Lingbing Feng[1,2] / Yanlin Shi[3]

Markov regime-switching autoregressive model with tempered stable distribution: simulation evidence

[1] Institute of Industrial Economics, Jiangxi University of Finance and Economics, Nanchang, Jiangxi, China. https://orcid.org/0000-0002-7157-3792.
[2] International Institute for Financial Studies, Jiangxi University of Finance and Economics, Nanchang, Jiangxi, China. https://orcid.org/0000-0002-7157-3792.
[3] Department of Actuarial Studies and Business Analytics, Macquarie University, NSW 2109, Australia, Phone: +61 2 9850 4750, E-mail: yanlin.shi@mq.edu.au

Abstract:
Markov regime-switching (MRS) autoregressive model is a widely used approach to model the economic and financial data with potential structural breaks. The innovation series of such MRS-type models are usually assumed to follow a Normal distribution, which cannot accommodate fat-tailed properties commonly present in empirical data. Many theoretical studies suggest that this issue can lead to inconsistent estimates. In this paper, we consider the tempered stable distribution, which has the attractive stability under aggregation property missed in other popular alternatives like Student's t-distribution and General Error Distribution (GED). Through systematically designed simulation studies with the MRS autoregressive models, our results demonstrate that the model with tempered stable distribution uniformly outperforms those with Student's t-distribution and GED. Our empirical study on the implied volatility of the S&P 500 options (VIX) also leads to the same conclusions. Therefore, we argue that the tempered stable distribution could be widely used for modelling economic and financial data in general contexts with an MRS-type specification.

Keywords: fat-tailed distribution, regime-switching, tempered stable distribution

JEL classification: C22, C51, G11

DOI: 10.1515/snde-2018-0008

1 Introduction

Regime-switching models are widely employed to study various types of financial data such as interest rate, equity volatility and inflation rate. The popularity of those models is due to their attractive attributes that allow for the effects of structural breaks and business cycle (Hamilton 1988; 1994; Marcucci, 2005; Shi and Ho, 2015). Hence, more accurate estimates and predictions can be generated, which are essential to financial practice like derivative pricing.

As one of the most popular models accommodating structural breaks, the Markov regime-switching autoregressive (MRS) model is proposed in the seminal work of Hamilton (1989). It is originally developed to investigate the impacts of business cycle on financial time series. In this work, Hamilton (1989) argues that the behaviour of financial data tends to be dependent on the states (regimes) of business cycles. To take this impact into consideration, parameters of a general ARIMA model are allowed to switch between states. The MRS model is then constructed by assuming that the latent states follow a Markov chain. Over the past few decades, this model has become one of the standard approaches to model regime switching, structural breaks and business cycle (Diebold and Inoue, 2001).

The original MRS model is derived with the assumption that the innovation sequence follows a Normal (Gaussian) distribution. However, significant evidence suggests that the financial time series is rarely Gaussian but typically leptokurtic and exhibits fat-tail behaviour (Bollerslev, 1987; Susmel and Engle, 1994; Shi and Ho, 2015). For the ARIMA model, it is well-known that Maximum Likelihood Estimation (MLE) based on Gaussian distribution will always still lead to consistent estimators.[1] As for a time-homogeneous MRS-type models, Mevel and Finesso (2004) and Douc and Moulines (2012) comprehensively discuss asymptotic properties of the MLE of such models under misspecification. It is found that the consistency of estimates and the information

Yanlin Shi is the corresponding author.

matrix equality do not necessarily hold. A recent study by Pouzo, Psaradakis, and Sola (2016) extends the discussion to time-inhomogeneous MRS-type models and suggest that MLE is consistent only for the parameters of the model having the lowest Kullback–Leibler divergence from the data-generating process. In particular to the misspecified distribution assumption, 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). Haas and Paolella (2012) specifically argue that MLE of MRS-type model with Gaussian assumption is inconsistent when fitted to fat-tailed data. To support this, simulation evidence in the case of a GARCH specification is provided in Shi and Feng (2016). Consequently, the sought of an appropriate distribution to accommodate the excess kurtosis of the empirical economic and financial data becomes an essential issue to use the MRS-type models.

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) (Klaassen, 2002; Haas, Mittnik and Paolella, 2004; Marcucci, 2005). However, a recent study by Calzolari, Halbleib, and Parrini (2014) argues that those alternatives can be problematic. Their most serious drawback is that they lack in stability under aggregation, which is of particular importance in portfolio applications and risk management in finance study. As a replacement of them, the α-stable distribution is recommended by Calzolari, Halbleib, and Parrini (2014) to model the fat-tailed behaviour of financial time series. Unfortunately, the second moment of the α-stable distribution does not exist in usual cases. Hence, the application of it to the MRS model will lead to questionable interpretation and problematic statistical properties of the MLE.

The tempered stable distribution is a natural substitution of the α-stable (Feng and Shi, 2016; Shi and Feng, 2016). Firstly introduced[2] in Koponen (1995), tempered stable distribution covers a range of well-known subclasses like Variance Gamma, bilateral Gamma and CGMY distributions (Küchler and Tappe, 2013). Its most outstanding advantage is that the tempered stable distribution retains most of the attractive properties of the α-stable and has defined moments. Additionally, it has a flexible form which accommodates both the symmetric and asymmetric shapes of density.

In this paper, we employ the tempered stable distribution within the MRS framework and argue that it outperforms the commonly used Student's t and GED. To demonstrate that, we conduct a series of simulation studies to compare the performance [3] of the MRS models with the Normal, Student's t, GED and tempered stable distributions. First, we investigate two scenarios where the true distributions are Student's t and GED, respectively. Via twelve combinations of different parameter and sample size settings, the MRS models with four distinct distributions are systematically analysed. It is demonstrated that when the true distribution is Student's t or GED, the MRS model with tempered stable distribution generates very similar results to those of the true model. More importantly, it clearly outperforms the competitors (the MRS models with distributions other than the true density) in terms of biasness, efficiency and overall performance. Second, we set the tempered stable as the true distribution. Twelve sets of simulations are further constructed, including different choices of transition probabilities and tempered stable distribution parameters. In this scenario, none of the MRS models with Normal, Student's t and GED can perform as well as that with tempered stable distribution. Such findings are also observed for the correctiveness in identifying the underlying latent Markov states. Nevertheless, when comparing ratios of sampling standard error to estimated standard error, MRS models with three fat-tailed distributions exhibit similar behaviours across different combinations for all parameters except for the standard deviation. Therefore, we argue that the tempered stable distribution could be widely used for modelling financial data in general contexts with an MRS-type specification.

To empirically analyse the MRS model with different fat-tailed distributions, we work on the daily VIX index, which is the implied volatility of the S&P 500 options. Since the VIX index is skewed to some extent after taking the logarithm transformation, we consider and compare five cases: (skewed) Student's t, (skewed) GED and tempered stable distributions. The results suggest that the MRS model with tempered stable distribution outperforms that with all the other distributions, whereas MRS model with (skewed) Student's t-distribution is preferred to that with (skewed) GED. Besides, the estimated parameters across different distributions are relatively close to each other. The fitted smoothing probability series of the five MRS models also have similar general patterns, although identified states of the MRS models with (skewed) GED differ to some degree from those of the other three fitted models. Hence, our empirical result is robust with the simulation evidence, suggesting that the MRS model with tempered stable distribution is a useful tool in the practical financial study.

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

2 MRS model

Financial time series may behave differently depending on the states (regimes) of the business cycle (Hamilton 1989). For instance, the stock return tends to be less volatile in a peak state and be more volatile in a trough state (Ho, Shi, and Zhang 2013). To take this impact into consideration, the MRS model is proposed with state-dependent parameters (Hamilton 1988, 1989, and 1994)

Let $\{s_t\}$ be a stationary, irreducible Markov process with discrete state space $\{1, 2\}$ and transition matrix $P = [p_{jk}]$, where $p_{jk} = \mathrm{P}(s_{t+1} = k|s_t = j)$ is the one-step transition probability of moving from state j to state k $(j, k \in \{1, 2\})$. Then, one popular type of MRS model is

$$y_t = \mu_{s_t} + \gamma y_{t-1} + \varepsilon_t, \varepsilon_t = \eta_t \sqrt{\sigma_\varepsilon^2} \text{ and } \eta_t \overset{iid}{\sim} N(0, 1) \tag{1}$$

where s_t is the state of y_t at time t. γ is the autoregressive parameter. μ_{s_t} is the intercept (mean) parameter in state s_t. ε_t is the error at time t. η_t is an identical and independent innovation sequence following a Normal distribution, with 0 mean and unit standard deviation. In particular, it is constrained that $\mu_1 < \mu_2$, which indicates that the overall mean in state 2 is greater than that in state 1.

Although the MRS model can have more than two states, we only consider the two-state case in this paper for the following reasons. To the best of our knowledge, there is no effective test developed so far to determine the optimal number of states. Studies like Cho and White (2007) can only test the null hypothesis of one regime against the alternative of regime switching between two states. A recent research of Faias and Nunes (2012) comprehensively reviews the existing regime-switching tests, but none of them test the case of three or more states. Other literature like Carrasco, Hu, and Ploberger (2014) tests parameters of the MRS model rather than the number of states, which is irrelevant to this purpose. Furthermore, the classic trinity (Wald, likelihood ratio and Lagrange multiplier) tests are not applicable here, due to the fact that the MRS-type model with more states does not nest the one with fewer states (Wilfling 2009). On the other hand, although there are many existing tests examining the number and/or locations of the structural changes (such as the non-parametric test used in Ross et al. (2011)), it is not straightforward to extend them to the MRS cases. Another potential problem is that even if the optimal number of states can be tested, their interpretation might not be straightforward. Hamilton's (1989) original work suggests that the two-state case can be linked to the business cycle. For instance, low- and high-mean states can be referred to as the trough and peak periods, respectively. However, the interpretation for a large number of such states might not be meaningful.

The parameters of the MRS model can be estimated using MLE, and the conditional density of ε_t is given as follows.

$$\begin{aligned} \Omega_{t-1} &= \{\varepsilon_{t-1}, \varepsilon_{t-2}, \dots, \varepsilon_1\} \\ \theta &= (\mu_1, \mu_2, p_{11}, p_{22}, \gamma, \sigma_\varepsilon^2)' \\ f(\varepsilon_t|s_t = j, \theta, \Omega_{t-1}) &= f(\varepsilon_{s_t,t}|s_t = j, \theta, \Omega_{t-1}) \\ &= \frac{1}{\sqrt{2\pi\sigma_\varepsilon^2}} e^{-\frac{\varepsilon_{j,t}^2}{2\sigma_\varepsilon^2}} \end{aligned} \tag{2}$$

where Ω_{t-1} is the information set at time $t-1$. θ is the vector of parameters. $\Gamma(\cdot)$ is the Gamma function and $f(\varepsilon_t|s_t = j, \theta, \Omega_{t-1})$ is the conditional density of ε_t when s_t is known. Since the s_t is equal to j and thus no longer latent, we have that $\varepsilon_t = \varepsilon_{s_t,t} = y_t - \mu_j - \gamma y_{t-1}$, and hence $f(\varepsilon_t|s_t = j, \theta, \Omega_{t-1}) = f(\varepsilon_{s_t,t}|s_t = j, \theta, \Omega_{t-1})$. The specification of $f(\varepsilon_{s_t,t}|s_t = j, \theta, \Omega_{t-1})$ then stems from the fact that ε_t follows a Normal distribution with mean 0 and variance σ_ε^2 given the information at time $t-1$.

Plugging the filtered probability in state j at time $t-1$, $\omega_{j,t-1} = \mathrm{P}(s_{t-1} = j|\theta, \Omega_{t-1})$, into equation (2) and integrating out the state variable s_{t-1}, the density function in the equation (2) becomes

$$f(\varepsilon_{s_t,t}|\theta, \Omega_{t-1}) = \sum_{j=1}^{2}\sum_{k=1}^{2} p_{jk}\omega_{j,t-1} f(\varepsilon_{s_t,t}|s_t = j, \theta, \Omega_{t-1}). \tag{3}$$

$\omega_{j,t-1}$ can be obtained by an integrative algorithm given in Hamilton (1989). The log-likelihood function corresponding to equation (3) is as follows:

$$L(\theta|\varepsilon) = \sum_{t=2}^{T} \ln f(\varepsilon_{s_t,t}|\theta, \Omega_{t-1}) \text{ where } \varepsilon = \left(\varepsilon_{s_t,1}, \varepsilon_{s_t,2}, \dots, \varepsilon_{s_t,T}\right)', \tag{4}$$

and the ML estimator $\hat{\theta}$ is obtained by maximizing equation (4).

In order to identify which state y_t lies in at time t, we extract the smoothing probability of state 1 as follows (Hamilton 1988, 1989, and 1994):

$$P(s_t = 1|\theta, \Omega_T) = \omega_{1,t}\left[\frac{p_{11}P(s_{t+1} = 1|\theta, \Omega_T)}{P(s_{t+1} = 1|\theta, \Omega_t)} + \frac{p_{12}P(s_{t+1} = 2|\theta, \Omega_T)}{P(s_{t+1} = 2|\theta, \Omega_t)}\right], \tag{5}$$

Using the fact that $P(s_T = 1|\theta, \Omega_T) = \omega_{1,T}$, the smoothing probability series $P(s_t = 1|\theta, \Omega_T)$ can be generated by iterating equation (5) backward from T to 1. According to Hamilton (1989), it is expected that y_t lies in state 1 (2) at time t, if $P(s_t = 1|\theta, \Omega_t)$ is greater (smaller) than 0.5.

3 Alternative distributions for the innovation sequence

Although the original MRS 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 behaviour (Bollerslev, 1987; Susmel and Engle, 1994; Stanley, Plerou and Gabaix, 2008; Ho, Shi and Zhang, 2013). Mevel and Finesso (2004) and Douc and Moulines (2012) comprehensively discuss asymptotic properties of the MLE of MRS models under misspecification. It is found that the consistency of estimates and the information matrix equality do not necessarily hold. A recent study by Pouzo, Psaradakis, and Sola (2016) extends the discussion to time-inhomogeneous MRS-type models. In particular, as noted by Klaassen (2002), Ardia (2009), and Haas (2009), if regimes are not Normal but leptokurtic, the use of within-regime normality can seriously affect the identification of the regime process. Haas and Paolella (2012) specifically argue that MLE of MRS-type model with Gaussian assumption is inconsistent when fitted to fat-tailed data. A recent study by Shi and Feng (2016) provides simulation evidence to support this argument in the case of a GARCH specification. Hence, the sought of an alternative distribution to accommodate the excess kurtosis of the financial time series becomes an essential issue for the MRS model.

