Computational Statistics (2019) 34:1215–1245
https://doi.org/10.1007/s00180-019-00873-3

ORIGINAL PAPER

Two-sided exponential–geometric distribution: inference and volatility modeling

Emrah Altun[1]

Received: 12 September 2018 / Accepted: 16 February 2019 / Published online: 23 February 2019

Abstract
In this paper, *two-sided exponential–geometric* (TSEG) distribution is proposed and its statistical properties are studied comprehensively. The proposed distribution is applied to the GJR-GARCH model to introduce a new conditional model in forecasting Value-at-Risk (VaR). Nikkei-225 and BIST-100 indexes are analyzed to demonstrate the VaR forecasting performance of GJR-GARCH-TSEG model against the GJR-GARCH models defined under normal, Student-t, skew-T and generalized error innovation distributions. The backtesting methodology is used to evaluate the out-of-sample performance of VaR models. Empirical findings show that GJR-GARCH-TSEG model produces more accurate VaR forecasts than other competitive models.

Keywords GARCH · GJR-GARCH · Exponential–geometric distribution · Value-at-risk · Volatility

1 Introduction

In the last decade, financial institutions have been exposed to unpredictable big losses because of the economic instability and political events. To minimize the effects of unexpected events on financial institutions, risk management is an important tool to identify, measure and control the relevant risks that effect the business cycles. Value-at-Risk (VaR), plays an essential role in risk management systems and is a powerful risk measure. The common risk measure, VaR, is widely used to measure and quantify the level of risk for the single asset or portfolio under a given confidence level and holding period. The VaR could be explained as a quantile estimation of financial return series. Therefore, true quantile estimation of the financial returns is a key point for increasing the accuracy of VaR forecasts. The financial economists and researchers have shown great interest to develop new models in forecasting VaR. The objective of this research is to introduce one potential risk management

✉ Emrah Altun
emrahaltun123@gmail.com

[1] Department of Statistics, Bartin University, 74100 Bartin, Turkey

tool by proposing a flexible distribution to capture the non-normal characteristics of financial return series.

Cont (2001) summarized the stylized facts of the financial returns. It is an important research to understand the properties of financial returns. According to Cont (2001) (it is also widely documented in literature), the financial return series exhibit significant skewness and excess kurtosis. Thus, alternative statistical distributions, enables to model both skewness and excess kurtosis, are needed to increase the modeling ability of financial returns. Forecasting VaR with inaccurate distribution causes to underestimation or overestimation of the real market risk. In recent years, researchers has showed a great interest to more flexible distributions for modeling and forecasting the financial risk. Angelidis et al. (2004) discussed the performance of Generalized Autoregressive Conditional Heteroskedasticity (GARCH) models with three innovation distributions: normal, Student-t and generalized error distributions and concluded that the leptokurtic distributions increase the accuracy of VaR forecasts. Dendramis et al. (2014) suggested to use of parametric volatility models with skewed distributions to increase the accuracy of VaR forecasts. Lyu et al. (2017) examined the performance of GARCH model under eight innovation distributions and concluded that the financial institutions should take into consideration the more flexible distributions in their internal risk management system to increase the accuracy of their internal risk system. Chen et al. (2012) applied the asymmetric Laplace distribution, introduced by Hinkley and Revankar (1977), to GARCH models and concluded that asymmetric Laplace distribution provides better VaR forecasts than normal and Student-t distributions. So and Yu studied the seven GARCH models in VaR estimation and concluded that it is important to consider a model with fat-tailed errors in forecasting VaR. Recently, Altun et al. (2018, 2019) concluded that the skewed and fat-tailed distributions have an essential role for increasing the forecasting accuracy of VaR.

The goal of this study is to provide an alternative distribution to increase the modeling accuracy of financial return series. To achieve this goal, two-sided exponential–geometric (TSEG) distribution is proposed and applied to GARCH models. The usefulness of proposed distribution is demonstrated in forecasting VaR. The proposed distribution enables to model skewness and excess kurtosis simultaneously. This property of proposed distribution increases the accuracy of VaR forecasts.

The rest of the paper is organized as follows: in Sect. 2, the main statistical properties of TSEG distribution are obtained. GARCH models with different innovation distributions such as normal, Student-t, generalized error distribution and TSEG are presented in Sect. 3. Backtesting methodology is given in Sect. 4. Empirical findings and model comparisons are presented in Sect. 5. Some concluding remarks are given in Sect. 6.

2 Two-sided exponential–geometric distribution

The exponential–geometric (EG) distribution, introduced by Adamidis and Loukas (1998), has been studied by many researchers such as Louzada et al. (2011, 2014, 2016) and Bidram and Nadarajah (2016). The objective of these researches is to add additional shape parameters to EG distribution for increasing the modeling ability of EG distribution especially in lifetime data modeling. The probability density function (pdf) of EG distribution is given by

$$f(x;\beta,p) = \beta(1-p)\exp(-\beta x)(1-p\exp(-\beta x))^{-2},\ x>0 \tag{1}$$

where $\beta > 0$ and $p \in (0,1)$. The corresponding cumulative distribution function (cdf) to (1) is

$$F(x) = (1-\exp(-\beta x))(1-p\exp(-\beta x))^{-1}. \tag{2}$$

The raw moments of EG distribution are given by

$$E(X^r) = (1-p)r!(\beta^r p)^{-1}L(p;r). \tag{3}$$

where $L(p;r) = \sum_{j=1}^{\infty} p^j j^{-r}$ is the generalization of Euler's dilogarithm function of p, $L(p;2)$ (see, Erdelyi et al. 1953, p. 31). This function is also known as polylogarithm function and available in **copula** package of R software. The quantile function (qf) is widely used to generate random variables from statistical distributions. The qf of EG distribution is given by

$$Q(q) = -\frac{1}{\beta}\ln\left(\frac{q-1}{qp-1}\right) \tag{4}$$

where $0 \le q \le 1$. TSEG distribution is introduced by means of EG distribution.

Proposition 1 *Let X random variable follows TSEG distribution, denoted as* $X \sim TSEG(x;\beta_1,p)$*, with pdf given below*

$$f(x;\beta_1,p) = \begin{cases} (1-p)\exp(\beta_1 x)(1-p\exp(\beta_1 x))^{-2}, & x<0 \\ (1-p)\exp(-\beta_2 x)(1-p\exp(-\beta_2 x))^{-2}, & x \ge 0 \end{cases} \tag{5}$$

where $p \in (0,1)$ *is the scale parameter and* $\beta_1 > 1$ *and* $\beta_2 > 1$ *are the shape parameters. To ensure that the pdf (5) integrates to 1, the below constraint on parameters is required.*

$$\frac{1}{\beta_1} + \frac{1}{\beta_2} = 1 \tag{6}$$

The above constraint allows to being free only for one parameter. The parameter β_2 *is restricted by (6) and* $\beta_2 = \frac{\beta_1}{\beta_1 - 1}$*. Note that* $P(X<0) = 1/\beta_1$*. Therefore,* $\beta_1 > 1$*. If* $\beta_1 > 2$*, TSEG distribution is right skewed, otherwise, left skewed. If* $\beta_1 = 2$*, the distribution is symmetric around the zero.*

Figure 1 displays the plots of density functions of the TSEG distribution. As seen in Fig. 1, the pdf of TSEG distribution can be symmetric, left-skewed and right skewed for several parameter values. As seen from Fig. 1, the proposed distribution is a good choice to capture the stylized facts of conditional distribution of return series such as

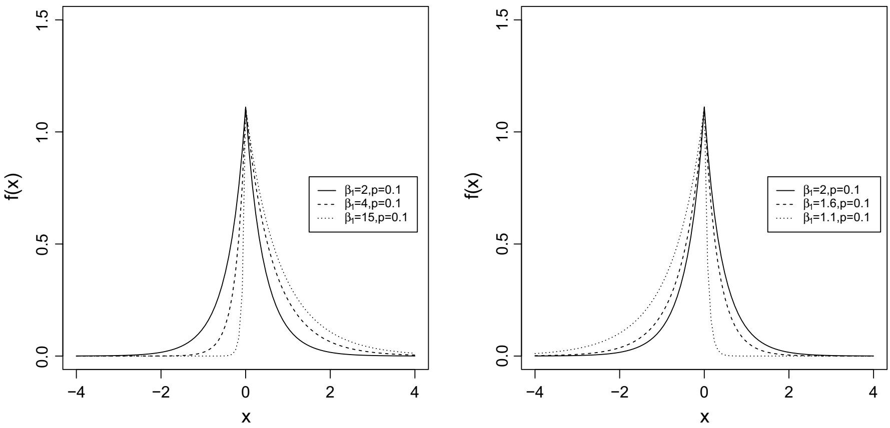

Fig. 1 Plots of density functions for the TSEG distribution for several parameter values

fat-tail and skewness. Accurately modeling of the tails of conditional distribution is essential to increase the accuracy of VaR forecasts.

