying covariances. Econ. Rev. 11 (2), 143–172.
Cai, J., 1994. A markov model of switching-regime arch. J. Bus. Econ. Stat. 12 (3), 309–316.
Calzolari, G., Halbleib, R., Parrini, A., 2014. Estimating garch-type models with symmetric stable innovations: indirect inference versus maximum likelihood. Comput. Stat. Data Anal. 76, 158–171.
Carr, P., Geman, H., Madan, D.B., Yor, M., 2002. The fine structure of asset returns: an empirical investigation*. J. Bus. 75 (2), 305–333.
Chkili, W., Aloui, C., Nguyen, D.K., 2012. Asymmetric effects and long memory in dynamic volatility relationships between stock returns and exchange rates. J. Int. Financ. Mark. Inst. Money 22 (4), 738–757.
Constantinides, A., Savel'ev, S.E., 2013. Modelling price dynamics: a hybrid truncated lévy flight¨cgarch approach. Phys. A 392 (9), 2072–2078.
Cont, R., Tankov, P., 2004. Financial Modelling with Jump Processes. Chapman and Hall/CRS Press, United Kingdom.
Dueker, M.J., 1997. Markov switching in garch processes and mean-reverting stock-market volatility. J. Bus. Econ. Stat. 15 (1).
Engle, R.F., 1982. Autoregressive conditional heteroskedasticity with estimates of variance of united kingdom inflation. Economerica 50 (4), 987–1007.
Fan, Y., Zhang, Y.-J., Tsai, H.-T., Wei, Y.-M., 2008. Estimating 'value at risk' of crude oil price and its spillover effect using the ged-garch approach. Energy Econ. 30 (6), 3156–3171.
Franses, P.H., Van Dijk, D., 1996. Forecasting stock volatility using (non-linear) GARCH models. J. Forecast. 15 (1996), 229–235.
Gray, S.F., 1996. Modeling the conditional distribution of interest rates as a regime-switching process. J. Financ. Econ. 42 (1), 27–62.
Haas, M., 2009. Value-at-risk via mixture distributions reconsidered. Appl. Math. Comput. 215 (6), 2103–2119.
Haas, M., Paolella, M.S., 2012. Mixture and regime-switching garch models. In: Bauwens, L., Hafner, C., Laurent, S. (Eds.), Handbook of Volatility Models and Their Applications. Wiley, United Kingdom, pp. 71–102.
Haas, M., Mittnik, S., Paolella, M.S., 2004. A new approach to markov-switching garch models. J. Financ. Econ. 2 (4), 493–530.
Hamilton, J.D., 1988. Rational-expectations econometric analysis of changes in regime: an investigation of the term structure of interest rates. J. Econ. Dyn. Control. 12 (2–3), 385–423.
Hamilton, J.D., 1989. A new approach to the economic analysis of nonstationary time series and the business cycle. Econometrica 57 (2), 357–384.
Hamilton, J.D., Susmel, R., 1994. Autoregressive conditional heteroskedasticity and changes in regime. J. Econ. 64 (1), 307–333.
Ho, K.-Y., Shi, Y., Zhang, Z., 2013. How does news sentiment impact asset volatility? Evidence from long memory and regime-switching approaches. N. Am. J. Econ. Financ. 26, 436–456.
Kim, Y.S., Rachev, S.T., Bianchi, M.L., Fabozzi, F.J., 2008. Financial market models with lévy processes and time-varying volatility. J. Bank. Financ. 32 (7), 1363–1378.
Klaassen, F., 2002. Improving garch volatility forecasts with regime-switching garch. Empir. Econ. 27 (2), 363–394.
Koponen, I., 1995. Analytic approach to the problem of convergence of truncated lévy flights towards the gaussian stochastic process. Phys. Rev. E 52 (1), 1197.
Küchler, U., Tappe, S., 2013. Tempered stable distributions and processes. Stoch. Process. Appl. 123 (12), 4256–4293.
Lamoureux, C.G., Lastrapes, W.D., 1990. Heteroskedasticity in stock return data: volume versus garch effects. J. Financ. 45 (1), 221–229.
Lin, G., 1998. Nesting regime-switching garch models and stock market volatility, returns and the business cycle (PhD dissertation) University of California.
Mabrouk, S., Saadi, S., 2012. Parametric value-at-risk analysis: evidence from stock indices. Q. Rev. Econ. Finance 52 (3), 305–321.
Marcucci, J., 2005. Forecasting stock market volatility with regime-switching GARCH models. Stud. Nonlinear Dyn. Econ. 94 (2).
Mittnik, S., Doganoglu, T., Chenyao, D., 1999. Computing the probability density function of the stable paretian distribution. Math. Comput. Model. 29 (10), 235–240.
Mullen, K., Ardia, D., Gil, D., Windover, D., Cline, J., 2011. Deoptim: an r package for global optimization by differential evolution. J. Stat. Softw. 40 (6), 1–26.
Narayan, P.K., Narayan, S., 2007. Modelling oil price volatilit. Energy Policy 35 (12), 6549–6553.
Narayan, S., Narayan, P.K., 2011. Are shocks to national income persistent? New global evidence. J. Econ. Stud. 38 (2), 218–230.
Narayan, P.K., Sharma, S.S., 2014. Firm return volatility and economic gains: the role of oil prices. Econ. Model. 38, 142–151.
Shi, Y., Ho, K.-Y., 2015a. Long memory and regime switching: A simulation study on the Markov regime-switching ARFIMA model. J. Bank. Financ. 61 (S2), S189–S204.
Shi, Y., Ho, K.-Y., 2015b. Modeling high-frequency volatility with three-state figarch models. Econ. Model. 51, 473–483.
Stanley, H.E., Plerou, V., Gabaix, X., 2008. A statistical physics view of financial fluctuations: evidence for scaling and universality. Phys. A 387 (15), 3967–3981.
Susmel, R., Engle, R.F., 1994. Hourly volatility spillovers between international equity markets. J. Int. Money Financ. 13 (1), 3–25.
Westerlund, J., Narayan, P.K., 2012. Does the choice of estimator matter when forecasting returns? J. Bank. Financ. 36 (9), 2632–2640.
Yang, S.-R., Brorsen, B.W., 1993. Nonlinear dynamics of daily futures prices: conditional heteroskedasticity or chaos? J. Futur. Mark. 13 (2), 175–191.
Zhu, D., Galbraith, J.W., 2011. Modeling and forecasting expected shortfall with the generalized asymmetric student-t and asymmetric exponential power distributions. J. Empir. Financ. 18 (4), 765–778.