3.1 Student's t and General Error Distribution

Among the existing literature, Student's t-distribution and GED are two widely employed alternatives in finance studies using MRS-type models (Klaassen, 2002; Haas, Mittnik and Paolella, 2004; Marcucci, 2005). Both of them can capture leptokurtic and heavy-tail behaviours. When they are applied to the MRS model, the corresponding density functions of equation (2) are changed as follows.

$$\begin{aligned}
&\text{Student's t: } f(\varepsilon_{s_t,t}|s_t = j, \theta, \Omega_{t-1}) = \frac{\Gamma\left(\frac{v+1}{2}\right)}{\Gamma\left(\frac{v}{2}\right)\sqrt{\pi(v-2)\sigma_\varepsilon^2}}\left[1 + \frac{\varepsilon_{j,t}^2}{(v-2)\sigma_\varepsilon^2}\right]^{\frac{v+1}{2}} \\
&\text{GED: } f(\varepsilon_{s_t,t}|s_t = j, \theta, \Omega_{t-1}) = \frac{ve^{-\frac{1}{2}\left|\frac{\varepsilon_{j,t}}{\lambda\sigma_\varepsilon}\right|^v}}{\lambda 2^{(v+1)/v}\Gamma(1/v)} \\
&\text{where } \lambda = \left[\frac{2^{(-2/v)}\Gamma(1/v)}{\Gamma(3/v)}\right]^{\frac{1}{2}}
\end{aligned} \tag{6}$$

and v is the degrees of freedom. Then, the ML estimator $\hat{\theta}$ can be obtained in the same way as described in Section 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. Calzolari, Halbleib, and Parrini (2014) suggests that the Student's t-distribution lacks the stability-under-addition property. This problem also exists for non-stable distributions like GED. Since stability is desirable in portfolio applications and risk management, a distribution could overcome this issue is of particular importance.

3.2.1 α-stable distribution

As suggested by Calzolari, Halbleib, and Parrini (2014), the α-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 α-stable distribution overcomes the stability problem. Additionally, the α-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

$$kx + d \stackrel{d}{=} c_1 x_1 + c_2 x_2 \tag{7}$$

where x_1 and x_2 are independent variables and have the same distribution as x. The notation $\stackrel{d}{=}$ indicates equality in distribution. The property described by Equation (7) is also known as the *stability-under-addition* property (Tankov 2003). In particular, if $d = 0$, x is said to be strictly stable. According to Calzolari, Halbleib, and Parrini (2014), theoretical foundations of α-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 behaviour of the tails.

Since α-stable distribution does not have a close form of density function, the best way to describe it is by means of its characteristic function, which has the following form.

$$\phi(t) = \exp\{i\delta t - \sigma^\alpha |t|^\alpha (1 - i\beta sign(t)\Phi)\} \tag{8}$$

where Φ is $-(2/\pi)\log|t|$ when $\alpha = 1$ and is $\tan(\pi\alpha/2)$ when $\alpha \neq 1$. $\alpha \in [0, 2]$ is the index of stability or characteristic exponent that describes the tail-thickness of the distribution (small values indicating thick tails). $\sigma \in \mathrm{R}^+$ is the scale parameter. $\delta \in \mathrm{R}$ is the location parameter. β describe its skewness. Calzolari, Halbleib, and Parrini (2014) only consider the symmetric α-stable distribution ($\beta = 0$), which is then characterized by (α, σ, δ) and is denoted as $S(\alpha, \sigma, \delta)$. Therefore, the standardized symmetric version is $S(\alpha, 1, 0)$ with the following characteristic function

$$\phi(t) = \exp\{-|t|^\alpha\} \tag{9}$$

Despite its attractive properties like stability-under-addition, the second moment of the α-stable distribution does not exist in most cases.[4] Consequently, the application of this distribution to the MRS-type model will cause serious problems. For instance, the interpretation of σ_ε^2 would fail. More importantly, the existence of σ_ε^2 is required for the asymptotic properties of the MLE estimator to hold (Douc, Moulines, and Ryden 2004). Therefore, the sought of a substitute of the α-stable distribution, which has similar attractive properties and defined moments, would be of particular interest and importance.

3.2.2 Tempered stable distribution

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

$$\nu(x) = \frac{c^-}{|x|^{1+\alpha^-}} e^{-\lambda^-|x|} 1_{x<0} + \frac{c^+}{|x|^{1+\alpha^+}} e^{-\lambda^+|x|} 1_{x>0} \tag{10}$$

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

$$\begin{aligned} \phi(t) = \exp\Bigg\{&\Gamma(-\alpha^+)(\lambda^+)^{\alpha^+} C^+ \left[\left(1 - \frac{it}{\lambda^+}\right)^{\alpha^+} - 1 + \frac{it\alpha^+}{\lambda^+}\right] \\ &+ \Gamma(-\alpha^-)(\lambda^-)^{\alpha^-} C^- \left[\left(1 - \frac{it}{\lambda^-}\right)^{\alpha^-} - 1 + \frac{it\alpha^-}{\lambda^-}\right]\Bigg\} \end{aligned} \tag{11}$$

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

$$\begin{aligned}\kappa_1 &= 0\\ \kappa_2 &= \Gamma(2-\alpha^+)C^+(\lambda^+)^{\alpha^+-2} + \Gamma(2-\alpha^-)C^-(\lambda^-)^{\alpha^--2}\\ \kappa_3 &= \Gamma(3-\alpha^+)C^+(\lambda^+)^{\alpha^+-3} + \Gamma(3-\alpha^-)C^-(\lambda^-)^{\alpha^--3}\\ \kappa_4 &= \Gamma(4-\alpha^+)C^+(\lambda^+)^{\alpha^+-4} + \Gamma(4-\alpha^-)C^-(\lambda^-)^{\alpha^--4}\end{aligned} \tag{12}$$

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

$$\begin{aligned}m_1 &= \kappa_1\\ m_2 &= \kappa_2 + \kappa_1^2\\ m_3 &= \kappa_3 + 3\kappa_2\kappa_1 + \kappa_1^3\\ m_4 &= \kappa_4 + 4\kappa_3\kappa_1 + 3\kappa_2^2 + 6\kappa_2\kappa_1^2 + \kappa_1^4\end{aligned} \tag{13}$$

Clearly, the tempered stable distribution has defined moments, which enables it to be further employed to describe the innovation of the MRS-type model.

Küchler and Tappe (2013) suggest that sample paths of x can be classified as some well-known processes, which mainly depend on values of α^+ and α^-. 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 $x_t = \sum_{s\leq t} \Delta x_s$.
- 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 model, we require x to be standardized. Bianchi et al. (2010) argues that one way to achieve the standardization is letting

$$C^+ = \frac{p(\lambda^+)^{2-\alpha^+}}{\Gamma(2-\alpha^+)} \text{ and } C^- = \frac{(1-p)(\lambda^-)^{2-\alpha^-}}{\Gamma(2-\alpha^-)} \tag{14}$$

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

Combining equations (11) and (14), we now have a standardized tempered stable distribution. This distribution is expected to retain all the attractive properties similar to those of α-stable distribution and has defined moments.

3.2.3 MRS model with tempered stable distribution

Since the tempered stable distribution has defined moments, the sufficient condition to ensure the asymptotic properties of MLE of the MRS model will not be affected.[5] As the tempered stable distribution can effectively accommodate extreme behaviours like fat tails, we expect that the MLE of MRS model with this density would provide consistent estimators for financial data.

In terms of estimation, the procedures of MLE as discussed in Section 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, Doganoglu, and Chenyao (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 $f(\varepsilon_{s_t,t}|s_t, \theta, \Omega_{t-1})$ via the following iteration steps suggested by Shi and Feng (2016):

1. Acquiring the minimum and maximum of $\eta_t = \varepsilon_{s_t,t}/\sigma_\varepsilon$ as η_1 and η_2, respectively;
2. Calculating the values of $\phi(t)$ for the tempered stable distribution determined by the estimates of α^+, C^+, λ^+, α^-, C^- and λ^- via equation (11), 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 η_t that fall between the prespecified equally-spaced density values of t.

Hence, the interpolated values will be the estimates of $f(\varepsilon_{s_t,t}|s_t, \theta, \Omega_{t-1})$. By further applying them to equations (3) and (4), 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 the MRS models with Normal, Student's t, GED and tempered stable distributions. The data generation process is the 2-state MRS model with regime-switching in mean only as described equation (1) in all cases.[7] True distributions of the three studies are Student's t, GED and tempered stable, respectively. Additionally, following Pouzo, Psaradakis, and Sola (2016), we analyze the simulation results with different sample sizes (500, 1000 and 3000). A summary of the specific scenarios considered in this section can be found in Table 1.

Table 1: Simulation scenarios.

Scenarios	Student's t and GED			Tempered stable						
	p_{11}	p_{22}	T	p_{11}	p_{22}	T	α^+	α^-	λ^+	λ^-
1	0.999	0.99	500	0.999	0.99	3000	0.5	0.5	1.0	1.0
2			1000			3000	0.2	1.2	1.2	0.2
3			3000			3000	1.2	0.2	1.2	0.2
4	0.99	0.999	500	0.99	0.999	3000	0.5	0.5	1.0	1.0
5			1000			3000	0.2	1.2	1.2	0.2
6			3000			3000	1.2	0.2	1.2	0.2
7	0.99	0.99	500	0.99	0.99	3000	0.5	0.5	1.0	1.0
8			1000			3000	0.2	1.2	1.2	0.2
9			3000			3000	1.2	0.2	1.2	0.2
10	0.999	0.999	500	0.999	0.999	3000	0.5	0.5	1.0	1.0
11			1000			3000	0.2	1.2	1.2	0.2
12			3000			3000	1.2	0.2	1.2	0.2

4.1 Simulation study: student's t-distribution

First, we set the true distribution as Student's t with 3 degrees of freedom.[8] Altogether, twelve sets of simulations of the MRS process with different p_{11}, p_{22}[9] and sample sizes T are generated, where $\mu_1 = -0.5$, $\mu_2 = 0.5$, $\sigma_\varepsilon = 1$, $\gamma = 0.9$ and the number of replicates for each set is 1000. To avoid the starting bias, 10,000 points are generated for each simulation, and then only the last 500, 1000 or 3000 are kept. Moreover, to avoid simulation bias, 1500 such replicates are produced for each combination, while the first 500 are discarded.

The simulated data are fitted into the MRS models with Normal (MRS-N), Student's t (MRS-t), GED (MRS-G) and tempered stable (MRS-S) distributions, respectively. In Table 2, the mean log-likelihood (LL), bias, standard error (SE) and root-mean-square-error (RMSE) of μ_1, μ_2, p_{11}, p_{22} and σ_ε^2 obtained from the MRS-N, MRS-t and MRS-G models are reported. Bias is the mean difference between the true parameter and its estimates. SE is the standard error of the estimates. RMSE is the square root of the mean of squared difference between the true parameter and its estimates. The results of MRS-S model are presented in panel A of Table 5.

Although tempered stable distribution has four (five) more parameters than Student's t and GED (Normal), the log-likelihood is still a preliminary indicator of the model performance. Not surprisingly, MRS-N has the smallest log-likelihood in all cases, as the true distribution is fat-tailed. Besides, MRS-G leads to smaller log-likelihood than the true model. However, it is worth noticing that MRS-S can yield slightly greater log-likelihood compared with MRS-t. Therefore, it seems that MRS-S model might provide quite satisfied results, even when the true distribution is not tempered stable.

Table 2: Simulation results: student's t-distribution.