2.1 Moments

Proposition 2 *The moments of the TSEG distribution can be obtained from the moments of EG distribution. Let* $X \sim TSEG(x;\beta_1,p)$, *then the rth moment of X is given by*

$$E(X^r) = \frac{1}{\beta_2}(1-p)r!\left(\beta_2^r p\right)^{-1}L(p;r) + (-1)^r\frac{1}{\beta_1}(1-p)r!\left(\beta_1^r p\right)^{-1}L(p;r) \tag{7}$$

The first four raw moments of the TSEG distribution are obtained using (7) as follows:

$$\begin{aligned} E(X) &= (1-p)L(p;1)\left(\frac{1}{\beta_2^2 p} - \frac{1}{\beta_1^2 p}\right) \\ E\left(X^2\right) &= 2(1-p)L(p;2)\left(\frac{1}{\beta_2^3 p} + \frac{1}{\beta_1^3 p}\right) \\ E\left(X^3\right) &= 6(1-p)L(p;3)\left(\frac{1}{\beta_2^4 p} - \frac{1}{\beta_1^4 p}\right) \\ E\left(X^4\right) &= 24(1-p)L(p;4)\left(\frac{1}{\beta_2^5 p} + \frac{1}{\beta_1^5 p}\right) \end{aligned} \tag{8}$$

Using the first four moments of the TSEG distribution, skewness and kurtosis can be obtained easily by, respectively,

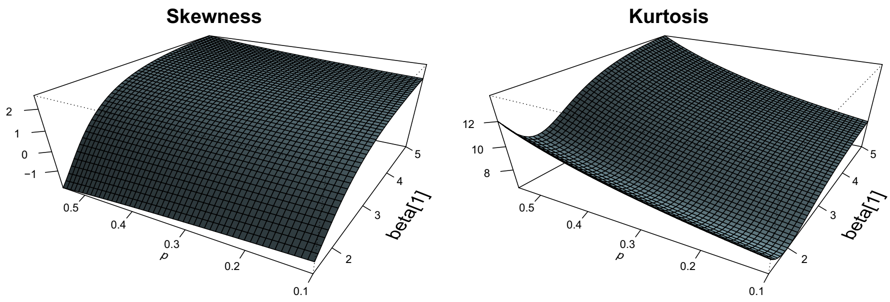

Fig. 2 Plots of skewness, and kurtosis of the TSEG distribution

$$\gamma_1 = \frac{E(X-\mu)^3}{\sigma^3},$$
$$\gamma_2 = \frac{E(X-\mu)^4}{\sigma^4} - 3,$$

where $\mu = E(X)$ and σ is given by,

$$\sigma = \sqrt{2(1-p)L(p;2)\left(\frac{1}{\beta_2^3 p} + \frac{1}{\beta_1^3 p}\right) - \left\{(1-p)L(p;1)\left(\frac{1}{\beta_2^2 p} - \frac{1}{\beta_1^2 p}\right)\right\}^2} \tag{9}$$

Figure 2 displays skewness, and kurtosis of TSEG distribution for some parameter values. As seen from Fig. 2, TSEG distribution could be used to model skewed and leptokurtic data sets.

Proposition 3 *Let γ_1 denotes the skewness of TSEG distribution. The range of γ_1 is given by,*

$$-2.612356 \le \gamma_1 \le 2.612356 \tag{10}$$

The range of γ_1 is obtained by means of numerical calculation.

2.2 Distribution function

Proposition 4 *Let $X \sim TSEG\left(x;\beta_1,p\right)$, then the cdf of X is given by*

$$F(x) = \begin{cases} \frac{1}{\beta_1}\left[1-\left(1-\exp\left(\beta_1 x\right)\right)\left(1-p\exp\left(\beta_1 x\right)\right)^{-1}\right], & x<0 \\ \frac{1}{\beta_1} + \frac{1}{\beta_2}\left(1-\exp\left(-\beta_2 x\right)\right)\left(1-p\exp\left(-\beta_2 x\right)\right)^{-1}, & x \ge 0 \end{cases} \tag{11}$$

2.3 Quantile function

Proposition 5 *Let* $X \sim TSEG(x;\beta_1,p)$*, then the qf of X is given by*

$$Q(q)=\begin{cases}\frac{1}{\beta_1}\ln\left(\frac{1-(1-q\beta_1)}{1-p(1-q\beta_1)}\right), & 0<q<1/\beta_1\\ -\frac{1}{\beta_2}\ln\left(\frac{1-\frac{\beta_2}{\beta_1}(q\beta_1-1)}{1-\frac{\beta_2}{\beta_1}p(q\beta_1-1)}\right), & 1/\beta_1\le q<1\end{cases} \tag{12}$$

where $0 \le q \le 1$.

Here, an algorithm is given for generating random variables from the TSEG distribution.

Proposition 6 *(Inverse Transform Algorithm)*

Let $Q(q)$ denotes the qf of TSEG distribution. The below algorithm can be used to generate random observations from $X \sim TSEG(x;\beta_1,p)$.

1. Generate $p \sim uniform(0,1)$,
2. If $0<q<1/\beta_1, X=\frac{1}{\beta_1}\ln\left(\frac{1-(1-q\beta_1)}{1-p(1-q\beta_1)}\right)$, otherwise $X=-\frac{1}{\beta_2}\ln\left(\frac{1-\frac{\beta_2}{\beta_1}(q\beta_1-1)}{1-\frac{\beta_2}{\beta_1}p(q\beta_1-1)}\right)$.

2.4 Standardized TSEG distribution

The standardized TSEG distribution is obtained using the transformed random variable $\varepsilon=(x-\mu)/\sigma$ where $E(\varepsilon)=0$ and $\mathrm{var}(\varepsilon)=1$. The random variable X with distribution given in (5) can be expressed as $x=\varepsilon\sigma+\mu$ and $\partial x/\partial\varepsilon=\sigma$. Thus, the pdf of standardized TSEG distribution is given by

$$f(\varepsilon;\beta_1,p)=\begin{cases}\sigma(1-p)\exp\left(\beta_1(\varepsilon\sigma+\mu)\right)\left(1-p\exp\left(\beta_1(\varepsilon\sigma+\mu)\right)\right)^{-2}, & \varepsilon<-\mu/\sigma\\ \sigma(1-p)\exp\left(-\beta_2(\varepsilon\sigma+\mu)\right)\left(1-p\exp\left(-\beta_2(\varepsilon\sigma+\mu)\right)\right)^{-2}, & \varepsilon\ge-\mu/\sigma\end{cases} \tag{13}$$

where $p \in (0,1)$, $\beta_1 > 1$ and $\beta_2 > 1$. Note that μ and σ are mean and standard deviation of TSEG distribution given in (7) and (9), respectively.

2.5 Estimation and simulation study

2.5.1 Estimation

Let $x_1,x_2,\dots,x_n$ be a random sample from $X \sim TSEG(x;\beta_1,p)$ distribution. Using Eq. (5), the log-likelihood function of TSEG distribution is given by

$$\ell(\Theta) = \left\{ n\ln(1-p) + \sum_{i=1}^{n} \beta_1 x_i - 2\sum_{i=1}^{n} \ln\left(1 - p\exp\left(\beta_1 x_i\right)\right) \right\} I_{(x_i<0)} + \left\{ n\ln(1-p) - \sum_{i=1}^{n} \beta_2 x_i - 2\sum_{i=1}^{n} \ln\left(1 - p\exp\left(-\beta_2 x_i\right)\right) \right\} I_{(x_i\geq 0)} \tag{14}$$

where $\Theta = \left(\beta_1, p\right)$ denotes the parameter vector and $I(\cdot)$ is the indicator function. Since the parameter β_2 is restricted by (6), there is no need to estimate this parameter. Taking partial derivatives from (14) with respect to parameters, the following normal equations are obtained as

$$\frac{\partial \ell}{\partial p} = \left\{ -\frac{n}{1-p} - 2\sum_{i=1}^{n} \frac{\exp\left(\beta_1 x_i\right)}{1 - p\exp\left(\beta_1 x_i\right)} \right\} I_{(x_i<0)} + \left\{ -\frac{n}{1-p} - 2\sum_{i=1}^{n} \frac{\exp\left(-\beta_2 x_i\right)}{1 - p\exp\left(-\beta_2 x_i\right)} \right\} I_{(x_i<0)} \tag{15}$$

$$\frac{\partial \ell}{\partial \beta_1} = \left\{ \sum_{i=1}^{n} x_i + 2\sum_{i=1}^{n} \frac{p x_i \exp\left(\beta_1 x_i\right)}{1 - p x_i \exp\left(\beta_1 x_i\right)} \right\} I_{(x_i<0)} \tag{16}$$

The maximum likelihood estimates (MLEs) of $\left(\beta_1, p\right)$, say, $\left(\hat{\beta}_1, \hat{p}\right)$, are the simultaneous solutions of the equations: $\frac{\partial \ell}{\partial \beta_1} = 0$, and $\frac{\partial \ell}{\partial p} = 0$. Since the likelihood equations contain non-linear functions, it is not possible to obtain explicit forms of the MLEs. Therefore, they have to be solved by using numerical methods. S-Plus, R or MATLAB can be used for obtaining the MLEs of the parameters. Here, **constrOptim** function of R software is used to minimize the minus log-likelihood function of TSEG distribution. The observed information matrix, $\boldsymbol{I}_F(\Theta)$ evaluated at $\hat{\Theta}$ is used to obtain corresponding standard errors. The elements of $\boldsymbol{I}_F(\Theta)$ is upon request from the authors.