Scenarios	$Mean_{ll}$	$Bias_{\mu_1}$	SE_{μ_1}	$RMSE_{\mu_1}$	$Bias_{\mu_2}$	SE_{μ_2}	$RMSE_{\mu_2}$	$Bias_{p_{11}}$	$SE_{p_{11}}$	$RMSE_{p_{11}}$	$Bias_{p_{22}}$	$SE_{p_{22}}$	$RMSE_{p_{22}}$	$Bias_{\sigma}$	SE_{σ}	$RMSE_{\sigma}$	$Bias_{\gamma}$	SE_{γ}	$RMSE_{\gamma}$
Panel A: Normal distribution																			
1	−687.88	−0.1140	0.1671	0.2023	0.8043	2.0820	2.2320	−0.0067	0.0078	0.0103	−0.1895	0.3517	0.3995	−0.1043	0.3371	0.3529	0.0082	0.0172	0.0191
2	−1395.4	−0.0658	0.1315	0.1470	0.7902	1.8628	2.0235	−0.0033	0.0039	0.0051	−0.1789	0.3481	0.3914	−0.0733	0.1802	0.1945	0.0077	0.0130	0.0151
3	−4185.8	−0.0510	0.0953	0.1081	0.6487	1.4552	1.5932	−0.0025	0.0024	0.0035	−0.1670	0.3384	0.3774	−0.0775	0.1199	0.1428	0.0049	0.0093	0.0105
4	−700.52	−1.0157	3.0359	3.2013	0.1338	0.1768	0.2217	−0.1968	0.3496	0.4012	−0.0158	0.0902	0.0916	−0.0668	0.2789	0.2868	0.0082	0.0176	0.0194
5	−1399.4	−0.7996	2.4238	2.5523	0.0795	0.1453	0.1656	−0.1570	0.3282	0.3638	−0.0039	0.0043	0.0058	−0.0659	0.2478	0.2564	0.0065	0.0152	0.0165
6	−4199.8	−0.7350	1.8139	1.9572	0.0591	0.1007	0.1168	−0.1252	0.2859	0.3121	−0.0024	0.0026	0.0035	−0.0584	0.1335	0.1457	0.0056	0.0104	0.0118
7	−705.09	−0.1987	0.6332	0.6636	0.2463	1.0775	1.1053	−0.0445	0.1598	0.1659	−0.0544	0.1945	0.2020	−0.0852	0.2270	0.2425	0.0083	0.0149	0.0171
8	−1404.9	−0.0579	0.0962	0.1123	0.0529	0.0957	0.1093	−0.0128	0.0065	0.0144	−0.0134	0.0073	0.0153	−0.0709	0.1781	0.1917	0.0082	0.0092	0.0123
9	−4304.4	−0.0590	0.0557	0.0811	0.0488	0.0587	0.0763	−0.0119	0.0041	0.0126	−0.0117	0.0035	0.0122	−0.0326	0.1378	0.1416	0.0065	0.0057	0.0086
10	−690.58	−0.3009	1.0737	1.1151	0.1669	0.3911	0.4252	−0.0466	0.1904	0.1960	−0.0211	0.1045	0.1066	−0.0915	0.2328	0.2501	0.0078	0.0173	0.0190
11	−1382.5	−0.1595	0.4313	0.4598	0.1164	0.3235	0.3438	−0.0202	0.1010	0.1030	−0.0156	0.1001	0.1013	−0.0889	0.1710	0.1927	0.0071	0.0138	0.0155
12	−4223.5	−0.0753	0.0931	0.1197	0.0698	0.0905	0.1143	−0.0028	0.0068	0.0074	−0.0024	0.0052	0.0057	−0.0329	0.1496	0.1532	0.0048	0.0072	0.0087
Panel B: Student's t-distribution																			
1	−612.72	−0.0425	0.0763	0.0873	0.0168	0.1823	0.1831	−0.0064	0.0296	0.0303	−0.0495	0.1340	0.1429	0.0534	0.2008	0.2078	−0.0045	0.0075	0.0087
2	−1236.2	−0.0075	0.0509	0.0514	0.0090	0.0841	0.0846	−0.0009	0.0012	0.0015	−0.0210	0.0243	0.0321	0.0229	0.1050	0.1075	0.0032	0.0051	0.0060
3	−3681.1	−0.0072	0.0319	0.0327	0.0068	0.0586	0.0590	−0.0002	0.0007	0.0007	−0.0141	0.0141	0.0199	0.0026	0.0570	0.0571	0.0028	0.0028	0.0040
4	−620.3	−0.0034	0.1370	0.1370	0.0301	0.0901	0.0950	−0.0302	0.0539	0.0618	−0.0027	0.0035	0.0044	0.0477	0.1783	0.1846	0.0060	0.0081	0.0101
5	−1238	0.0055	0.1181	0.1182	0.0185	0.0644	0.0670	−0.0231	0.0330	0.0403	−0.0015	0.0030	0.0034	0.0245	0.1097	0.1124	0.0036	0.0065	0.0074
6	−3699.6	0.0008	0.0879	0.0879	0.0037	0.0280	0.0282	−0.0155	0.0150	0.0216	−0.0003	0.0006	0.0007	0.0093	0.0532	0.0540	−0.0032	0.0028	0.0043
7	−627.1	−0.0247	0.0862	0.0897	0.0158	0.0710	0.0727	−0.0161	0.0236	0.0286	−0.0172	0.0358	0.0397	0.0539	0.1885	0.1961	0.0060	0.0067	0.0090
8	−1253.6	−0.0056	0.0560	0.0563	−0.0019	0.0584	0.0584	−0.0103	0.0053	0.0116	−0.0106	0.0056	0.0120	0.0245	0.1063	0.1091	0.0032	0.0055	0.0064
9	−3780.2	−0.0021	0.0261	0.0262	0.0005	0.0304	0.0304	−0.0099	0.0030	0.0103	−0.0098	0.0029	0.0102	0.0071	0.0514	0.0519	0.0031	0.0025	0.0040
10	−613.56	−0.0333	0.0976	0.1031	0.0359	0.0926	0.0993	−0.0132	0.0915	0.0924	−0.0072	0.0303	0.0311	0.0499	0.1936	0.1999	−0.0036	0.0083	0.0090
11	−1229.4	−0.0221	0.0800	0.0830	0.0059	0.0584	0.0587	−0.0045	0.0211	0.0216	−0.0018	0.0049	0.0052	0.0346	0.1076	0.1130	0.0027	0.0063	0.0069
12	−3684.2	−0.0097	0.0356	0.0369	0.0092	0.0409	0.0419	−0.0005	0.0012	0.0013	−0.0004	0.0013	0.0014	0.0095	0.0617	0.0624	0.0020	0.0033	0.0039
Panel C: GED																			
1	−618.91	−0.0285	0.0944	0.0986	0.0177	0.2172	0.2179	−0.0087	0.0438	0.0447	−0.0378	0.1018	0.1086	−0.1740	0.1329	0.2189	0.0082	0.0086	0.0119
2	−1250.3	−0.0051	0.0634	0.0636	0.0108	0.1573	0.1577	−0.0009	0.0014	0.0017	−0.0209	0.0247	0.0324	−0.1695	0.0861	0.1901	0.0079	0.0054	0.0096
3	−3728.3	−0.0071	0.0349	0.0356	0.0049	0.1098	0.1099	−0.0002	0.0007	0.0007	−0.0147	0.0127	0.0194	−0.1567	0.0473	0.1637	0.0052	0.0034	0.0062
4	−627.36	0.0082	0.2083	0.2085	0.0269	0.0937	0.0975	−0.0429	0.0917	0.1012	−0.0028	0.0042	0.0050	−0.1623	0.1400	0.2143	0.0083	0.0091	0.0123
5	−1253	−0.0039	0.2166	0.2166	0.0135	0.0719	0.0732	−0.0375	0.0560	0.0674	−0.0011	0.0040	0.0041	−0.1511	0.0857	0.1737	0.0069	0.0072	0.0100
6	−3745.1	0.0024	0.1467	0.1467	0.0060	0.0330	0.0335	−0.0197	0.0210	0.0288	0.0008	0.0010	0.0013	−0.1423	0.0509	0.1511	0.0059	0.0033	0.0068
7	−633.43	−0.0082	0.0962	0.0965	0.0059	0.0878	0.0880	−0.0159	0.0243	0.0290	−0.0146	0.0344	0.0374	−0.1728	0.1108	0.2053	0.0094	0.0074	0.0120
8	−1266	−0.0017	0.0641	0.0641	−0.0017	0.0661	0.0661	−0.0101	0.0053	0.0114	−0.0105	0.0059	0.0120	−0.1579	0.0813	0.1776	0.0084	0.0062	0.0104
9	−3827.6	0.0008	0.0353	0.0353	−0.0009	0.0340	0.0340	−0.0097	0.0029	0.0101	−0.0096	0.0029	0.0100	−0.1566	0.0527	0.1652	0.0073	0.0031	0.0079
10	−620.14	−0.0289	0.1498	0.1526	0.0325	0.1486	0.1521	−0.0078	0.0322	0.0331	−0.0171	0.0969	0.0984	−0.1749	0.1097	0.2065	0.0080	0.0094	0.0123
11	−1242.6	−0.0229	0.1031	0.1056	0.0083	0.0663	0.0668	−0.0069	0.0308	0.0316	−0.0017	0.0048	0.0051	−0.1654	0.0940	0.1902	0.0070	0.0066	0.0096

12	−3730.7	−0.0074	0.0429	0.0435	0.0096	0.0394	0.0406	−0.0005	0.0012	0.0013	−0.0004	0.0013	0.0014	−0.1635	0.0482	0.1705	0.0049	0.0037	0.0061

4.1.1 Results of mean, standard deviation and autoregressive parameter

As discussed in Section 3, ML estimates of the MRS-N model may not be consistent for fat-tailed data. It is evidenced by the large absolute biases observed in all cases of μ_1 and μ_2. Some of them even exceed 1. As to the fat-tail models, both MRS-G and MRS-S generate much better estimates. Most of the absolute biases are less than 0.04. For σ_ε, however, the story is different. MRS-N and MRS-S yield relatively similar results of bias (although biases produced by MRS-S are smaller), which are not too far from those of the MRS-t. The estimates of MRS-G are much smaller than the true values with absolute bias around 0.17 in all cases. Turning to γ, biases produced by all models are basically negligible, the absolute values of which are all below 0.01. Hence, the biases produced by MRS-S are small and close to those of the true model in all cases. Biases generated by MRS-G (MRS-N) are substantial for standard deviation (mean) when the true model is MRS-t. Nevertheless, in most cases, the absolute biases reduce with the increasing sample sizes. This is consistently observed across different MRS models.

SE is a measurement of the estimation efficiency. It is well known that the 'wrong' model is generally less efficient than the true model. For MRS-N, its SEs are on a larger scale than the other models for μ_1 and μ_2. Some large SEs are even greater than 1. On the other hand, MRS-G and MRS-S still lead to similar results. Although their SEs of the μ_1 and μ_2 are slightly greater than those of the true model, they are basically on the same scale, and very few of them are over 0.2. For the estimation of standard deviation, MRS-G generates smaller SEs than the true model in basically all cases. The results of MRS-S are comparatively better than those of MRS-N and are not far from those of MRS-t. As for γ, MRS-N leads to slightly larger SEs, whereas those of MRS models with three fat-tailed distributions are close to each other. Therefore, MRS-S model can generate similar SEs to the true model regarding the estimates of mean, standard deviation and autoregressive parameter. The efficiency of MRS-N is low for the estimation of mean and is acceptable for the estimation of standard deviation and γ. MRS-G has similar efficiency to MRS-S for estimating mean and γ, and it can be more efficient than the true model to estimate standard deviation. However, since the bias of standard deviation produced by MRS-G is quite large, its overall performance of estimating standard deviation may not be optimal. Nonetheless, SEs generated by all models consistently reduce with increase in sample sizes.

Finally, RMSE is a combination of bias and SE, which is widely employed as the overall performance indicator in existing literature. Accordant with the above results, MRS-N is the least preferred model in the estimation of μ_1 and μ_2, while the performance of MRS-G and MRS-S are quite similar and close to that of the true model. As to the standard deviation, MRS-N and MRS-S lead to acceptable results, which are not too far from those of MRS-t. MRS-G, however, tends to generate the largest RMSEs in all cases. As for γ, RMSEs of MRS models with three fat-tailed distribution are close to each other, whereas those of MRS-N are slightly greater.

4.1.2 Results of transition probabilities

Biases of MRS-N are considerably substantial for both transition probabilities in all cases. Its absolute biases are much larger than those of the other models, with the maximum exceeding 0.19. Biases of MRS-G and MRS-S are minimal in all sets, and they are very close to those produced by MRS-t. Those absolute biases are almost all under 0.03. Also, it is worth noticing that MRS-S performs slightly better than MRS-G in most cases.

The efficiency of MRS-N is still the worst among all models. In scenarios that $p_{11} \neq p_{22}$, the SEs produced by MRS-N can be a couple of times greater than those of the other MRS models. The three fat-tailed MRS models have similar efficiency, and estimates of MRS-S are still overall more efficient than those of MRS-G. Finally, the RMSE suggests that MRS-N is the least preferred model to estimate the transition probabilities in all cases. The performance of the other three are close to each other, while MRS-S slightly outperforms MRS-G. With growing sample sizes, SEs drop rapidly in most scenarios, whereas the reductions in biases are less obvious.

In conclusion, with respect to estimating mean, standard deviation, autoregressive parameter and transition probabilities, MRS-S model can uniformly generate very similar estimates to those of the true model. MRS-N model cannot perform as well as the fat-tailed models in most cases. The estimates of mean, autoregressive parameter and transition probabilities generated by MRS-G are mostly acceptable, but it leads to the worst estimates of standard deviation. Therefore, considering the overall performance of all parameters, MRS-S is preferred to its competitors MRS-G and MRS-N. Additionally, its results are very close to those of the true model in all cases.

Table 3: Simulation results: GED.

Scenarios	$Mean_{ll}$	$Bias_{\mu_1}$	SE_{μ_1}	$RMSE_{\mu_1}$	$Bias_{\mu_2}$	SE_{μ_2}	$RMSE_{\mu_2}$	$Bias_{p_{11}}$	$SE_{p_{11}}$	$RMSE_{p_{11}}$	$Bias_{p_{22}}$	$SE_{p_{22}}$	$RMSE_{p_{22}}$	$Bias_{\sigma}$	SE_{σ}	$RMSE_{\sigma}$	$Bias_{\gamma}$	SE_{γ}	$RMSE_{\gamma}$
Panel A: Normal distribution																			
1	−716	−0.0631	0.1242	0.1393	0.2043	0.5129	0.5521	−0.0055	0.0095	0.0110	−0.1065	0.2605	0.2814	−0.0220	0.1541	0.1557	0.0071	0.0113	0.0133
2	−1426	−0.0460	0.1065	0.1160	0.1832	0.2802	0.3348	−0.0035	0.0056	0.0066	−0.1108	0.2465	0.2703	−0.0188	0.0750	0.0773	0.0049	0.0096	0.0108
3	−4267	−0.0205	0.0534	0.0572	0.1566	0.4905	0.5149	−0.0017	0.0046	0.0049	−0.0864	0.2413	0.2563	−0.0182	0.0510	0.0542	0.0025	0.0057	0.0062
4	−716	−0.2379	0.5722	0.6197	0.0520	0.1365	0.1461	−0.1120	0.2478	0.2719	−0.0058	0.0085	0.0103	−0.0285	0.1316	0.1347	0.0076	0.0139	0.0158
5	−1421	−0.1669	0.5326	0.5581	0.0627	0.0886	0.1085	−0.1023	0.2280	0.2499	−0.0028	0.0053	0.0060	−0.0218	0.0828	0.0856	0.0033	0.0089	0.0095
6	−4263	−0.1337	0.4086	0.4299	0.0163	0.0513	0.0538	−0.0957	0.1962	0.2183	−0.0020	0.0054	0.0058	−0.0185	0.0567	0.0596	0.0031	0.0052	0.0061
7	−715	−0.0615	0.1187	0.1337	0.0664	0.1163	0.1339	−0.0158	0.0147	0.0216	−0.0134	0.0123	0.0182	−0.0438	0.1293	0.1365	0.0067	0.0103	0.0123
8	−1440	−0.0281	0.0895	0.0938	0.0394	0.0900	0.0982	−0.0115	0.0075	0.0137	−0.0116	0.0074	0.0138	−0.0180	0.0602	0.0628	0.0052	0.0086	0.0100
9	−4351	−0.0255	0.0466	0.0531	0.0200	0.0447	0.0490	−0.0103	0.0038	0.0110	−0.0103	0.0033	0.0108	0.0003	0.0398	0.0398	0.0020	0.0044	0.0048
10	−712	−0.0771	0.1331	0.1538	0.0957	0.2170	0.2372	−0.0223	0.1070	0.1093	−0.0160	0.1134	0.1145	−0.0173	0.1347	0.1358	0.0071	0.0131	0.0149
11	−1421	−0.0767	0.2704	0.2811	0.0362	0.2141	0.2171	−0.0100	0.0443	0.0454	−0.0146	0.0998	0.1009	−0.0144	0.0763	0.0776	0.0044	0.0077	0.0089
12	−4275	−0.0160	0.0512	0.0536	0.0350	0.0949	0.1011	−0.0009	0.0022	0.0024	−0.0008	0.0041	0.0042	−0.0004	0.0416	0.0416	0.0012	0.0040	0.0042
Panel B: Student's t-distribution																			
1	−685	−0.0286	0.0948	0.0990	0.0682	0.2872	0.2952	−0.0048	0.0164	0.0171	−0.0599	0.1694	0.1797	0.2842	0.4030	0.4931	0.0076	0.0083	0.0113
2	−1365	−0.0215	0.0659	0.0693	0.0510	0.2518	0.2569	−0.0019	0.0057	0.0060	−0.0490	0.1229	0.1323	0.2196	0.1558	0.2693	0.0069	0.0052	0.0086
3	−4085	−0.0105	0.0377	0.0391	0.0276	0.1520	0.1545	−0.0003	0.0010	0.0010	−0.0245	0.0452	0.0514	0.1833	0.0763	0.1985	0.0054	0.0035	0.0064
4	−685	−0.0492	0.2990	0.3030	0.0246	0.0973	0.1004	−0.0646	0.1377	0.1521	−0.0063	0.0160	0.0172	0.2704	0.3204	0.4193	0.0087	0.0089	0.0124
5	−1360	−0.0264	0.2103	0.2120	0.0183	0.0725	0.0748	−0.0223	0.0330	0.0398	−0.0012	0.0019	0.0022	0.1892	0.1674	0.2526	0.0066	0.0073	0.0098
6	−4083	0.0016	0.0790	0.0790	0.0110	0.0347	0.0364	−0.0211	0.0329	0.0391	−0.0003	0.0009	0.0009	0.1797	0.0854	0.1990	0.0061	0.0032	0.0069
7	−686	−0.0383	0.0995	0.1066	0.0476	0.1043	0.1146	−0.0141	0.0118	0.0184	−0.0121	0.0096	0.0154	0.1939	0.2244	0.2966	0.0075	0.0077	0.0107
8	−1383	−0.0150	0.0686	0.0702	0.0156	0.0664	0.0682	−0.0109	0.0068	0.0128	−0.0106	0.0061	0.0122	0.1817	0.1256	0.2209	0.0070	0.0049	0.0085
9	−4172	−0.0134	0.0330	0.0356	0.0091	0.0352	0.0364	−0.0096	0.0035	0.0102	−0.0096	0.0028	0.0100	0.1670	0.0878	0.1887	0.0041	0.0029	0.0050
10	−681	−0.0396	0.1471	0.1523	0.0494	0.1539	0.1616	−0.0052	0.0129	0.0139	−0.0047	0.0118	0.0127	0.2168	0.2418	0.3248	0.0071	0.0099	0.0122
11	−1359	−0.0259	0.0960	0.0994	0.0246	0.0737	0.0777	−0.0034	0.0131	0.0135	−0.0011	0.0021	0.0024	0.1996	0.1606	0.2562	0.0058	0.0056	0.0081
12	−4088	−0.0075	0.0498	0.0504	0.0084	0.0431	0.0439	−0.0007	0.0017	0.0018	−0.0009	0.0018	0.0020	0.1915	0.0789	0.2071	0.0034	0.0032	0.0047
Panel C: GED																			
1	−680	−0.0168	0.0801	0.0818	0.0465	0.1417	0.1491	−0.0041	0.0108	0.0116	−0.0532	0.1530	0.1620	0.0161	0.1056	0.1068	0.0080	0.0076	0.0110
2	−1355	−0.0135	0.0566	0.0582	0.0224	0.1153	0.1175	−0.0013	0.0023	0.0026	−0.0469	0.1042	0.1143	0.0027	0.0698	0.0699	0.0075	0.0054	0.0092
3	−4056	−0.0067	0.0320	0.0327	0.0043	0.0613	0.0615	−0.0002	0.0008	0.0008	−0.0209	0.0395	0.0447	−0.0037	0.0340	0.0342	0.0066	0.0032	0.0073
4	−680	−0.0218	0.1345	0.1363	0.0160	0.0898	0.0912	−0.0505	0.1212	0.1313	−0.0038	0.0067	0.0077	0.0113	0.1098	0.1104	0.0080	0.0082	0.0115
5	−1351	−0.0192	0.1139	0.1155	0.0109	0.0653	0.0662	−0.0203	0.0310	0.0371	−0.0009	0.0013	0.0016	−0.0102	0.0756	0.0763	0.0068	0.0065	0.0094
6	−4054	0.0045	0.0713	0.0714	0.0087	0.0285	0.0298	−0.0189	0.0291	0.0347	−0.0003	0.0008	0.0009	−0.0060	0.0432	0.0436	0.0058	0.0028	0.0064
7	−681	−0.0256	0.0829	0.0868	0.0295	0.0868	0.0917	−0.0140	0.0122	0.0186	−0.0118	0.0091	0.0149	−0.0264	0.0912	0.0949	0.0082	0.0071	0.0108
8	−1373	−0.0054	0.0582	0.0584	0.0084	0.0577	0.0583	−0.0108	0.0068	0.0128	−0.0106	0.0060	0.0122	−0.0070	0.0595	0.0599	0.0080	0.0053	0.0096
9	−4145	−0.0042	0.0282	0.0285	0.0022	0.0306	0.0307	−0.0095	0.0035	0.0101	−0.0095	0.0028	0.0099	0.0066	0.0401	0.0406	0.0059	0.0025	0.0064
10	−676	−0.0252	0.1085	0.1114	0.0330	0.1065	0.1115	−0.0048	0.0125	0.0134	−0.0048	0.0125	0.0134	−0.0053	0.1068	0.1069	0.0077	0.0094	0.0122
11	−1349	−0.0256	0.0873	0.0910	0.0207	0.0634	0.0667	−0.0031	0.0104	0.0109	−0.0010	0.0020	0.0022	0.0025	0.0715	0.0715	0.0064	0.0054	0.0084

12	−4060	−0.0056	0.0440	0.0444	0.0070	0.0395	0.0401	−0.0007	0.0019	0.0020	−0.0009	0.0043	0.0044	−0.0006	0.0399	0.0399	0.0052	0.0030	0.0060

4.2 Simulation study: GED

Next, we set the true distribution as GED with 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 3 and panel B of Table 5. As preliminary evidence of model performance, log-likelihood results are accordant with those presented in Section 4.1. More specifically, MRS-N is the least preferred model, while MRS-t cannot perform as well as the true model in all cases. The results of MRS-S are overall slightly better than those of MRS-G.

4.2.1 Results of mean, standard deviation and autoregressive parameter

In the case of bias comparison, MRS-N still leads to much larger absolute values than the others for μ_1 and μ_2. Some of them are greater than 0.2. MRS-t and MRS-S have similar results, which are very close to those of the true model. Most of those absolute biases are less than 0.05. Turning to the standard deviation, biases of MRS-t are remarkably large at around 0.2. MRS-N and MRS-S generate similar biases as the true model (below 0.02 in most cases). As for γ, the biases produced in all models are fairly small and minimal (below 0.01). Hence, the biases produced by MRS-S are small and close to those of the true model in all cases. Biases generated by MRS-S (MRS-N) are substantial for standard deviation (mean), when the true model is MRS-G. Nevertheless, in most cases, the absolute biases reduce with the increasing sample sizes. This is consistently observed across different MRS models.

Regarding efficiency, MRS-N cannot perform as well as the others to estimate mean. Some of its SEs are above 0.5. The results of MRS-t and MRS-S are again similar and close to those of the true model. Most of them are below 0.2. As to the standard deviation, different from our findings in Section 4.1, SEs of MRS-t are greater than those produced by other MRS models in all cases. As for γ, MRS-N still leads to slightly larger SEs. Consequently, RMSE suggests that MRS-N (MRS-t) is the least preferred model in the estimation of the mean (standard deviation). MRS-N is the least preferred when estimating γ. The overall performance of MRS-S is very close to that of the true model in all cases. Nonetheless, as observed previously, all SEs are reducing with the increment of sample size across different MRS models.

4.2.2 Results of transition probabilities

Most of the absolute biases obtained from the MRS-N model are below 0.02, but are still overall the largest among all models. Taking the sets where $p_{11} \neq p_{22}$ as examples, the absolute biases of MRS-N can be 0.11. The results of the other models are basically less than 0.05. Hence, the results of MRS-N are still relatively worse than those of the fat-tailed models.

SEs of the MRS-N are small for the sets where $p_{11} = p_{22}$, but are still considerably large in the rest cases. Additionally, all of them are greater than those of the fat-tailed models. MRS-t and MRS-S lead to similar results which are close to those of MRS-G. It is worth noticing that the SEs of MRS-S are relatively smaller than those of MRS-t, especially in the cases where $p_{11} \neq p_{22}$. Therefore, as indicated by RMSE, MRS-S is preferred to the MRS-N and MRS-t in the estimation of transition probabilities. The results of MRS-S are also close to those of the true model in all cases. It is also worth noticing that both SEs and absolute biases drop with increasing sample sizes in almost all cases, whereas biases reduce at a much smaller scale.

In conclusion, with respect to estimating mean, standard deviation, autoregressive parameter and transition probabilities, MRS-S model can uniformly generate very similar estimates to those of MRS-G. MRS-N model cannot perform as well as the fat-tailed models in most cases. The estimates of mean, transition probabilities and autoregressive parameters generated by MRS-t are mostly acceptable, but it has the worst estimates of standard deviation. Therefore, considering the overall performance of all parameters, MRS-S outperforms its competitors MRS-t and MRS-N. Apart from that, the results of MRS-S are very close to those of the true model in all cases.

4.3 Simulation study: tempered stable distribution

Finally, we consider the scenario that the true distribution is tempered stable with three sets of different parameters (where all the ps are set to be 0.5), including one case of the CGMY distribution (when $\alpha^+ = \alpha^- = 0.5$ and $\lambda^+ = \lambda^- = 1.0$) and two general cases. Altogether, twelve sets of simulations are constructed. All the sample sizes are set to 3000. True values of p_{11}, p_{22}, μ_1, μ_2 and σ_ε^2 are not changed. 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 4 and panel C of Table 5. Not surprisingly, log-likelihood values of MRS-N, MRS-t and MRS-G are all smaller than those of the true model. Also, MRS-G outperforms MRS-t only in the CGMY distribution cases, while MRS-t generates lager log-likelihood in the other two general cases.

4.3.1 Results of mean, standard deviation and autoregressive parameter

In the case of bias comparison, MRS-N leads to the worst results for μ_1 and μ_2 with the largest absolute values over 0.8. MRS-t and MRS-G have similar absolute biases which are not far from those of the true model. For standard deviation, the absolute biases of MRS-N is smaller than those of the MRS-t and MRS-G. Most absolute biases of MRS-t (MRS-G) can be as large as 0.5 (0.1). Therefore, small biases can only be produced by MRS-N (MRS-t and MRS-G) in the case of standard deviation (mean). As for γ, all biases are small and are below 0.1.

The story of SE is slightly different. MRS-N is the least efficient model in the estimation of mean in most cases. Its largest SE is even above 1. The SEs of MRS-t and MRS-G are close to those of the true model. Most of them are at around 0.04. For the estimation of standard deviation, MRS-N and MRS-G are more efficient than MRS-t. Most of their SEs are below 0.1, while SEs of MRS-t are generally greater than 0.2. As for γ, SEs produced by all three fat-tailed models are close to each other and below 0.05. Those generated by MRS-N are slightly larger and can be greater than 0.1. In terms of overall performance, MRS-N is not preferred in the estimation of mean and autoregressive parameter, whereas it outperforms MRS-t and MRS-G to estimate standard deviation in most cases.

4.3.2 Results of transition probabilities

The absolute biases of MRS-N are below 0.11 when $p_{11} = p_{22}$, but can be above 0.2 when $p_{11} \neq p_{22}$. The results of MRS-t and MRS-G are similar and close to those of the true model (almost all of them are below 0.05). Hence, in terms of estimating transition probabilities, MRS-N still cannot provide small biases as the fat-tailed models.

The situation of SEs is basically the same as that of the biases. The SEs of MRS-N are below 0.01 when $p_{11} = p_{22}$, but can be as large as 0.4 when $p_{11} \neq p_{22}$. MRS-t and MRS-G lead to similar SEs which are not far from those of MRS-S. Their SEs are generally smaller than those of MRS-N. Therefore, RMSE suggests MRS-N is the least preferred model in the estimation of transition probabilities. MRS-t and MRS-G can provide similar results to those of the true model.

4.4 Summary of accuracy of estimated standard errors

Standard errors are crucial in measuring the significance of the parameters. To estimate those standard errors from the MRS model, we adopt a general approach and employ the empirical Hessian matrix.[11] To measure the accuracy of this approach, we follow Pouzo, Psaradakis, and Sola (2016) and consider the ratios of sampling standard error (SEs presented in Table 2, Table 3, Table 4 and Table 5) to those estimated standard errors from empirical Hessian matrix. The ratios are summarized in Table 6.

Table 4: Simulation results: tempered stable distribution.