2.5.2 Simulation study

In this subsection, Monte-Carlo simulation study is conducted to evaluate the performance of MLEs of TSEG distribution based on the $N = 10{,}000$ samples of sizes $n = 50, 150$ and 500 from TSEG distribution. Inverse transform algorithm given in Sect. 2.3 is used to generate random variables from TSEG distribution. The accuracy of the MLEs is discussed by means of the following measures: averages of the estimates (AEs), biases and mean square errors (MSEs). The simulation results are summarized in Table 1. Based on the results given in Table 1, the parameter estimates are quite stable and closed to their nominal values. The

Table 1 The AEs, biases and MSEs based on 10,000 simulations of the TSEG distribution for $n = 50, 150$ and 500

n	Parameters	$\beta_1 = 2 \quad p = 0.5$ AE	Bias	MSE	$\beta_1 = 1.3 \quad p = 0.5$ AE	Bias	MSE
50	β_1	2.0153	0.0153	0.0530	1.2998	− 0.0002	0.0059
	p	0.4892	− 0.0108	0.0168	0.4869	− 0.0131	0.0165
150	β_1	2.0132	0.0132	0.0166	1.2983	− 0.0017	0.0020
	p	0.4954	− 0.0046	0.0051	0.4973	− 0.0027	0.0048
500	β_1	2.0015	0.0015	0.0049	1.3023	0.0023	0.0006
	p	0.4994	− 0.0006	0.0015	0.4997	− 0.0003	0.0017
n	Parameters	$\beta_1 = 3 \quad p = 0.3$ AE	Bias	MSE	$\beta_1 = 4 \quad p = 0.7$ AE	Bias	MSE
50	β_1	3.0851	0.0851	0.3317	4.1629	0.1629	0.8608
	p	0.2927	− 0.0073	0.0250	0.6961	− 0.0039	0.0062
150	β_1	3.0352	0.0352	0.2997	4.0918	0.0918	0.2772
	p	0.2918	− 0.0082	0.0107	0.6947	− 0.0053	0.0018
500	β_1	3.0147	0.0147	0.0692	4.0184	0.0184	0.0696
	p	0.2997	− 0.0003	0.0032	0.6996	− 0.0004	0.0006

MSEs and biases approach to zero when the sample sizes increase. It is an evidence that the asymptotic normal distribution provides an adequate approximation for the finite sample distribution of MLEs of the TSEG distribution. Figure 3 displays the Q–Q plots of MLEs of TSEG distributions for $n = 500$. Figure 3 verifies the asymptotic normality property of MLE. It is clear that the MLEs of TSEG distribution are near to normal distribution.

3 GARCH models in VaR forecast

Let $r_t = \ln(p_t) - \ln(p_{t-1})$ denotes the daily log-returns and p_t is the closing price of asset at time t. The benchmark model, GARCH(1,1) introduced by Bollerslev (1986), is given by

$$\begin{aligned} r_t &= m_t + e_t, \\ e_t &= \varepsilon_t h_t,\ \varepsilon_t \sim i.i.d., \\ h_t^2 &= \omega + \gamma_1 e_{t-1}^2 + \gamma_2 h_{t-1}^2, \end{aligned} \tag{17}$$

where $\omega > 0, \gamma_1 > 0, \gamma_2 > 0$. Here, m_t and h_t^2 are the conditional mean and variance, respectively, and ε_t is the sequence of independently and identically distributed random variables with zero mean and unit variance. The mean process of log-returns are near the zero. Therefore, there is no need to use any autoregressive process for mean and it can be omitted or taken as a. The common choice is to take the mean process

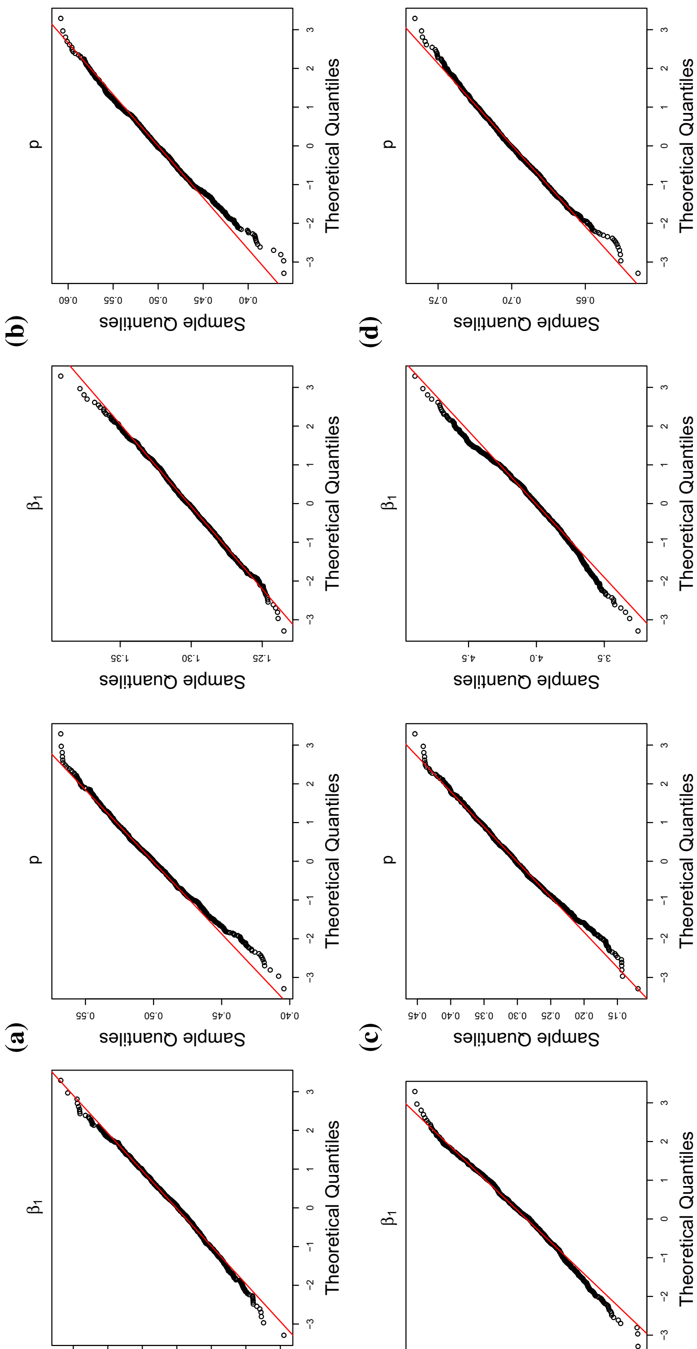

Fig. 3 QQ normality plots of MLEs for $n = 500$, **a** $\beta_1 = 2, p = 0.5$, **b** $\beta_1 = 1.3, p = 0.5$, **c** $\beta_1 = 3, p = 0.3$ and **d** $\beta_1 = 4, p = 0.3$

as zero. The h_t^2, conditional variance of e_t, is a time-varying, positive and measurable function of the information set at time $t-1$. When the $\gamma_1 + \gamma_2 < 1$, the process of e_t is covariance stationary and the unconditional variance of e_t is $\omega/\left(1-\gamma_1-\gamma_2\right)$. The GARCH model is generalization of ARCH model introduced by Engle (1982). When the parameter $\gamma_2 = 0$, GARCH model reduces to ARCH model.

ARCH and GARCH models do not enable to capture the asymmetric volatility dynamics in financial return series. For this reason, asymmetric GARCH models have been introduced. Here, the most used asymmetric volatility model, GJR-GARCH, introduced by Glosten et al. (1993), is used to model the asymmetric effects in volatility process. The GJR-GARCH(1,1) model is given by

$$h_t^2 = \omega + \gamma_1 e_{t-1}^2 + \gamma_3 I_{t-1} e_{t-1}^2 + \gamma_2 h_{t-1}^2, \tag{18}$$

where $\omega > 0, \gamma_1 > 0, \gamma_2 > 0$ and $\gamma_1 + \gamma_3 > 0$. When the innovation distribution is normal, $\gamma_1 + \gamma_2 + \frac{1}{2}\gamma_3 < 1$ for covariance stationary. The parameter γ_3 represents the leverage effect. I_{t-1} is an indicator function and $I_{t-1} = 1$ for $e_{t-1} < 0$, otherwise, $I_{t-1} = 0$. The parameter γ_3 represents the asymmetry effect on volatility. The positive γ_3 parameter indicates that the bad news yields higher volatility than good news. Note that when $\gamma_3 = 0$, GJR-GARCH model reduces to model of Bollerslev (1986). The unconditional variance of e_t for GJR-GARCH models is given by

$$Var\left(e_t\right) = \frac{\omega}{1-\gamma_1-\gamma_2-\kappa\gamma_3}, \tag{19}$$

where κ is

$$\kappa = E\left(I_{t-1} e_{t-1}^2\right) \int_{-\infty}^{0} f(\varepsilon) d\varepsilon. \tag{20}$$

It is easy to see that $\kappa = 1/2$ for standard normal distribution.