Scenarios	$Mean_{ll}$	$Bias_{\mu_1}$	SE_{μ_1}	$RMSE_{\mu_1}$	$Bias_{\mu_2}$	SE_{μ_2}	$RMSE_{\mu_2}$	$Bias_{p_{11}}$	$SE_{p_{11}}$	$RMSE_{p_{11}}$	$Bias_{p_{22}}$	$SE_{p_{22}}$	$RMSE_{p_{22}}$	$Bias_{\sigma}$	SE_{σ}	$RMSE_{\sigma}$	$Bias_{\gamma}$	SE_{γ}	$RMSE_{\gamma}$
Panel A: Normal distribution																			
1	−608.19	−0.0261	0.0939	0.0975	−0.0279	0.1778	0.1800	−0.0060	0.0294	0.0300	−0.0364	0.0842	0.0917	0.0586	0.2104	0.2184	0.0083	0.0078	0.0114
2	−1231.7	−0.0077	0.0583	0.0588	−0.0219	0.1070	0.1092	−0.0009	0.0012	0.0015	−0.0197	0.0270	0.0334	0.0232	0.1130	0.1154	0.0062	0.0051	0.0080
3	−3676.5	−0.0026	0.0377	0.0378	0.0022	0.0734	0.0734	−0.0002	0.0007	0.0007	−0.0138	0.0158	0.0210	−0.0072	0.0495	0.0500	0.0045	0.0029	0.0054
4	−615.4	0.0408	0.1355	0.1415	0.0284	0.1019	0.1058	−0.0300	0.0491	0.0575	−0.0029	0.0030	0.0042	0.0560	0.2072	0.2146	0.0084	0.0079	0.0115
5	−1233.3	0.0408	0.1071	0.1146	0.0141	0.0714	0.0728	−0.0247	0.0461	0.0523	−0.0016	0.0015	0.0022	0.0231	0.1076	0.1101	0.0069	0.0064	0.0094
6	−3695.8	0.0015	0.0680	0.0680	0.0044	0.0357	0.0360	−0.0154	0.0148	0.0214	−0.0002	0.0006	0.0006	−0.0041	0.0636	0.0637	0.0052	0.0030	0.0060
7	−622.49	−0.0193	0.1040	0.1058	−0.0017	0.0712	0.0712	−0.0163	0.0243	0.0293	−0.0170	0.0279	0.0327	0.0644	0.1981	0.2083	0.0087	0.0073	0.0114
8	−1249.4	0.0032	0.0648	0.0649	−0.0015	0.0636	0.0636	−0.0103	0.0054	0.0116	−0.0107	0.0056	0.0121	0.0313	0.1102	0.1146	0.0074	0.0056	0.0093
9	−3775.6	−0.0022	0.0297	0.0298	−0.0013	0.0345	0.0345	−0.0099	0.0030	0.0103	−0.0098	0.0029	0.0102	−0.0089	0.0615	0.0621	0.0044	0.0025	0.0051
10	−609.15	−0.0142	0.1112	0.1121	0.0058	0.1058	0.1060	−0.0061	0.0213	0.0222	−0.0093	0.0556	0.0564	0.0541	0.2016	0.2087	0.0079	0.0090	0.0120
11	−1225	−0.0105	0.0843	0.0850	0.0024	0.0657	0.0657	−0.0049	0.0198	0.0204	−0.0018	0.0045	0.0048	0.0300	0.1177	0.1215	0.0075	0.0063	0.0098
12	−3679.7	−0.0091	0.0405	0.0415	0.0069	0.0443	0.0448	−0.0005	0.0012	0.0013	−0.0004	0.0013	0.0014	−0.0107	0.0528	0.0539	0.0044	0.0033	0.0055
Panel B: Student's t-distribution																			
1	−4058	−0.0077	0.0377	0.0385	−0.0291	0.0824	0.0874	−0.0003	0.0008	0.0009	−0.0255	0.0751	0.0793	0.1541	0.0892	0.1781	0.0079	0.0034	0.0086
2	−3785	−0.0459	0.0314	0.0556	−0.0216	0.1056	0.1078	−0.0058	0.0208	0.0216	−0.0698	0.2041	0.2157	0.5170	0.3028	0.5991	0.0080	0.0028	0.0085
3	−3768	0.0351	0.0305	0.0465	0.0393	0.0632	0.0744	−0.0002	0.0006	0.0006	−0.0386	0.1567	0.1614	0.4820	0.2530	0.5444	0.0086	0.0024	0.0089
4	−4064	−0.0126	0.1525	0.1530	0.0105	0.0384	0.0398	−0.0497	0.1611	0.1686	−0.0009	0.0039	0.0040	0.1616	0.0993	0.1897	0.0079	0.0035	0.0086
5	−3778	−0.0380	0.0721	0.0815	−0.0298	0.0292	0.0417	−0.0167	0.0279	0.0325	−0.0004	0.0006	0.0007	0.5040	0.2588	0.5666	0.0079	0.0025	0.0083
6	−3776	0.0283	0.0964	0.1005	0.0378	0.0318	0.0494	−0.0469	0.1554	0.1623	−0.0006	0.0027	0.0028	0.4634	0.2702	0.5364	0.0087	0.0028	0.0091
7	−4137	−0.0107	0.0326	0.0343	0.0079	0.0274	0.0285	−0.0091	0.0029	0.0096	−0.0093	0.0029	0.0097	0.1536	0.0776	0.1721	0.0080	0.0027	0.0084
8	−3857	−0.0410	0.0312	0.0515	−0.0328	0.0246	0.0410	−0.0093	0.0029	0.0097	−0.0094	0.0030	0.0099	0.4723	0.2316	0.5260	0.0086	0.0024	0.0089
9	−3856	0.0322	0.0358	0.0482	0.0409	0.0264	0.0487	−0.0095	0.0036	0.0102	−0.0101	0.0030	0.0105	0.4732	0.2691	0.5444	0.0080	0.0027	0.0084
10	−4053	−0.0102	0.0417	0.0429	0.0098	0.0477	0.0487	−0.0003	0.0012	0.0012	−0.0005	0.0017	0.0018	0.1620	0.0803	0.1808	0.0079	0.0034	0.0086
11	−3776	−0.0455	0.0370	0.0586	−0.0238	0.0377	0.0446	−0.0007	0.0015	0.0017	−0.0009	0.0039	0.0040	0.4979	0.3030	0.5828	0.0078	0.0029	0.0083
12	−3760	0.0302	0.0358	0.0468	0.0397	0.0352	0.0531	−0.0006	0.0017	0.0018	−0.0005	0.0014	0.0015	0.4396	0.2149	0.4893	0.0080	0.0029	0.0085
Panel C: GED																			
1	−4050	−0.0045	0.0378	0.0381	−0.0082	0.0991	0.0994	−0.0002	0.0008	0.0008	−0.0285	0.1052	0.1090	−0.0227	0.0467	0.0519	0.0082	0.0035	0.0089
2	−3799	−0.0695	0.0355	0.0780	−0.0511	0.1120	0.1231	−0.0022	0.0146	0.0148	−0.0314	0.0870	0.0925	−0.0902	0.0997	0.1344	0.0078	0.0033	0.0085
3	−3780	0.0641	0.0325	0.0719	0.0694	0.1115	0.1313	−0.0003	0.0006	0.0007	−0.0202	0.0729	0.0756	−0.0999	0.0861	0.1319	0.0088	0.0030	0.0093
4	−4054	0.0247	0.1513	0.1533	0.0088	0.0386	0.0396	−0.0277	0.0847	0.0891	−0.0003	0.0007	0.0008	−0.0183	0.0519	0.0550	0.0080	0.0036	0.0088
5	−3790	−0.0603	0.1371	0.1498	−0.0593	0.0327	0.0677	−0.0324	0.1316	0.1355	−0.0004	0.0006	0.0007	−0.0912	0.1036	0.1380	0.0084	0.0027	0.0088
6	−3790	0.0628	0.1524	0.1648	0.0592	0.0338	0.0682	−0.0280	0.0781	0.0830	−0.0007	0.0023	0.0024	−0.0980	0.0926	0.1348	0.0090	0.0033	0.0096
7	−4128	−0.0052	0.0323	0.0327	0.0044	0.0281	0.0284	−0.0090	0.0029	0.0095	−0.0091	0.0029	0.0096	−0.0175	0.0676	0.0698	0.0083	0.0029	0.0088
8	−3869	−0.0605	0.0280	0.0667	−0.0678	0.0312	0.0746	−0.0091	0.0028	0.0095	−0.0091	0.0029	0.0096	−0.0942	0.0885	0.1293	0.0093	0.0029	0.0097
9	−3868	0.0647	0.0356	0.0738	0.0610	0.0309	0.0684	−0.0092	0.0036	0.0099	−0.0100	0.0030	0.0104	−0.0963	0.0940	0.1346	0.0086	0.0031	0.0091
10	−4045	−0.0104	0.0397	0.0410	0.0116	0.0508	0.0521	−0.0004	0.0012	0.0013	−0.0005	0.0018	0.0019	−0.0179	0.0605	0.0631	0.0080	0.0035	0.0087
11	−3788	−0.0723	0.0393	0.0823	−0.0477	0.0411	0.0630	−0.0007	0.0015	0.0017	−0.0008	0.0033	0.0034	−0.0935	0.0665	0.1147	0.0077	0.0034	0.0084

12	−3773	0.0590	0.0369	0.0696	0.0610	0.0354	0.0705	−0.0006	0.0016	0.0017	−0.0004	0.0014	0.0015	−0.1039	0.0672	0.1237	0.0084	0.0032	0.0090

Table 5: Summary of the MRS-S models.

Scenarios	$Mean_{ll}$	$Bias_{\mu_1}$	SE_{μ_1}	$RMSE_{\mu_1}$	$Bias_{\mu_2}$	SE_{μ_2}	$RMSE_{\mu_2}$	$Bias_{p_{11}}$	$SE_{p_{11}}$	$RMSE_{p_{11}}$	$Bias_{p_{22}}$	$SE_{p_{22}}$	$RMSE_{p_{22}}$	$Bias_{\sigma}$	SE_{σ}	$RMSE_{\sigma}$	$Bias_{\gamma}$	SE_{γ}	$RMSE_{\gamma}$
Panel A: Student's t-distribution																			
1	−608	−0.0261	0.0939	0.0975	−0.0279	0.1778	0.1800	−0.0060	0.0294	0.0300	−0.0364	0.0842	0.0917	0.0586	0.2104	0.2184	0.0083	0.0078	0.0114
2	−1232	−0.0026	0.0583	0.0584	−0.0219	0.1070	0.1092	−0.0009	0.0012	0.0015	−0.0197	0.0270	0.0334	0.0232	0.1130	0.1154	0.0062	0.0051	0.0080
3	−3677	−0.0077	0.0377	0.0385	0.0022	0.0734	0.0734	−0.0002	0.0007	0.0007	−0.0138	0.0158	0.0210	−0.0072	0.0495	0.0500	0.0045	0.0029	0.0054
4	−615	0.0372	0.1355	0.1405	0.0284	0.1019	0.1058	−0.0300	0.0491	0.0575	−0.0029	0.0030	0.0042	0.0560	0.2072	0.2146	0.0084	0.0079	0.0115
5	−1233	0.0408	0.1071	0.1146	0.0141	0.0714	0.0728	−0.0247	0.0461	0.0523	−0.0016	0.0015	0.0022	0.0231	0.1076	0.1101	0.0069	0.0064	0.0094
6	−3696	0.0015	0.0680	0.0680	0.0044	0.0357	0.0360	−0.0154	0.0148	0.0214	−0.0002	0.0006	0.0006	−0.0041	0.0636	0.0637	0.0052	0.0030	0.0060
7	−622	−0.0193	0.1040	0.1058	−0.0013	0.0712	0.0712	−0.0163	0.0243	0.0293	−0.0170	0.0279	0.0327	0.0644	0.1981	0.2083	0.0087	0.0073	0.0114
8	−1249	0.0032	0.0648	0.0649	−0.0017	0.0636	0.0636	−0.0103	0.0054	0.0116	−0.0107	0.0056	0.0121	0.0313	0.1102	0.1146	0.0074	0.0056	0.0093
9	−3776	−0.0022	0.0297	0.0298	−0.0015	0.0345	0.0345	−0.0099	0.0030	0.0103	−0.0098	0.0029	0.0102	−0.0089	0.0615	0.0621	0.0044	0.0025	0.0051
10	−609	−0.0142	0.1112	0.1121	0.0058	0.1058	0.1060	−0.0061	0.0213	0.0222	−0.0093	0.0556	0.0564	0.0541	0.2016	0.2087	0.0079	0.0090	0.0120
11	−1225	−0.0105	0.0843	0.0850	0.0024	0.0657	0.0657	−0.0049	0.0198	0.0204	−0.0018	0.0045	0.0048	0.0300	0.1177	0.1215	0.0075	0.0063	0.0098
12	−3680	−0.0091	0.0405	0.0415	0.0069	0.0443	0.0448	−0.0005	0.0012	0.0013	−0.0004	0.0013	0.0014	−0.0107	0.0528	0.0539	0.0044	0.0033	0.0055
Panel B: GED																			
1	−677	−0.0151	0.0895	0.0908	−0.0465	0.1528	0.1597	−0.0067	0.0258	0.0267	−0.0530	0.1423	0.1518	0.0241	0.1087	0.1113	0.0081	0.0087	0.0119
2	−1352	−0.0143	0.0579	0.0596	−0.0305	0.1185	0.1224	−0.0023	0.0078	0.0081	−0.0305	0.0508	0.0593	0.0163	0.0761	0.0778	0.0070	0.0058	0.0091
3	−4054	−0.0077	0.0359	0.0367	−0.0043	0.0896	0.0897	−0.0002	0.0008	0.0008	−0.0193	0.0221	0.0293	−0.0015	0.0490	0.0490	0.0061	0.0036	0.0071
4	−677	0.0445	0.1689	0.1747	0.0025	0.0926	0.0926	−0.0410	0.0808	0.0906	−0.0051	0.0151	0.0159	0.0230	0.1210	0.1232	0.0080	0.0087	0.0118
5	−1348	0.0243	0.1291	0.1314	0.0124	0.0719	0.0730	−0.0209	0.0363	0.0419	−0.0009	0.0014	0.0017	−0.0062	0.0767	0.0770	0.0078	0.0067	0.0103
6	−4052	0.0077	0.0790	0.0794	0.0075	0.0324	0.0333	−0.0180	0.0270	0.0324	−0.0003	0.0008	0.0009	0.0060	0.0436	0.0440	0.0062	0.0030	0.0069
7	−678	−0.0162	0.0838	0.0854	0.0212	0.0889	0.0914	−0.0136	0.0110	0.0175	−0.0118	0.0096	0.0152	0.0080	0.0947	0.0950	0.0083	0.0079	0.0115
8	−1370	−0.0043	0.0610	0.0612	0.0082	0.0611	0.0616	−0.0108	0.0067	0.0127	−0.0106	0.0060	0.0122	0.0002	0.0574	0.0574	0.0080	0.0058	0.0099
9	−4142	−0.0011	0.0326	0.0326	0.0008	0.0307	0.0307	−0.0094	0.0035	0.0100	−0.0095	0.0028	0.0099	0.0001	0.0391	0.0391	0.0069	0.0028	0.0074
10	−673	−0.0143	0.1043	0.1053	0.0152	0.1130	0.1140	−0.0053	0.0131	0.0141	−0.0078	0.0137	0.0158	0.0182	0.1064	0.1079	0.0079	0.0092	0.0121
11	−1346	−0.0223	0.0811	0.0841	0.0098	0.0674	0.0681	−0.0032	0.0106	0.0111	−0.0010	0.0021	0.0023	0.0033	0.0763	0.0764	0.0066	0.0056	0.0087
12	−4057	−0.0012	0.0451	0.0451	0.0027	0.0375	0.0376	−0.0008	0.0021	0.0022	−0.0009	0.0043	0.0044	0.0006	0.0403	0.0403	0.0050	0.0033	0.0060
Panel C: Tempered stable distribution																			
1	−4042	−0.0027	0.0361	0.0362	−0.0031	0.0802	0.0803	−0.0002	0.0008	0.0008	−0.0151	0.0161	0.0221	−0.0020	0.0467	0.0467	0.0072	0.0033	0.0079
2	−3747	−0.0082	0.0299	0.0310	−0.0081	0.0753	0.0757	−0.0002	0.0006	0.0006	−0.0174	0.0234	0.0292	0.0339	0.0517	0.0618	0.0064	0.0027	0.0069
3	−3727	0.0080	0.0280	0.0291	−0.0082	0.0622	0.0627	−0.0002	0.0007	0.0007	−0.0148	0.0233	0.0276	0.0524	0.0462	0.0699	0.0041	0.0024	0.0048
4	−4047	−0.0032	0.0771	0.0772	0.0065	0.0378	0.0384	−0.0222	0.0489	0.0537	−0.0003	0.0008	0.0009	−0.0005	0.0463	0.0463	0.0065	0.0036	0.0074
5	−3737	0.0138	0.0692	0.0706	−0.0016	0.0334	0.0334	−0.0282	0.0880	0.0924	−0.0003	0.0004	0.0005	0.0535	0.0482	0.0720	0.0032	0.0024	0.0040
6	−3738	0.0273	0.0744	0.0793	0.0012	0.0311	0.0311	−0.0212	0.0563	0.0602	−0.0002	0.0008	0.0008	0.0476	0.0507	0.0695	0.0070	0.0026	0.0075
7	−4121	−0.0051	0.0294	0.0298	0.0034	0.0293	0.0295	−0.0090	0.0029	0.0095	−0.0092	0.0029	0.0096	0.0004	0.0386	0.0386	0.0065	0.0026	0.0070
8	−3817	−0.0027	0.0264	0.0265	−0.0015	0.0276	0.0276	−0.0090	0.0028	0.0094	−0.0093	0.0028	0.0097	0.0481	0.0449	0.0658	0.0041	0.0022	0.0047
9	−3818	0.0033	0.0321	0.0323	0.0048	0.0295	0.0299	−0.0089	0.0036	0.0096	−0.0098	0.0029	0.0102	0.0445	0.0500	0.0669	0.0038	0.0025	0.0045
10	−4037	−0.0083	0.0382	0.0391	0.0055	0.0471	0.0474	−0.0004	0.0012	0.0013	−0.0005	0.0017	0.0018	0.0022	0.0429	0.0430	0.0068	0.0035	0.0076
11	−3736	−0.0098	0.0331	0.0345	0.0011	0.0361	0.0361	−0.0007	0.0015	0.0017	−0.0007	0.0026	0.0027	0.0383	0.0456	0.0596	0.0031	0.0027	0.0041

12	−3721	0.0047	0.0326	0.0329	0.0047	0.0328	0.0331	−0.0005	0.0015	0.0016	−0.0005	0.0015	0.0016	0.0432	0.0426	0.0607	0.0048	0.0029	0.0056

Table 6: Simulation results: sampling standard error to estimated standard error.

Scenarios	Student's t						GED						Tempered Stable					
	μ_1	μ_2	p_{11}	p_{22}	σ	γ	μ_1	μ_2	p_{11}	p_{22}	σ	γ	μ_1	μ_2	p_{11}	p_{22}	σ	γ
Panel A: Normal distribution																		
1	1.7082	1.7519	1.8460	1.6705	1.3953	1.3716	1.8239	1.5332	1.6452	1.6138	1.3211	1.3730	1.2796	1.0647	1.0885	1.2636	1.0722	1.0782
2	1.3444	1.4201	1.4741	1.4627	1.2976	1.1183	1.3489	1.4081	1.3486	1.4479	1.1859	1.1353	1.0624	1.2541	1.0515	1.0884	1.0143	0.9833
3	1.2541	1.1896	1.0527	1.0551	1.1332	1.0919	1.1230	1.1830	1.2390	1.1470	1.1051	1.0171	1.2410	1.2112	1.1335	1.2711	1.0741	1.0130
4	1.8410	1.6567	1.8497	1.6280	1.4974	1.4445	1.6571	1.7770	1.8624	1.6595	1.3066	1.4727	1.0793	1.2377	1.2578	1.0832	1.1170	1.0254
5	1.4357	1.4569	1.3766	1.3613	1.1282	1.1428	1.3757	1.3048	1.3515	1.4662	1.1675	1.2270	1.2327	1.1256	1.2147	1.1254	1.0462	1.0942
6	1.2924	1.1159	1.0803	1.1511	1.0235	1.0341	1.2764	1.1149	1.1192	1.1223	1.0549	0.9986	1.2975	1.1821	1.2600	1.0754	1.0989	0.9805
7	1.8371	1.9948	1.8492	1.5478	1.3691	1.3191	1.7802	1.6952	1.6766	1.8724	1.4781	1.3535	1.1153	1.0612	1.2068	1.2967	1.0293	1.1147
8	1.4235	1.3300	1.3613	1.4588	1.2647	1.2674	1.4419	1.4744	1.3922	1.4509	1.1924	1.1133	1.0747	1.1819	1.0530	1.2853	1.0908	0.9996
9	1.0770	1.2564	1.1189	1.1692	1.0569	1.1445	1.1704	1.1693	1.1218	1.1366	0.9953	1.1287	1.1160	1.1374	1.0514	1.2707	1.1346	1.0038
10	1.8439	1.8160	1.5598	1.9364	1.3497	1.3760	1.5818	1.7633	1.5442	1.6746	1.4547	1.3538	1.2989	1.2804	1.0782	1.2082	1.0963	1.0175
11	1.3362	1.3069	1.4173	1.3378	1.2494	1.1407	1.3339	1.3420	1.3149	1.3461	1.2500	1.0682	1.1509	1.1799	1.1670	1.2729	1.1273	1.1469
12	1.0595	1.1013	1.2441	1.1850	1.0753	0.9864	1.1725	1.1006	1.0734	1.2534	1.0883	1.0484	1.1995	1.0939	1.0822	1.1484	1.1168	1.1435
Panel B: Student's t-distribution																		
1	1.4269	1.3260	1.3610	1.4662	1.4788	1.3036	1.4911	1.4791	1.4399	1.4525	1.8987	1.4463	1.0250	1.0005	1.1368	1.0725	1.1927	0.9985
2	1.2551	1.1035	1.1841	1.2019	1.0941	1.0529	1.1562	1.1985	1.1023	1.1719	1.3304	1.2860	1.0076	1.1092	0.9945	1.0770	1.1689	1.0950
3	1.0695	1.0753	1.0254	1.0032	1.0664	1.1185	1.0818	1.0887	1.0454	1.0285	1.1580	1.1297	1.1187	1.0883	1.1196	1.0121	1.1619	1.0345
4	1.3207	1.3699	1.4101	1.4126	1.4094	1.4529	1.4625	1.4778	1.3673	1.4957	1.6527	1.3442	1.0976	1.0246	1.1258	1.0052	1.2393	1.0049
5	1.1520	1.0512	1.2227	1.1359	1.1424	1.1356	1.1361	1.1610	1.2410	1.1389	1.3590	1.2635	1.0274	1.0191	1.1197	1.1272	1.2710	1.0426
6	1.0478	1.0562	1.1027	1.1160	1.0737	1.0092	1.1403	1.1135	1.1008	1.0357	1.1993	1.0111	1.0611	1.1482	1.0829	1.1251	1.2076	1.0258
7	1.3085	1.3103	1.4134	1.4999	1.4847	1.3865	1.4079	1.4775	1.4888	1.3647	1.9186	1.4150	1.1475	0.9897	1.0968	1.1383	1.2422	1.0184
8	1.1895	1.1536	1.0713	1.2765	1.0731	1.2386	1.2677	1.2462	1.1002	1.2885	1.4470	1.2375	1.0176	1.0865	1.1299	1.0757	1.2487	1.0413
9	1.1351	1.0105	1.0672	1.1218	1.0941	1.0083	1.0737	1.0939	1.0710	1.0349	1.2346	1.0427	1.0458	1.0420	1.1154	1.1369	1.2577	1.0531
10	1.3223	1.4318	1.3540	1.3663	1.3863	1.4894	1.4833	1.3617	1.4508	1.4236	1.9744	1.3545	1.1427	1.1075	0.9948	1.0916	1.2534	1.1467
11	1.0537	1.1854	1.1263	1.2866	1.1797	1.2720	1.2570	1.2563	1.2025	1.1868	1.3794	1.2379	1.1074	1.0690	1.0680	1.1195	1.2301	1.0161
12	1.1465	1.1192	1.1117	1.0322	1.0025	1.1339	0.9841	1.0995	1.1304	1.0691	1.2057	1.1124	1.1261	0.9880	1.0301	1.0157	1.1830	1.0492
Panel C: GED																		
1	1.4440	1.4565	1.4764	1.4566	1.8727	1.4148	1.3180	1.3466	1.3513	1.4693	1.3458	1.3665	1.1360	1.0856	1.1056	1.1085	1.2911	1.0050
2	1.2373	1.2668	1.2870	1.2035	1.4818	1.1528	1.1017	1.2476	1.2916	1.2555	1.2917	1.1164	1.0292	1.0825	1.0348	1.0153	1.1730	1.1125
3	1.1165	0.9862	1.0064	1.0862	1.2230	1.0785	1.0552	1.0991	1.1329	0.9981	1.0934	1.0835	0.9940	1.1160	1.0835	0.9903	1.1719	1.1327
4	1.4664	1.4908	1.3028	1.4651	1.6670	1.3994	1.3300	1.4290	1.4610	1.3256	1.4387	1.3251	0.9888	0.9852	1.0592	1.0809	1.1586	0.9943
5	1.1673	1.1602	1.1545	1.2984	1.3471	1.1333	1.0845	1.0746	1.2506	1.1898	1.1670	1.1313	1.1124	1.0742	1.0620	1.1013	1.1734	1.0726
6	1.0873	1.1211	1.0483	1.1163	1.2918	1.0632	1.0775	1.1237	1.0228	1.0843	1.1109	1.1324	1.0932	1.0753	1.0075	1.1285	1.2072	1.0715
7	1.3662	1.3214	1.4557	1.3067	1.9483	1.4236	1.3734	1.4890	1.3935	1.4913	1.4501	1.3584	1.1014	1.0940	1.0969	0.9939	1.1904	1.0037
8	1.0513	1.2395	1.2529	1.2431	1.3274	1.2664	1.2629	1.0997	1.0535	1.1596	1.1143	1.1400	1.0672	1.1042	0.9841	1.0009	1.2423	1.0961
9	1.1320	0.9842	1.0305	1.0743	1.2597	1.0219	0.9985	1.0633	1.0983	0.9863	1.1257	1.1416	1.0597	0.9922	1.1350	0.9893	1.2527	1.0480
10	1.3409	1.3314	1.4500	1.4803	1.7879	1.3991	1.3419	1.4672	1.4026	1.3845	1.3774	1.3527	0.9996	1.0466	1.1079	1.1085	1.2100	0.9970

11	1.1550	1.2370	1.1022	1.0763	1.3266	1.2201	1.1784	1.2245	1.0768	1.2532	1.1500	1.1428	1.0936	0.9902	1.1036	1.0310	1.2992	0.9944
12	1.0857	1.0700	0.9889	1.0338	1.2172	1.1157	1.0367	1.1263	1.0231	1.1281	1.1125	1.0831	0.9895	1.1388	1.0878	1.0552	1.1967	1.1297
Panel D: Tempered stable distribution																		
1	1.4394	1.3401	1.4325	1.3729	1.3376	1.4966	1.4627	1.3508	1.3134	1.3096	1.3528	1.3487	1.1426	1.1379	1.0207	1.0191	1.0903	1.1335
2	1.1542	1.0725	1.0853	1.2110	1.2986	1.1158	1.2017	1.0763	1.1846	1.2257	1.1567	1.2148	1.0101	1.0970	1.1467	0.9986	0.9846	1.0073
3	1.0928	1.0049	1.1300	1.0343	1.0478	1.0697	1.0906	1.1116	1.0541	0.9874	1.0203	1.1016	1.0960	1.0016	1.0608	1.0117	1.0917	1.0653
4	1.4633	1.3012	1.4647	1.4411	1.3261	1.4770	1.3413	1.4306	1.3607	1.3691	1.3785	1.3028	1.0417	1.0703	1.1114	1.1202	1.1217	1.0084
5	1.2055	1.2402	1.0568	1.0713	1.0646	1.1041	1.2776	1.1449	1.1561	1.1040	1.1637	1.0542	1.0800	1.0717	1.0650	1.1337	1.1175	1.1320
6	1.1355	0.9975	1.0105	1.0939	1.0504	1.1356	1.0588	1.0120	1.0134	1.1329	0.9827	1.1262	1.0431	1.0994	0.9863	1.0475	1.1134	1.1258
7	1.4723	1.4384	1.4906	1.3257	1.4291	1.4949	1.4475	1.3717	1.4147	1.4891	1.4547	1.3266	1.1112	0.9852	1.0656	1.0142	1.0152	1.0158
8	1.2182	1.1757	1.1789	1.1959	1.1393	1.0783	1.1504	1.2147	1.0955	1.0654	1.1300	1.2631	1.1264	1.0500	0.9995	1.0408	1.0602	1.0079
9	1.1491	1.0077	1.1287	1.0174	1.0731	1.1300	1.0303	1.0001	1.0101	1.1132	1.0006	1.0077	1.0193	1.0942	1.0328	0.9853	1.1436	1.1170
10	1.4572	1.3813	1.4356	1.3831	1.4635	1.4710	1.3388	1.3302	1.4415	1.4711	1.3108	1.4245	1.0302	1.1393	1.1433	1.0800	1.0186	1.1188
11	1.0841	1.0935	1.2664	1.2305	1.1899	1.0541	1.0996	1.0764	1.1126	1.1317	1.1097	1.0504	1.0193	0.9923	1.0227	0.9818	1.0952	1.1316
12	1.0608	1.0139	0.9912	1.0559	1.0160	1.1258	1.0160	1.0616	1.1196	1.0376	1.0307	1.1096	1.0586	0.9886	1.0199	1.1495	1.0852	0.9815

Our previous results suggest that MRS-N (MRS-t and MRS-G) can produce considerable biases when estimating mean and transition probabilities (standard deviation). Therefore, it is expected that the accuracy of the corresponding estimated standard errors may be problematic. From Table 6, results of MRS-N suggest that the ratios are substantially large for μ_1, μ_2, p_{11} and p_{22}, especially when the sample size is small. Some ratios are over 1.5 under such circumstances. When sample size increases, the ratios drop quickly and are closer to 1. As for σ and γ, the ratios are comparatively smaller, suggesting the estimated standard errors are more accurate. For MRS-t and MRS-G models, when they are misspecified, most ratios are smaller than or close to those of MRS-N for all parameters except for σ. This is again consistent with our expectation, as those misspecified models cannot produce small bias for standard deviation. Nevertheless, ratios are closer to 1 when the sample size increases.