The distributional assumption on innovation process of volatility models directly affects the both accuracy of volatility and VaR forecasts. Therefore, the rest of this section is devoted to present GJR-GARCH models with normal and skewed and fat-tailed distributions.

3.1 Normal innovation distribution

The log-likelihood function of the r_t specified under normal innovations can be given by

$$\ell(\boldsymbol{\psi}) = -0.5\left(T \ln 2\pi + \sum_{t=1}^{T} \ln h_t^2 + \sum_{t=1}^{T} \varepsilon_t^2\right), \tag{21}$$

where $\boldsymbol{\psi} = (m, \omega, \gamma_1, \gamma_2, \gamma_3)$ denotes the parameter vector of the GJR-GARCH-normal (GJR-GARCH-N) model and $h_t^2 = \omega + \gamma_1 e_{t-1}^2 + \gamma_3 I_{t-1} e_{t-1}^2 + \gamma_2 h_{t-1}^2$.

The one-day-ahead VaR forecast based on normal distribution is given by

$$VaR_{t+1} = \hat{m}_{t+1} + F_q^{-1}\hat{h}_{t+1}, \tag{22}$$

where $\hat{m}_{t+1}$ and $\hat{h}_{t+1}$ are forecasts of mean and conditional standard deviation, respectively. F_q^{-1} is the quantile function (qf) of the normal distribution at the q level.

3.2 Student-t innovation distribution

Since financial return series have fatter tails than normal distribution, Bollerslev (1986, 1987) proposed the GARCH model with the Student-t innovations. GARCH model with the Student-t innovations enables to model both fat-tail and excess kurtosis observed in financial return series. The log-likelihood function of the GJR-GARCH-Student-t (GJR-GARCH-T) model is given as

$$\begin{aligned}\ell(\boldsymbol{\psi}) =& T\left[\ln\Gamma\left(\frac{\upsilon+1}{2}\right) - \ln\Gamma\left(\frac{\upsilon}{2}\right) - \frac{1}{2}\ln\left[\pi(\upsilon-2)\right]\right] \\ & - \frac{1}{2}\sum_{t=1}^{T}\left[\ln h_t^2 + (1+\upsilon)\ln\left(1 + \frac{\varepsilon_t^2}{\upsilon-2}\right)\right]\end{aligned}$$

where $\boldsymbol{\psi} = (m, \omega, \gamma_1, \gamma_2, \gamma_3, \upsilon)$ denotes the parameter vector, $\Gamma(\upsilon)$ is the gamma function and parameter υ controls the tails of the distribution. The one-day-ahead VaR forecast based on the Student-t distribution is given by

$$VaR_{t+1} = \hat{m}_{t+1} + F_q^{-1}(\varepsilon_t, \upsilon)\hat{h}_{t+1},$$

where $F_q^{-1}(\varepsilon_t, \upsilon)$ is the qf of the Student-t distribution at the q level. The **qt** function of R is used to obtain quantile estimation of standardized Student-t distribution.

3.3 Skew-T innovation distribution

Fernandez and Steel (1998) introduced the skew generalization of Student-t distribution, called as skew-T distribution. Lambert and Laurent (2001) applied the skew-T distribution to GARCH models. The log-likelihood function of the GJR-GARCH-skew-T (GJR-GARCH-ST) model is given as

$$\begin{aligned}\ell(\psi) =& T\left[\ln\left(\Gamma\left(\frac{\upsilon+1}{2}\right)\right) - \ln\left(\frac{\upsilon}{2}\right) - \frac{1}{2}\ln\left(\pi(\upsilon-2)\right)\right. \\ & \left. + \ln\left(\frac{2}{\kappa + 1/\kappa}\right) + \ln(s)\right] \\ & - \frac{1}{2}\sum_{i=1}^{T}\left[\ln\left(h_t^2\right) + (1+\upsilon)\ln\left(1 + \frac{\left(s\varepsilon_t + m\right)^2}{\upsilon-2}\kappa^{-2I_t}\right)\right]\end{aligned} \tag{23}$$

where $\boldsymbol{\psi} = (m, \omega, \gamma_1, \gamma_2, \gamma_3, \upsilon, \kappa)$ denotes the parameter vector, κ is the skewness parameter and,

$$I_t = \begin{cases} 1, \text{ if } z_t \geq -\frac{m}{s} \\ -1, \text{ if } z_t < -\frac{m}{s} \end{cases}$$

$$m = \frac{\Gamma(\upsilon + 1/2)\sqrt{\upsilon - 2}}{\sqrt{\pi}\Gamma(\upsilon/2)}(\kappa - 1/\kappa)$$

$$s = \sqrt{(\kappa^2 + 1/\kappa^2 - 1) - m^2}$$

The one-day-ahead VaR forecast of GJR-GARCH-ST model is given by

$$VaR_{t+1} = \hat{m}_{t+1} + F_q^{-1}(\varepsilon_t, \upsilon, \kappa)\hat{h}_{t+1}. \tag{24}$$

where $F_q^{-1}(\varepsilon_t, \upsilon, \kappa)$ is the qf of skew-T distribution at q level. The **qsstd** function of R is used to obtain quantile estimation of standardized skew-T distribution.

3.4 Generalized error innovation distribution

Nelson (1991) proposed the GED instead of assuming ε_t is normally distributed. Under this specification, the log-likelihood function of GJR-GARCH-GED model is given by

$$\ell(\boldsymbol{\psi}) = \sum_{t=1}^{T} \left[\ln\left(\frac{\kappa}{2}\right) - \frac{1}{2}\left|\frac{\varepsilon_t}{\delta}\right|^{\kappa} - (1 + \kappa^{-1})\ln(2) - \ln\Gamma\left(\frac{1}{2}\right) - \frac{1}{2}\ln\left(h_t^2\right) \right], \tag{25}$$

where $\boldsymbol{\psi} = (m, \omega, \gamma_1, \gamma_2, \gamma_3, \kappa)$ denotes the parameter vector, κ is tail-thickness parameter and,

$$\delta = \left(\frac{\Gamma\left(\frac{1}{\kappa}\right)}{2^{\frac{2}{\kappa}}\Gamma\left(\frac{3}{\kappa}\right)} \right)^{\frac{1}{2}}, \tag{26}$$

When the parameter $\kappa = 2$, the GED reduces to standard normal distribution. When $\kappa < 2$, GED has heavier tails than Gaussian distribution. The one-day-ahead VaR forecast of GJR-GARCH-GED model is given by

$$VaR_{t+1} = \hat{m}_{t+1} + F_q^{-1}(\varepsilon_t, \kappa)\hat{h}_{t+1}. \tag{27}$$

where $F_q^{-1}(\varepsilon_t, \kappa)$ is the qf of GED at q level. The **qged** function of R is used to obtain quantile estimation of standardized GED distribution.

3.5 Two-sided exponential–geometric innovation distribution

In this subsection, GJR-GARCH-TSEG model is introduced by means of standardized TSEG distribution with zero mean and unit variance as given in (13). The log-likelihood function of GJR-GARCH model with the TSEG innovation distribution is given by

$$\begin{aligned}\ell(\psi) =& \left\{ T\ln(\sigma(1-p)) + T\beta_1\mu + \beta_1\sigma\sum_{t=1}^{T}\varepsilon_t \right.\\ & \left. -2\sum_{t=1}^{T}\ln\left(1-p\exp\left(\beta_1\left(\varepsilon_t\sigma+\mu\right)\right)\right)\right\}_{I_{\varepsilon_t<-\mu/\sigma}}\\ & +\left\{ T\ln(\sigma(1-p)) - T\beta_2\mu - \beta_2\sigma\sum_{t=1}^{T}\varepsilon_t \right.\\ & \left. -2\sum_{t=1}^{T}\ln\left(1-p\exp\left(-\beta_2\left(\varepsilon_t\sigma+\mu\right)\right)\right)\right\}_{I_{\varepsilon_t\geq-\mu/\sigma}}\\ & -\frac{1}{2}\sum_{t=1}^{T}\ln\left(h_t^2\right)\end{aligned} \tag{28}$$

where $\boldsymbol{\psi} = (m, \omega, \gamma_1, \gamma_2, \gamma_3, \beta_1, p)$. When the GJR-GARCH model is used, the definition of conditional variance is $h_t^2 = \omega + \gamma_1 e_{t-1}^2 + \gamma_3 I_{t-1} e_{t-1}^2 + \gamma_2 h_{t-1}^2$. The stationary condition of GJR-GARCH model under TSEG distribution is given by

$$\gamma_1 + \gamma_2 + \kappa\gamma_3 < 1 \tag{29}$$

where

$$\begin{aligned}\kappa =& \int_{-\infty}^{0} f(\varepsilon)d\varepsilon\\ =&(p-1)\left\{ \left(\beta_1 p\right)^{-1} \right.\\ & \left. +\left(\beta_1 p\left[p\exp\left(\beta_1(1-p)L(p;1)\left(\frac{1}{\beta_2^2 p}-\frac{1}{\beta_1^2 p}\right)\right)-1\right]\right)^{-1}\right\}\end{aligned} \tag{30}$$

It is easy to see that $\kappa = \frac{1}{2}$ for symmetric distributions. The one-day-ahead VaR forecast based on TSEG distribution is given by,

$$VaR_{t+1} = \hat{m}_{t+1} + F_q^{-1}(\varepsilon_t, \beta_1, p)\hat{h}_{t+1}, \tag{31}$$

where $F_q^{-1}(\varepsilon_t, \beta_1, p)$ is the qf of TSEG distribution at q level. The quantile estimation of TSEG distribution is obtained by (12).