Finally, MRS-S model leads to satisfactory results for all parameters. Its ratios are not far from those of the true model when misspecified, indicating that the standard error estimated from empirical Hessian matrix is quite reliable. Also, with a relatively large sample size, those ratios are fairly close to 1.

4.5 Summary of states identification

As the underlying Markov states are latent, smoothing probabilities explained in Section 2 are important in identifying them. To produce the true values of those probabilities, we first use the true parameters and innovation distributions to generate the true likelihoods for each observation. Then, following the filtering process discussed in Hamilton (1989), the true filtered probabilities ($\omega_{1,t}$) can be produced. The true smoothing probabilities can then be derived using Equation (5). After all four MRS models are fitted for the replicate, corresponding smoothing probabilities are estimated and the absolute differences from the true probabilities for each t is recorded. As a measure of the correctiveness in state identification, we then compute the average of the absolute differences for each replicate (summation over the T absolute differences divided by T). The mean and SE of those averages across different MRS models for each simulation scenario are summarized in Table 7.

Table 7: Simulation results: smoothing probabilities.

Scenarios	$Mean_N$	SE_N	$Mean_t$	SE_t	$Mean_G$	SE_G	$Mean_{TS}$	SE_{TS}
Panel A: Student's t-distribution								
1	0.0308	0.0242	0.0213	0.0495	0.0225	0.0524	0.0218	0.0469
2	0.0210	0.0155	0.0075	0.0054	0.0083	0.0070	0.0076	0.0056
3	0.0169	0.0086	0.0049	0.0032	0.0051	0.0033	0.0050	0.0032
4	0.0352	0.0716	0.0142	0.0139	0.0159	0.0489	0.0149	0.0155
5	0.0215	0.0139	0.0098	0.0105	0.0114	0.0154	0.0104	0.0140
6	0.0178	0.0117	0.0058	0.0047	0.0089	0.0073	0.0056	0.0032
7	0.0484	0.0290	0.0281	0.0159	0.0294	0.0172	0.0289	0.0153
8	0.0456	0.0187	0.0270	0.0124	0.0279	0.0136	0.0274	0.0129
9	0.0445	0.0096	0.0261	0.0065	0.0272	0.0068	0.0261	0.0065
10	0.0194	0.0330	0.0066	0.0069	0.0071	0.0074	0.0087	0.0073
11	0.0136	0.0135	0.0043	0.0035	0.0046	0.0038	0.0043	0.0036
12	0.0072	0.0051	0.0026	0.0018	0.0027	0.0019	0.0026	0.0019
Panel B: GED								
1	0.0282	0.0297	0.0237	0.0439	0.0197	0.0328	0.0190	0.0243
2	0.0219	0.0167	0.0159	0.0241	0.0128	0.0116	0.0136	0.0134
3	0.0131	0.0079	0.0077	0.0059	0.0071	0.0042	0.0076	0.0049
4	0.0293	0.0245	0.0246	0.0332	0.0211	0.0235	0.0233	0.0278
5	0.0181	0.0165	0.0107	0.0092	0.0106	0.0104	0.0119	0.0140
6	0.0136	0.0094	0.0077	0.0045	0.0077	0.0046	0.0080	0.0047
7	0.0466	0.0249	0.0337	0.0212	0.0330	0.0198	0.0336	0.0200
8	0.0459	0.0202	0.0339	0.0155	0.0334	0.0153	0.0335	0.0150
9	0.0458	0.0115	0.0324	0.0083	0.0322	0.0082	0.0329	0.0085
10	0.0143	0.0140	0.0096	0.0089	0.0093	0.0082	0.0086	0.0067
11	0.0084	0.0074	0.0051	0.0045	0.0049	0.0043	0.0052	0.0044
12	0.0054	0.0044	0.0036	0.0024	0.0036	0.0024	0.0036	0.0025
Panel C: Tempered stable distribution								
1	0.0148	0.0095	0.0076	0.0043	0.0076	0.0044	0.0076	0.0044
2	0.0153	0.0101	0.0123	0.0228	0.0093	0.0184	0.0051	0.0027
3	0.0233	0.0148	0.0050	0.0030	0.0053	0.0033	0.0052	0.0029
4	0.0140	0.0089	0.0084	0.0060	0.0075	0.0045	0.0077	0.0046
5	0.0221	0.0132	0.0051	0.0030	0.0052	0.0031	0.0056	0.0041

6	0.0143	0.0094	0.0093	0.0155	0.0074	0.0063	0.0056	0.0038
7	0.0442	0.0097	0.0307	0.0075	0.0304	0.0070	0.0303	0.0071
8	0.0452	0.0108	0.0247	0.0060	0.0259	0.0062	0.0232	0.0057
9	0.0460	0.0111	0.0252	0.0063	0.0265	0.0069	0.0238	0.0060
10	0.0058	0.0053	0.0033	0.0023	0.0031	0.0023	0.0031	0.0022
11	0.0073	0.0053	0.0028	0.0021	0.0028	0.0020	0.0025	0.0018
12	0.0069	0.0061	0.0027	0.0022	0.0026	0.0020	0.0024	0.0018

The subscript N, t, G and TS indicate that the statistic is for Normal, Student's t, GED and tempered stable distribution, respectively.

According to Equation (5), transition probabilities play an essential role in determining the smoothing probabilities. Hence, when the true distribution is fat-tailed, our previous findings suggest that MRS-N model may be the least accurate in identifying the latent states, whereas the three fat-tailed MRS models may have similar results. This is evident in Table 7. Comparing the mean absolute differences, MRS-N has the largest mean among all models across all cases. SEs of those averages are also the largest for MRS-N in almost all cases. In contrast, three fat-tailed models produce rather similar results with no material differences.

Hence, our results suggest that MRS-N model cannot identify the latent underlying Markov states as effectively as the fat-tailed MRS models. This is consistent with existing studies of Klaassen (2002), Ardia (2009), and Haas (2009), which suggest that for leptokurtic data, use of within-regime normality can seriously affect the identification of the regime process. This may be due to the fact that Normal distribution incorrectly treat negative and positive outliers as switching to low and high regimes, respectively. Therefore, the estimated transition probabilities are negatively biased, whereas the identified regime-switching is more frequent than the latent process.

To summarise, in terms of parameter estimates, when the true distribution is Student's t (GED), MRS-S model uniformly outperforms the competing models MRS-N and MRS-G (MRS-N and MRS-t). Also, the results of MRS-S and those of the true model are very close in most situations. Besides, MRS-S can even generate larger log-likelihood than the true model. When the true distribution is tempered stable, MRS-N, MRS-t and MRS-G cannot uniformly perform as well as the MRS-S model in the estimation of all parameters. Similar conclusions can be drawn when considering accuracy of estimated standard errors of parameters and the correctiveness of latent states identification. All the above observations are robust across various combinations of different parameter and sample size settings. We hereby argue that for a given financial dataset from an unknown fat-tailed distribution, the MRS-S model should be employed to accurately study its regime-switching properties.

5 Empirical results

To empirically compare MRS models with various distributions, we work on the dataset of the daily implied volatility of the S&P 500 options (VIX). The daily closing prices for VIX over the period between 1 January 2004 and 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). Additionally, since volatility is non-negative, we work on the logarithm transformation. Thus, no constraints to ensure non-negativity are needed.

The level of the logarithm of VIX is plotted in Figure 1. In the this plot, VIX dramatically increases from mid-2007, which is the start of global financial crisis (GFC) period. This suggests that the S&P 500 options return is more volatile during the GFC period. After 2010, the rough end of GFC, it tends to be less volatile with some turbulences around the end of 2010 and the beginning of 2012. Hence, with the presence of GFC and preliminary visual evidence, structural breaks may possibly exist in the VIX. Regarding the descriptive statistics, the mean and standard error of logarithm of VIX are 2.9175 and 0.3909, respectively. The skewness is 0.9741, indicating that the logarithm of VIX is moderately positively skewed. The excess kurtosis is 0.7911, which suggests that a non-Gaussian distribution may be appropriate. Thus, we perform the Kolmogorov–Smirnov and Jarque–Bera normality tests, where the null hypotheses indicating normality are rejected in both cases (p-values are 0.0000). Further, we perform the non-parametric change-point test proposed by Ross et al. (2011), which is robust for non-Gaussian data. The corresponding p-value is 0.0000, suggesting rejection of the null hypothesis assuming no structural breaks in the data. Consequently, it is interesting to study the regime-switching properties of VIX. As the data follow a non-Gaussian distribution, the MRS-N is not appropriate, and we will employ the MRS-t, MRS-G and MRS-S to investigate the logarithm of VIX. Besides, since its skewness is not negligible, we also consider MRS models with the skewed Student's t-distribution and GED for comparison.

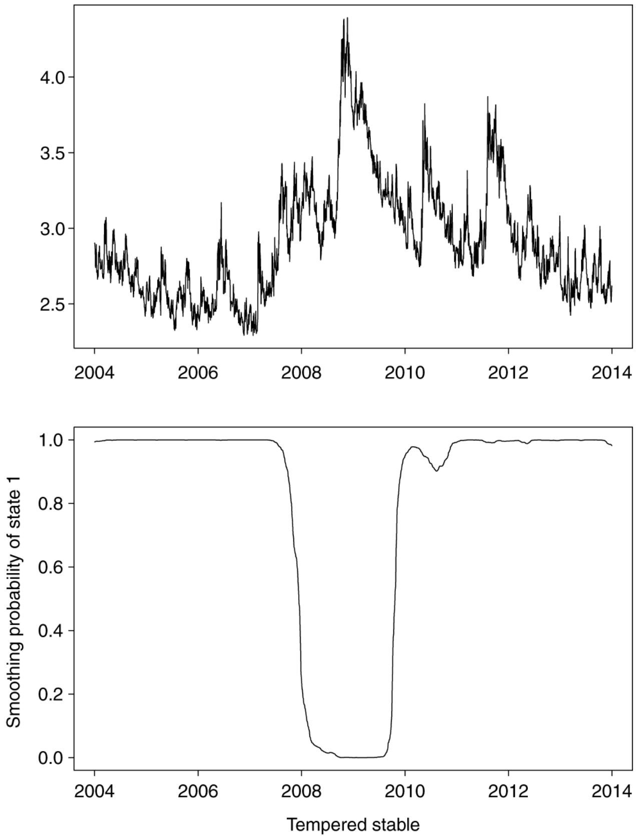

Figure 1: VIX index and identified states by the MRS-S model.

We fit the MRS models with (skewed) Student's t, (skewed) GED and tempered stable distributions for the logarithm of VIX. The results are presented in Table 8. Overall, all estimates are significant at 5% level in all models. Estimates of each parameter from MRS models with tempered stable, skewed Student's t (MRS-st) and skewed GED (MRS-sG) are quite close, which are somewhat different from those of MRS-t and MRS-G. More specifically, estimates of μ_1 and μ_2 are around 0.1 (0.07) and 0.12 (0.08), respectively, for MRS-S, MRS-st and MRS-sG (MRS-t and MRS-G). Both p_{11} and p_{22} are greater than 0.999 in all models, suggesting that the expectations of staying at both the low- and high-volatility states are quite long (Shi and Ho, 2015). The estimated standard deviation is slightly under 0.07 in all cases. Autoregressive parameter γ is over 0.95, suggesting that the log VIX index is quite persistent. To compare the model performance, log-likelihood, Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) are presented. It can be seen that MRS-t and MRS-G cannot perform as well as MRS-st and MRS-sG. More importantly, although MRS-S has the largest number of parameters, both AIC and BIC suggest it outperforms all the other four MRS models with fat-tailed distributions. Nevertheless, MRS-t and MRS-st outperforms MRS-G and MRS-sG, respectively.

Table 8: Empirical results: VIX index.

	Student's t	*GED*	*Skewed Student's t*	*Skewed GED*	*Temperd stable*
μ_1	0.0687	0.0704	0.0973	0.1012	0.0969
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
μ_2	0.0842	0.0871	0.1170	0.1166	0.1165
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
p_{11}	0.9995	0.9975	0.9995	0.9993	0.9995
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
p_{22}	0.9974	0.9957	0.9973	0.9981	0.9972
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
γ	0.9735	0.9716	0.9654	0.9635	0.9653
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
σ	0.0700	0.0661	0.0699	0.0661	0.0675
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
ν	3.4876	1.0382	3.2548	1.0827	
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	
α^+					0.3205

					(0.0000)
α^-					1.6108
					(0.0000)
λ^+					0.9452
					(0.0000)
λ^-					0.2404
					(0.0000)
p					0.6222
					(0.0000)
ξ			1.2556	1.2483	
			(0.0000)	(0.0000)	
Log Lik.	3426	3415	3459	3450	3474
AIC	−6839	−6815	−6901	−6883	−6926
BIC	−6798	−6774	−6855	−6837	−6862

Log Lik. stands for the log-likelihood. The values in the parentheses are the corresponding p-values.

To identify the underlying states over time, we plot the estimated smoothing probabilities of state 1 (low-volatility state) obtained from all models in Figure 1 and Figure 2. It is clear that the identified state patterns of VIX are roughly consistent across most models. It starts from the low-volatility state until the beginning of GFC. Since then, it switches to the high-volatility state and stays there for around 2 years. After that, MRS-S, MRS-t and MRS-st suggest that it returns to and stays at the low-volatility state. MRS-sG, however, indicates that the logged VIX stays at the high-volatility state for another year before switching back and staying at the low-volatility state. As for MRS-G model, the identified trend is not stable between 2010 and 2012. Despite the differences after 2010, these patterns are accordant with those of the real macroeconomic situation: the 2008 GFC causes high volatility, and its effects last for around 2 years.

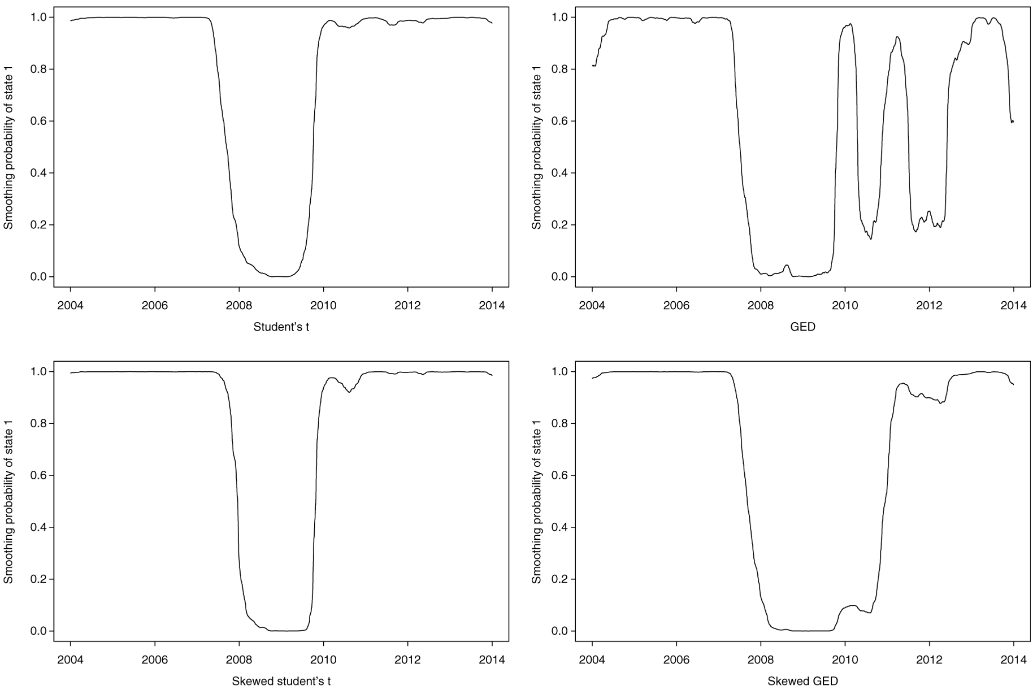

Figure 2: Smoothing probability of state 1 by the MRS models with (skewed) student' t-distribution and GED.

6 Conclusion

The MRS model has enjoyed particular popularity in the finance research related to structural breaks. However, the original MRS model is based on the Normal distribution, and its estimators will be inconsistent for fat-tail-distributed data. Unfortunately, the financial data are rarely Gaussian in practice. Hence, the sought of an appropriate distribution to accommodate their excess kurtosis becomes an essential issue for the application of the MRS-type model. Student's t-distribution and GED are two widely used fat-tailed densities in the existing literature. However, a recent study by Calzolari, Halbleib, and Parrini (2014) points out that due to the instability under aggregation, those distributions are not optimal choices. To overcome this problem, the α-stable distribution is introduced. Despite its attractive properties, the second moment of the α-stable distribution does not exist. Hence, it can cause more serious issues for the MRS model, which affect the validity of the asymptotic properties of its ML estimators and the interpretation of its estimated parameters.

To address those new problems, this paper suggests that the tempered stable distribution should be used instead of the α-stable. The reason is that tempered stable distribution retains all the attractive properties of the α-stable and has defined moments. Via three different simulation studies of the two-state MRS process, we systematically demonstrate the appropriateness of the tempered stable distribution applied within the MRS framework. The first two studies assume that the true distributions are the Student's t and GED, respectively. In these studies, results of MRS-S are close to those of the true models. Additionally, MRS-S generally outperforms its competitors (models other than the true specification) in terms of bias, efficiency and overall performance of the parameter estimation. Similar conclusions can be drawn for the estimated standard error and identified latent states. We consider three different combinations of the underlying tempered stable distribution in the third study. Our results suggest that none of the MRS-N, MRS-t and MRS-G can perform as well as the MRS-S model.

Finally, empirical evidence is further provided to evaluate the performance of MRS-S in practice. We fit the logarithm of daily VIX into MRS models with five fat-tailed distributions. Our results suggest that MRS-S is still preferred to the others, with respect to various model selection criteria. Also, the identified state structures of VIX obtained from the fitted models are largely accordant with the macroeconomic situation. Therefore, we argue that the tempered stable distribution could be widely used for modelling the financial data in general contexts with an MRS-type specification.

Existing studies including Mevel and Finesso (2004), Douc and Moulines (2012), and Pouzo, Psaradakis, and Sola (2016) have comprehensively analyzed the asymptotic properties of MLE under misspecified MRS models. Our results provide significant simulation evidence in the support of those studies. In particular, Pouzo, Psaradakis, and Sola (2016) point out that the MLE is consistent only for the parameters of the model having the lowest Kullback–Leibler divergence from the data-generating process. This may help explain the superiority of the MRS-S model under model misspecification. The relevant theoretical discussion remains for future research.

Acknowledgments

We are grateful to Macquarie University and Jiangxi University of Finance and Economics for their support. The authors would also like to thank Dave Allen, Felix Chan, Kin-Yip Ho, 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. We particularly thank the Editor-in-Chief Bruce Mizrach and one anonymous referee for providing valuable comments on earlier drafts. The usual disclaimer applies.

Notes

1 The estimators are inefficient if the underlying distribution of innovation sequence is not Gaussian.
2 In this case, the associated Levy processes are called "truncated Levy flights", the appropriateness of which to be applied in finance studies is discussed in Constantinides and Savel'ev (2013).
3 The performance comparison includes bias and efficiency of parameter estimates, correctiveness in identifying the underlying latent Markov states and accuracy of estimated standard errors as a measure of the sampling standard errors.
4 The second moment of the α-stable distribution only exists when $\alpha = 2$. In this case, the symmetric α-stable distribution collapses to a Gaussian distribution and cannot describe the fat fails.
5 Our MRS specification considered in this paper is netted by the general MRS autoregressive model discussed in Douc, Moulines, and Ryden (2004), where comprehensive asymptotic properties of MLE like consistency and normality are well established. Specific sufficient conditions relevant to those properties can be found in Douc, Moulines, and Ryden (2004).
6 As discussed in Mittnik, Doganoglu, and Chenyao (1999), discrete Fourier transform works most efficiently for N being expressed in terms of a power of 2.
7 The cases of regime-switching in variance only and in both mean and variance are also considered. The results are robust and available upon request.
8 We also consider the cases of 4 ,5 and 6 degrees of freedom. The results are robust and available upon request.
9 We also conduct studies with smaller transition probabilities of 0.25, 0.5, 0.75 and 0.9. The results are robust and available upon request.
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.

11 It is worth noticing that the estimated standard errors produced in this way are consistent only if the likelihood function is correctly specified.

References

Ardia, D. 2009. "Bayesian Estimation of a Markov-Switching Threshold Asymmetric GARCH Model with Student-t Innovations." *The Econometrics Journal* 12: 105–126.

Bianchi, M. L., S. T. Rachev, Y. S. Kim, and F. J. Fabozzi. 2010. "Tempered Stable Distributions and Processes in Finance: Numerical Analysis." In: *Mathematical and Statistical Methods for Actuarial Sciences and Finance*, edited by M. Corazza and C. Pizzi, 33–42. Italy: Springer.

Bollerslev, T. 1987. "A Conditional Heteroskedastic Time Series Model for Speculative Prices and Rates of Return." *Review of Economics and Statistics* 69: 542–547.

Calzolari, G., R. Halbleib, and A. Parrini. 2014. "Estimating GARCH-type Models with Symmetric Stable Innovations: Indirect Inference Versus Maximum Likelihood." *Computational Statistics & Data Analysis* 76: 158–171.

Carr, P., H. Geman, D. B. Madan, and M. Yor. 2002. "The Fine Structure of Asset Returns: An Empirical Investigation*." *The Journal of Business* 75: 305–333.

Carrasco, M., L. Hu, and W. Ploberger. 2014. "Optimal Test for Markov Switching Parameters." *Econometrica* 82: 765–784.

Cho, J. S., and H. White. 2007. "Testing for Regime Switching." *Econometrica* 75: 1671–1720.

Constantinides, A., and S. Savel'ev. 2013. "Modelling Price Dynamics: A Hybrid Truncated Lévy Flight-GARCH Approach." *Physica A* 392: 2072–2078.

Cont, R., and P. Tankov. 2004. *Financial Modelling with Jump Processes*. United Kingdom: Chapman and Hall/CRS Press.

Diebold, F. X., and A. Inoue. 2001. "Long Memory and Regime Switching." *Journal of Econometrics* 105: 131–159.

Douc, R., E. Moulines, and T. Ryden. 2004. "Asymptotic Properties of the Maximum Likelihood Estimator in Autoregressive Models with Markov Regime." *Annals of Statistics* 32: 2254–2304.

Douc, R., and E. Moulines. 2012. "Asymptotic Properties of the Maximum Likelihood Estimation in Misspecified Hidden Markov Models." *The Annals of Statistics* 40: 2697–2732.

Faias, J., and L. C. Nunes. 2012. "Testing Volatility Regime Switching." Working paper. Available at http://docentes.fe.unl.pt/~lcnunes/workingpapers/testing_volatility.pdf

Feng, L., and Y. Shi. 2016. "Fractionally Integrated Garch Model with Tempered Stable Distribution: A Simulation Study." *Journal of Applied Statistics* 44: 2837–2857.

Haas, M. 2009. "Value-at-Risk via Mixture Distributions Reconsidered." *Applied Mathematics and Computation* 215: 2103–2119.

Haas, M., and M. S. Paolella. 2012. "Mixture and Regime-Switching GARCH Models." In: *Handbook of Volatility Models and Their Applications*, edited by L. Bauwens, C. Hafner, and S. Laurent, 71–102. United Kingdom: Wiley.

Haas, M., S. Mittnik, and M. S. Paolella. 2004. "A New Approach to Markov-Switching GARCH Models." *Journal of Financial Econometrics* 2: 493–530.

Hamilton, J. D. 1988. "Rational-Expectations Econometric Analysis of Changes in Regime: An Investigation of the Term Structure of Interest Rates." *Journal of Economic Dynamics and Control* 12: 385–423.

Hamilton, J. D. 1989. "A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle." *Econometrica* 57: 357–384.

Hamilton, J. D. 1994. *Time Series Analysis*. Princeton: Princeton University Press.

Ho, K. Y., Y. Shi, and Z. Zhang. 2013. "How does News Sentiment Impact Asset Volatility? Evidence from Long Memory and Regime-Switching Approaches." *North American Journal of Economics and Finance* 26: 436–456.

Kim, Y. S., S. T. Rachev, M. L. Bianchi, and F. J. Fabozzi. 2008. "Financial Market Models with Lévy Processes and Time-Varying Volatility." *Journal of Banking & Finance* 32: 1363–1378.

Klaassen, F. 2002. "Improving GARCH Volatility Forecasts with Regime-Switching GARCH." *Empirical Economics* 27: 363–394.

Koponen, I. 1995. "Analytic Approach to the Problem of Convergence of Truncated Lévy Flights Towards the Gaussian Stochastic Process." *Physical Review E* 52: 1197.

Küchler, U., and S. Tappe. 2013. "Tempered Stable Distributions and Processes." *Stochastic Processes and their Applications* 123: 4256–4293.

Marcucci, J. 2005. "Forecasting Stock Market Volatility with Regime-Switching GARCH Models." *Studies in Nonlinear Dynamics & Econometrics* 9: 1–55.

Mevel, L., and L. Finesso. 2004. "Asymptotical Statistics of Misspecified Hidden Markov Models." *IEEE Transactions on Automatic Control* 49: 1123–1132.

Mittnik, S., T. Doganoglu, and D. Chenyao. 1999. "Computing the Probability Density Function of the Stable Paretian Distribution." *Mathematical and Computer Modelling* 29: 235–240.

Pouzo, D., Z. Psaradakis, and M. Sola. 2016. *Maximum Likelihood Estimation in Possibly Misspecified Dynamic Models with Time Inhomogeneous Markov Regimes*. Available at SSRN: https://ssrn.com/abstract=2887771.

Ross, G. J., N. M. Adams, D. K. Tasoulis, and D. J. Hand. 2011. "A Nonparametric Change Point Model for Streaming Data." *Technometrics* 53: 379–389.

Shi, Y., and K. Y. Ho. 2015. "Long Memory and Regime Switching: A Simulation Study on the Markov Regime-Switching ARFIMA Model." *Journal of Banking & Finance* 61: S189–S204.

Shi, Y., and L. Feng. 2016. "A Discussion on the Innovation Distribution of the Markov Regime-Switching GARCH Model." *Economic Modelling* 53: 278–288.

Stanley, H. E., V. Plerou, and X. Gabaix. 2008. "A Statistical Physics View of Financial Fluctuations: Evidence for Scaling and Universality." *Physica A* 387: 3967–3981.

Susmel, R., and R. F. Engle. 1994. "Hourly Volatility Spillovers between International Equity Markets." *Journal of International Money and Finance* 13: 3–25.
Tankov, P. 2003. Financial Modelling with Jump Processes. Vol. 2. Boca Raton, Florida: CRC press.
Wilfling, B. 2009. "Volatility Regime-Switching in European Exchange Rates Prior to Monetary Unification." *Journal of International Money and Finance* 28: 240–270.