4 Evaluation of VaR forecasts

Backtesting methodology is used to compare VaR forecast accuracy of the models introduced in Sect. 3. Statistical accuracy of the models is evaluated by backtests of Kupiec (1995), Christoffersen (1998) and Sarma et al. (2003).

4.1 Unconditional coverage

Kupiec (1995) proposed a likelihood ratio (LR) test of unconditional coverage (LR_{uc}) to evaluate the model accuracy. The test examines whether the failure rate is equal to the expected value. The LR test statistic is given by

$$LR_{uc} = -2\ln\left[\frac{p^{n_1}(1-p)^{n_0}}{\hat{\pi}^{n_1}(1-\pi)^{n_0}}\right] \sim \chi_1^2,$$

where $\hat{\pi} = n_1/(n_0 + n_1)$ is the MLE of p, n_1 represents the total violation and n_0 represents the total non-violations forecasts. Violation means that if $VaR_t > r_t$, violation occurs, opposite case indicates the non-violation. Under the null hypothesis ($H_0 : p = \hat{\pi}$), the LR statistic follows a chi-square distribution with one degree of freedom.

4.2 Conditional coverage

The LR_{uc} test fails to detect if violations are not randomly distributed. Christoffersen (1998) proposed a LR test of conditional coverage LR_{cc} to remove the lack of Kupiec (1995) test. The LR_{cc} test investigates both equality of failure rate and expected one and also independently distributed violations. The LR_{cc} test statistic under the null hypothesis shows that the failures are independent and equal to the expected one. It is given by

$$LR_{cc} = -2\ln\left[\frac{(1-p)^{n_0}p^{n_1}}{\left(1-\pi_{01}\right)^{n_{00}}\pi_{01}{}^{n_{01}}\left(1-\pi_{11}\right)^{n_{10}}\pi_{11}{}^{n_{11}}}\right] \sim \chi_2^2,$$

where n_{ij} is the number of observations with value i followed by j for $i,j = 0, 1$ and $\pi_{ij} = n_{ij}/\sum_j n_{ij}$ are the probabilities, for $i,j = 1$. It denotes that the violation occurred, otherwise indicates the opposite case. The LR_{cc} statistic follows a chi-square distribution with two degrees of freedom.

4.3 Dynamic quantile test

Dynamic Quantile (DQ) test, proposed by Engle and Manganelli (2004) examines if the violations is uncorrelated with any variable that belongs to information set Ω_{t+1} when the VaR is calculated. The main idea of DQ test is to regress the current violations on past violations in order to test for different restrictions on the parameters of the model. The estimated linear regression model is given by

$$I_t = \beta_0 + \sum_{i=1}^{p} \beta_i I_{t-i} + \sum_{i=1}^{q} \mu_j X_j + \varepsilon_t \tag{32}$$

where

$$I_t = \begin{cases} 1, & \mathrm{r_t} < \mathrm{VaR_t} \\ 0, & \mathrm{r_t} \geq \mathrm{VaR_t} \end{cases} \tag{33}$$

This regression model tests whether the probability of violation depends on the level of the VaR. Here, p and q are used as 5 and 1, respectively, for illustrative purpose.

4.4 Loss functions

In most instances, evaluating the performance of VaR models by means of LR_{uc}, LR_{cc} and DQ tests may not be sufficient to decide the most adequate model among others. For instance, some models may have the same violation number with different forecast errors. In this case, the forecast errors of all candidate VaR models should be compared by means of loss functions. Here, two widely used loss functions are given to compare forecast errors of VaR models.

4.4.1 Regulator's loss function

Sarma et al. (2003) defined a test on the basis of *regulator's loss function* (RLF) to take into account differences between realized returns and VaR forecasts. The RLF is given by

$$RLF_{t+1} = \begin{cases} \left(r_{t+1} - VaR_{t+1}\right)^2, & \text{if } r_{t+1} < VaR_{t+1} \\ 0, & \text{if } r_{t+1} \geq VaR_{t+1} \end{cases}$$

where VaR_{t+1} represents the one-day-ahead VaR forecast for a long position.

4.4.2 Unexpected loss

The unexpected loss (UL) is equal to average value of differences between realized return and VaR forecasts. The one-day-ahead magnitude of the violation for long position is given by

Table 2 Summary statistics for the Nikkei-225 index

Descriptive statistics	Nikkei-225	BIST-100
Number of observations	1031	1228
Minimum	− 0.0825290	− 0.0734795
Maximum	0.0742620	0.0525509
Mean	0.0003440	0.0002808
Median	0.0005950	0.0007379
Std. devation	0.0128050	0.0129171
Skewness	− 0.2294740	− 0.3304695
Kurtosis	5.8478530	4.7327860
Jarque–Bera	1486.800 (< 0.001)	175.980 (< 0.001)
ARCH-LM test	75.683 (< 0.001)	26.493 (0.009)
Ljung-box	6.537 (0.7683)	7.521 (0.6755)
ADF	− 10.381 (< 0.001)	− 11.445 (< 0.001)

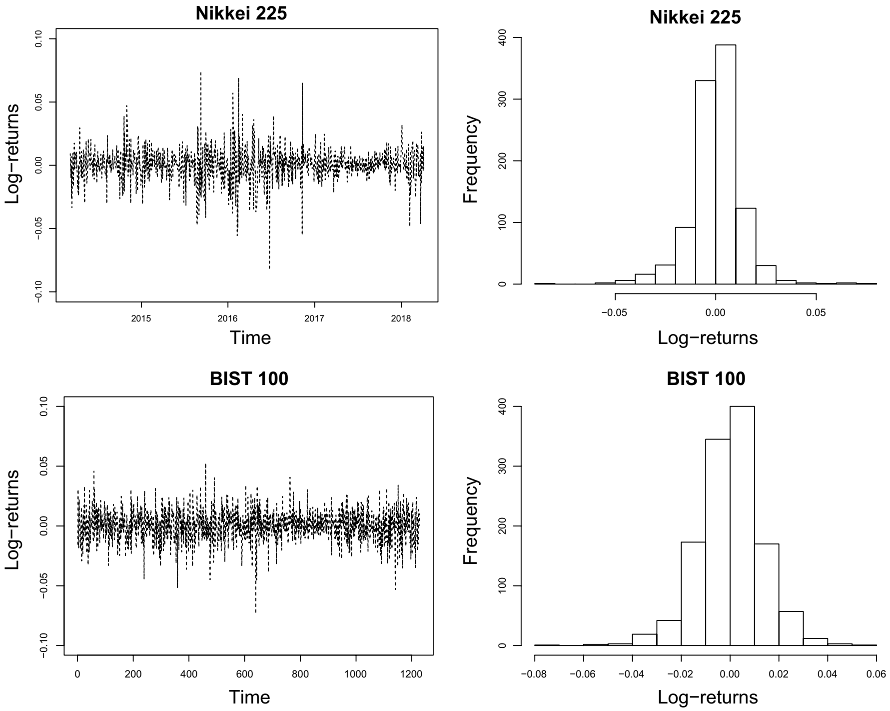

Fig. 4 Daily log-returns and corresponding histograms of Nikkei-225 (top) and BIST-100 (bottom) indexes

$$\mathrm{UL}_{t+1} = \begin{cases} \left(r_{t+1} - VaR_{t+1}\right), & \text{if} \quad r_{t+1} < VaR_{t+1} \\ 0, & \text{if} \quad r_{t+1} \geq VaR_{t+1} \end{cases}$$

5 Empirical results

5.1 Data description

To evaluate the performance of VaR models in forecasting daily VaR, the major indexes of Japan and Turkey, Nikkei-225 and BIST-100 indexes, respectively, are used. For two indexes, Nikkei-225 contains 1031 daily observations from 07.03.2014 to 06.04.2018 and BIST-100 contains 1228 daily observations from 03.01.2014 to 19.11.2018. The descriptive statistics for the log-returns of Nikkei-225 and BIST-100 indexes are given in Table 2. Figure 4 displays the daily log-returns of Nikkei-225 an BIST-100 indexes and corresponding histograms.

Table 2 shows that the mean returns are closed to 0. The Jarque–Bera statistic also shows that the null hypothesis of normality is rejected at any level of significance for both indexes. It is an evidence for high excess kurtosis and negative skewness. Thus, it is clear that log return of Nikkei-225 and BIST-100 indexes have non-normal characteristics, excess kurtosis, and fat tails. According to result of Ljung-box test, the null hypothesis of no autocorrelation up to 10th order is not rejected at any significance level for both indexes. Therefore, the used data sets have no autocorrelation problem. Finally, the result of Augmented Dickey–Fuller (ADF) test shows that the used time series are stationary.

The parameters of benchmark model, GJR-GARCH(1,1) specified under normal, Student-t, skew-T, GED and TSEG innovation distributions, are estimated by using the **constrOptim** function and **rugarch** package of R software. The textbfrugarch package is used to estimate parameters of GJR-GARCH-N, GJR-GARCH-T, GJR-GARCH-ST and GJR-GARCH-GED models. The log-likelihood function of GJR-GARCH-TSEG model is maximized by **constrOptim** function of R software. The initial values of parameters for GJR-GARCH-TSEG model are $\boldsymbol{\psi}_0 = (0.01, 0.1, 0.1, 0.89, 0.2, 2, 0.5)$. Tables 3 and 4 list the estimated parameters of GJR-GARCH(1,1) models for both indexes.

Tables 3 and 4 show that GJR-GARCH-TSEG model has the lowest log-likelihood value among others. It is clear that the TSEG distribution provides superior fits to standardized residuals of GJR-GARCH model. Based on the estimated parameters of GJR-GARCH-TSEG model, the conditional variance parameters γ_2 and γ_3 are obtained statistically significant at any significance level for both indexes. The estimated γ_3 parameter confirms the asymmetry effect on volatility. Therefore, the bad news yields higher volatility than good news. Figure 5 displays the P–P plots of the standardized residuals of GJR-GARCH models under normal and TSEG innovation distributions for Nikkei-225 and BIST-100 indexes. Based on the Fig. 5, it is concluded that use of TSEG distribution as an innovation process of GJR-GARCH yields more accurate modeling of the tails of conditional financial returns. The normal, Student-t, ST and GED innovation distributions fail

Table 3 Estimated parameters of GJR-GARCH(1, 1) model for the Nikkei-225 and assuming five different distributions for the standardized residuals, the corresponding standard errors are in second line and p-values in third line

Parameters	Nikkei-225				
	Normal	Student-t	GED	ST	TSEG
m	0.00039900	0.00061900	0.00050000	0.00031400	0.00012475
	0.00029300	0.00025500	0.00030200	0.00028600	0.00042897
	0.17395000	0.01526800	0.09743200	0.27271000	0.77118530
ω	0.00000500	0.00000600	0.00000500	0.00000600	0.00000787
	0.00000010	0.00000010	0.00000010	0.00000010	0.00000310
	< 0.0001	< 0.0001	< 0.0001	< 0.0001	0.01121836
γ_1	0.00214700	0.00309100	0.00276500	0.00268600	0.00012201
	0.00276500	0.00473700	0.00429900	0.00466200	0.01260354
	0.43751000	0.51413600	0.52009600	0.56460000	0.99227580
γ_2	0.84968700	0.82620300	0.84383300	0.82970400	0.80177570
	0.01289300	0.02147700	0.01816300	0.02035800	0.04260145
	< 0.0001	< 0.0001	< 0.0001	< 0.0001	< 0.0001
γ_3	0.23316200	0.28652100	0.24258500	0.28157200	0.36789590
	0.03318500	0.05613300	0.04554600	0.05361900	0.09518495
	< 0.0001	< 0.0001	< 0.0001	< 0.0001	0.00011106
υ	–	4.38363500	–	4.51627100	–
		0.60183800		0.65077700	
		< 0.0001		< 0.0001	
κ	–	–	1.18652100	0.90581900	–
			0.06654400	0.03751500	
			< 0.0001	< 0.0001	
β_1	–	–	–	–	1.83309000
					0.02812546
					<0.0001
p	–	–	–	–	0.00000025
					0.00000003
					< 0.0001
$-\ell$	− 3190.607	− 3238.017	− 3235.915	− 3240.814	− 3241.483
Ljung-box test					
Lag [1]	0.6746 (0.4115)	0.7734 (0.3792)	0.7130 (0.3985)	0.8652 (0.3523)	1.3288 (0.2490)
Lag [2]	0.7595 (0.5833)	0.8392 (0.5534)	0.7852 (0.5735)	0.9641 (0.5100)	1.5441 (0.4621)
Lag [5]	1.0009 (0.8595)	1.1183 (0.8324)	1.0245 (0.8541)	1.2181 (0.8087)	2.0516 (0.8420)
ARCH LM test					
Lag [3]	0.3102 (0.5776)	0.5605 (0.4541)	0.3717 (0.5421)	0.5531 (0.4570)	1.5861 (0.6625)
Lag [5]	1.5672 (0.5748)	1.9861 (0.4743)	1.6886 (0.5442)	1.9724 (0.4774)	4.2246 (0.5176)
Lag [7]	2.5722 (0.5977)	3.0604 (0.5022)	2.7567 (0.5606)	3.0023 (0.5131)	5.5463 (0.5936)

Table 4 Estimated parameters of GJR-GARCH (1, 1) model for the BIST-100 and assuming five different distributions for the standardized residuals, the corresponding standard errors are in second line and p values in third line

Parameters	BIST-100				
	Normal	Student-t	GED	ST	TSEG
m	0.000377	0.000522	0.00054	0.000334	0.000174116
	0.000343	0.00033	0.000323	0.000343	0.000851712
	0.272725	0.113675	0.094921	0.330182	0.838017012
ω	0.0000020	0.0000030	0.0000020	0.0000030	0.0000030
	0.0000010	0.0000010	0.0000010	0.0000001	0.0000047
	0.0014210	0.0000190	0.0088770	< 0.0001	0.5168298
γ_1	0.0000001	0.0000010	0.0000020	0.0000010	0.0000001
	0.0009170	0.0019160	0.0037660	0.0013710	0.0000015
	0.9999140	0.9994760	0.9995360	0.9992900	0.9402600
γ_2	0.9714810	0.9589250	0.9660830	0.9588170	0.9594033
	0.0099340	0.0047150	0.0043420	0.0045550	0.0438621
	< 0.0001	< 0.0001	< 0.0001	< 0.0001	< 0.0001
γ_3	0.0343260	0.0464750	0.0392120	0.0465390	0.0586954
	0.0100950	0.0112500	0.0081250	0.0115050	0.0149985
	0.0006730	0.0000360	0.0000010	0.0000520	0.0000910
υ	–	7.5371780	–	7.649428	–
		1.4136120		1.452825	
		< 0.0001		< 0.0001	
κ	–	–	1.4284380	0.923532	–
			0.0771590	0.036545	
			< 0.0001	< 0.0001	
β_1	–	–	–	–	1.8752540
					0.0434074
					< 0.0001
p	–	–	–	–	0.0000045
					0.0000014
					0.0011572
$-\ell$	− 3621.802	− 3646.067	− 3642.371	− 3648.067	− 3650.033
Ljung-box test					
Lag [1]	0.1024 (0.7489)	0.2037 (0.6518)	0.08263 (0.7738)	0.05235 (0.8190)	0.392 (0.5314)
Lag [2]	0.2430 (0.8283)	0.6967 (0.9235)	0.20637 (0.8503)	0.16700 (0.8749)	0.7469 (0.5881)
Lag [5]	0.5675 (0.9468)	2.4029 (0.8519)	0.52885 (0.9531)	0.48670 (0.9597)	1.7306 (0.6833)
ARCH LM test					
Lag [3]	1.4450 (0.2293)	0.7622 (0.3826)	1.124 (0.2891)	0.7138 (0.3982)	0.6713 (0.4115)
Lag [5]	1.5930 (0.5682)	0.8289 (0.7843)	1.219 (0.6693)	0.7880 (0.7967)	0.8341 (0.8021)
Lag [7]	2.9880 (0.5158)	2.2938 (0.6553)	2.650 (0.5819)	2.2653 (0.6612)	2.5472 (0.6041)

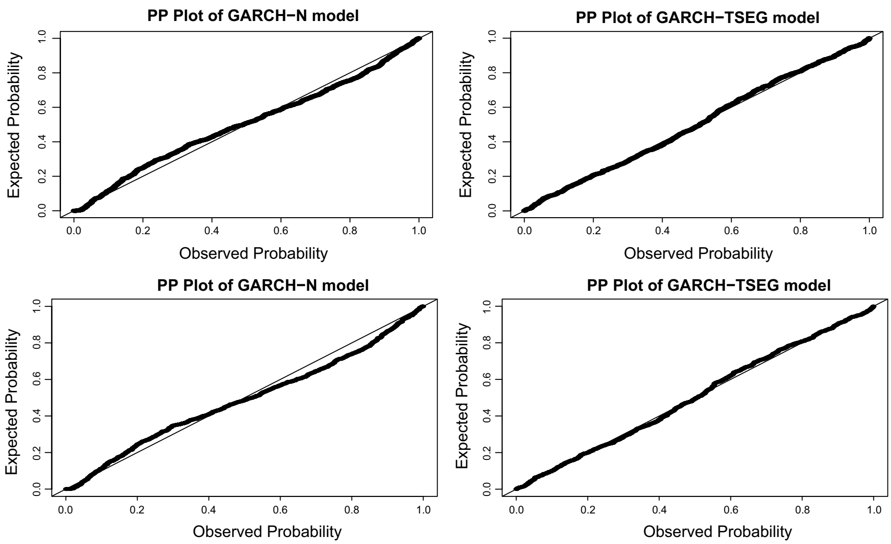

Fig. 5 P–P plots of GJR-GARCH-N and GJR-GARCH-TSEG models for Nikkei-225 (top) and BIST-100 indexes (bottom)

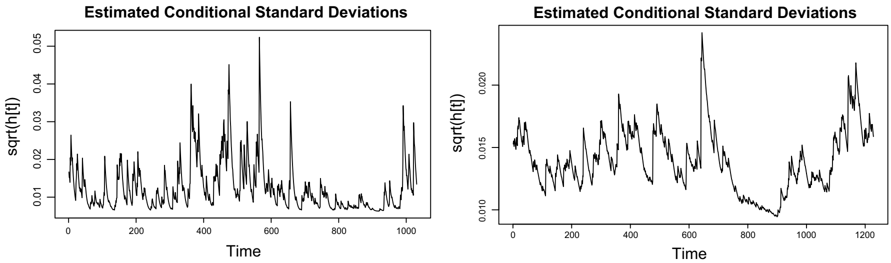

Fig. 6 Estimated conditional standard deviations of GJR-GARCH-TSEG model for Nikkei-225 (left) and BIST-100 indexes (right)

to represent the frequency of losses/gains and extreme events in tails of the conditional return series. Therefore, it is expected that forecasting VaR under normal, Student-t, and GED distributions yield to underestimated VaR forecasts for left-tail modeling. Figure 6 displays the estimated conditional standard deviations of GJR-GARCH-TSEG model for Nikkei-225 and BIST-100 indexes.

5.2 Backtesting results of Nikkei-225

Here, the out-of-sample performances of VaR models are compared based on the results of backtests and loss functions for Nikkei-225 index. The rolling window estimation method is used to estimate parameters of GJR-GARCH models and daily VaR forecasts. Rolling window estimation method gives an opportunity to Window length is determined as 331 and next 700 daily returns are used to evaluate the out of sample performance of VaR models. The below strategy is used to decide best model.

1. The VaR forecasts of all candidate models are obtained.
2. LR_{uc}, LR_{cc} and DQ tests are applied to decide which model produce consistent VaR forecasts for given q level.
3. The forecasting errors of VaR models, achieved to pass stage 2, are compared by means of loss functions.
4. The lowest values of the loss functions represent the best model.

Tables 5, 6 and 7 show the backtesting results of VaR models for following three quantile value: 0.05%, 0.025% and 0.1%, respectively. Based on the figures in Table 5, GJR-GARCH-N, GJR-GARCH-T, GJR-GARCH-ST, GJR-GARCH-GED and GJR-GARCH-TSEG models produce accurate VaR forecasts at 0.05% level on the basis of LR_{uc}, LR_{cc} and DQ results. Therefore, all VaR models achieve to pass stage 2. To decide the best model at 0.05% level, loss function results are investigated. Since the GJR-GARCH-TSEG model has the lowest values of average RLF (ARLF) and UL results at 0.05% level, it is the best model at 0.05% level and produces the most accurate VaR forecasts among others at 0.05% level.

Based on the figures in Table 6, GJR-GARCH-N, GJR-GARCH-T, GJR-GARCH-ST, GJR-GARCH-GED and GJR-GARCH-TSEG produce accurate VaR forecasts at 0.025% level on the basis of LR_{uc}, LR_{cc} and DQ results. Therefore, all VaR models also achieve to pass stage 2 for 0.025% level. The same strategy is used to decide best model. Since GJR-GARCH-TSEG model has the lowest forecasting error among others based on the ARLF and UL results, GJR-GARCH-TSEG model is chosen as the best model for 0.025% level.

Based on the figures in Table 7, all models produce underestimated VaR forecasts, except GJR-GARCH-TSEG model, at 0.01% level and fail to pass stage 2 on the basis of LR_{uc}, LR_{cc} and DQ results. GJR-GARCH model speficied under TSEG innovation distribution produce consistent VaR forecasts at 0.01% level. Moreover, GJR-GARCH-TSEG model has the lowest values of ARLF and UL results at 0.01% level. Therefore, GJR-GARCH-TSEG model is chosen as best model for 0.01% level. The reasons for success of TSEG distribution in modeling VaR can be summarized as follows: (i) TSEG distribution gives an opportunity for simultaneous modeling of skewness and excess kurtosis; (ii) provides more accurate representation for rare and extreme events; (iii) exhibits fat-tailed structure.

Figure 7 displays the VaR forecasts of GJR-GARCH models for Nikkei-225 index. This figure reveals that GJR-GARCH model with TSEG innovation distribution exhibits great consistency for estimating the true quantile value of conditional return series.

Table 5 Out-of-sample performances of VaR models for $p = 0.05$

Models	Mean VaR (%)	Failure rate	LR-uc	LR-cc	DQ	ARLF	UL
GJR-GARCH-N	− 2.026	0.049	0.030 (0.861)	1.027 (0.598)	5.993 (0.540)	0.1317194	− 0.05463311
GJR-GARCH-T	− 1.909	0.059	1.028 (0.310)	2.046 (0.359)	10.098 (0.183)	0.1353408	− 0.05989183
GJR-GARCH-ST	− 1.962	0.060	1.388 (0.238)	2.236 (0.327)	12.653 (0.08)	0.1382753	− 0.06034167
GJR-GARCH-GED	− 1.991	0.053	0.118 (0.731)	0.658 (0.731)	6.365 (0.497)	0.1296179	− 0.05585996
GJR-GARCH-TSEG	− 2.273	0.041	1.146 (0.284)	1.633 (0.442)	5.029 (0.656)	0.09478171	− 0.04136896

p values of LR-uc and LR-cc tests are presented in parentheses

Table 6 Out-of-sample performances of VaR models for $p = 0.025$

Models	Mean VaR (%)	Failure rate	LR-uc	LR-cc	DQ	ARLF	UL
GJR-GARCH-N	− 2.417	0.329	1.615 (0.203)	3.160 (0.205)	9.207 (0.238)	0.09444872	− 0.04002074
GJR-GARCH-T	− 2.462	0.034	2.223 (0.136)	3.541 (0.170)	9.840 (0.197)	0.08342246	− 0.03665815
GJR-GARCH-ST	− 2.570	0.034	2.223 (0.136)	3.541 (0.170)	10.038 (0.186)	0.08386006	− 0.03678371
GJR-GARCH-GED	− 2.523	0.030	0.675 (0.411)	2.743 (0.253)	9.633 (0.210)	0.0819643	− 0.03597006
GJR-GARCH-TSEG	− 2.975	0.020	0.768 (0.380)	1.951 (0.376)	4.377 (0.735)	0.04897422	− 0.02193595

p values of LR-uc and LR-cc tests are presented in parentheses

Table 7 Out-of-sample performances of VaR models for $p = 0.01$

Models	Mean VaR (%)	Failure rate	LR-uc	LR-cc	DQ	ARLF	UL
GJR-GARCH-N	− 2.870	0.024	10.313 (0.001)	13.763 (0.001)	38.753 (< 0.001)	0.06313683	− 0.02831519
GJR-GARCH-T	− 3.263	0.023	8.571 (0.003)	9.365 (0.009)	20.461 (0.004)	0.04079429	− 0.01913299
GJR-GARCH-ST	− 3.461	0.020	5.479 (0.019)	10.353 (0.005)	37.362 (< 0.001)	0.03874732	− 0.01953054
GJR-GARCH-GED	− 3.198	0.021	6.956 (0.008)	7.932 (0.018)	18.134 (0.013)	0.04495592	− 0.0205765
GJR-GARCH-TSEG	− 3.905	0.007	0.641 (0.423)	0.713 (0.700)	0.707 (0.998)	0.01995803	− 0.00910559

p values of LR-uc and LR-cc tests are presented in parentheses

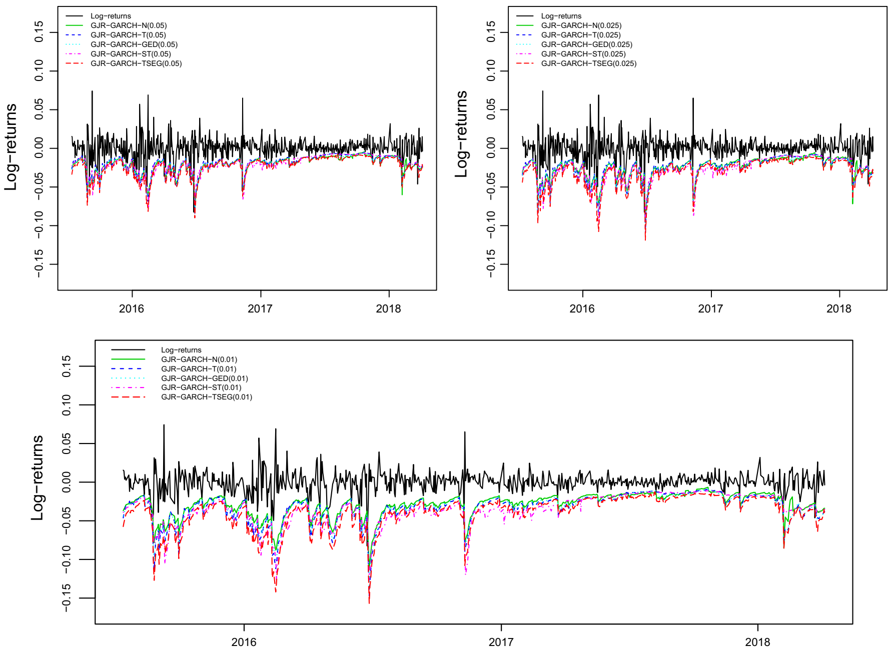

Fig. 7 Daily VaR forecasts of GJR-GARCH models for Nikkei-225 index

5.3 Backtesting results of BIST-100

Here, the results of out-of-sample performances of VaR models are compared for BIST-100 index. The same strategy with the Nikkei-225 index is used to decide the best VaR model for BIST-100 index. Tables 8, 9 and 10 show the backtesting results of VaR models for BIST-100 index. The following three quantile values are considered: 0.05%, 0.025% and 0.1%, respectively. Based on the figures in Table 8, as in Nikkei-225, all VaR models produce accurate VaR forecasts at 0.05% level on the basis of LR_{uc}, LR_{cc} and DQ results. So, all VaR models achieve to pass stage 2. It can be concluded that GJR-GARCH-TSEG model produce more accurate Var forecasts than other competitive models since it has the lowest values of ARLF and UL values at 0.05% level.

Based on the figures in Table 9, GJR-GARCH-T, GJR-GARCH-ST GJR-GARCH-GED and GJR-GARCH-TSEG produce accurate VaR forecasts at 0.025% level. GJR-GARCH-N model produces under-estimated VaR forecasts. As seen from Table 9, GJR-GARCH-TSEG model has the lowest values of loss functions results. Therefore, GJR-GARCH-TSEG model can be chosen as the best model for 0.025% level.

Based on the figures in Table 10, only GJR-GARCH-TSEG model produces the accurate VaR forecasts at 0.01% level and achieve to pass stage 2. As in Nikkei-225, GJR-GARCH-TSEG model provides better VaR forecasts than other competitive models at all quantiles. The superiority of TSEG distribution as an

Table 8 Out-of-sample performances of VaR models for $p = 0.05$

Models	Mean VaR (%)	Failure rate	LR-uc	LR-cc	DQ	ARLF	UL
GJR-GARCH-N	− 2.058	0.049	0.030 (0.861)	0.130 (0.937)	2.901 (0.893)	0.08619798	− 0.04556655
GJR-GARCH-T	− 1.983	0.053	0.118 (0.731)	0.123 (0.940)	3.397 (0.846)	0.08783585	− 0.04652643
GJR-GARCH-ST	− 1.954	0.051	0.029 (0.862)	0.053 (0.973)	5.214 (0.633)	0.09267473	− 0.05008057
GJR-GARCH-GED	− 2.002	0.049	0.030 (0.861)	0.130 (0.937)	2.326 (0.939)	0.08978827	− 0.04702769
GJR-GARCH-TSEG	− 2.306	0.037	2.664 (0.102)	2.669 (0.263)	5.799 (0.563)	0.06710882	− 0.03315294

p values of LR-uc and LR-cc tests are presented in parentheses

Table 9 Out-of-sample performances of VaR models for $p = 0.025$

Models	Mean VaR (%)	Failure rate	LR-uc	LR-cc	DQ	ARLF	UL
GJR-GARCH-N	− 2.461	0.039	4.549 (0.032)	5.393 (0.067)	9.748 (0.203)	0.05603735	− 0.02817416
GJR-GARCH-T	− 2.505	0.033	1.616 (0.203)	1.717 (0.423)	7.422 (0.386)	0.05060645	− 0.0245875
GJR-GARCH-ST	− 2.480	0.037	3.692 (0.054)	4.713 (0.094)	9.301 (0.232)	0.05154839	− 0.02657138
GJR-GARCH-GED	− 2.511	0.033	1.616 (0.203)	3.283 (0.193)	11.571 (0.115)	0.05240719	− 0.02565161
GJR-GARCH-TSEG	− 3.006	0.020	0.769 (0.380)	2.067 (0.355)	4.439 (0.727)	0.03592763	− 0.0141232

p values of LR-uc and LR-cc tests are presented in parentheses

Table 10 Out-of-sample performances of VaR models for $p = 0.01$

Models	Mean VaR (%)	Failure rate	LR-uc	LR-cc	DQ	ARLF	UL
GJR-GARCH-N	− 2.929	0.020	5.479 (0.019)	6.660 (0.035)	17.322 (0.015)	0.03558488	− 0.01587621
GJR-GARCH-T	− 3.222	0.009	0.151 (0.696)	4.508 (0.105)	20.416 (0.004)	0.02808731	− 0.01004373
GJR-GARCH-ST	− 3.206	0.013	0.529 (0.466)	3.254 (0.196)	22.397 (0.002)	0.0258652	− 0.01058993
GJR-GARCH-GED	− 3.145	0.016	1.966 (0.161)	3.951 (0.138)	18.001 (0.012)	0.02956035	− 0.01151757
GJR-GARCH-TSEG	− 3.932	0.004	2.939 (0.086)	2.965 (0.227)	2.896 (0.894)	0.01926053	− 0.0061429

p values of LR-uc and LR-cc tests are presented in parentheses

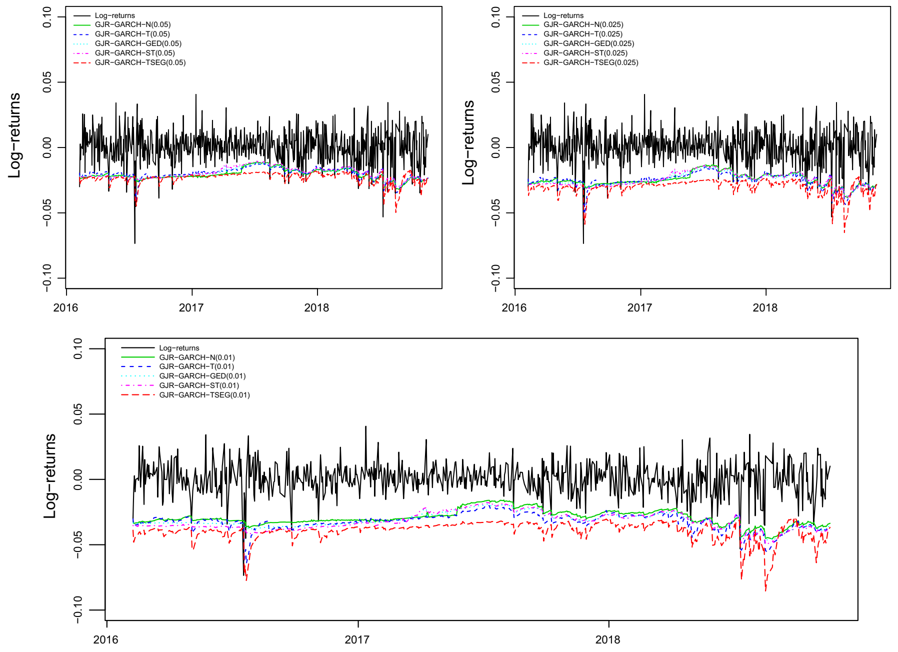

Fig. 8 Daily VaR forecasts of GJR-GARCH models for BIST-100 index

innovation process of GJR-GARCH models comes from its ability to model skewness and excess kurtosis simultaneously. The proposed distribution provides better representation for rare and extreme events than normal, Student-t, GED and skew-T distributions.

Figure 8 displays the VaR forecasts of GJR-GARCH models for BIST-100 index. As seen from Fig. 8, GJR-GARCH-N, GJR-GARCH-T, GJR-GARCH-ST and GJR-GARCH-GED models produce under-estimated VaR models. GJR-GARCH model with TSEG innovation distribution responds the volatility dynamics better than other models.

6 Conclusion

In this paper, a new skewed and fat-tailed distribution is proposed. The proposed distribution is applied to GJR-GARCH model as an innovation process. The common risk measure, VaR is modeled by GJR-GARCH-TSEG model. The usefulness of proposed financial risk model is demonstrated by means of real data applications on Nikkei-225 and BIST-100 indexes. The empirical findings of this study can be summarized as follows: Rolling window estimation method is used to obtain parameter estimation of GJR-GARCH model and VaR forecasts for in sample and out-of-sample periods. Empirical results show that GJR-GARCH model with TSEG innovation distribution produces more realistic VaR forecasts than normal, Student-t, skew-T and generalized error distributions for all confidence levels. Consequently, TSEG distribution opens new opportunities for modeling both skewness and excess kurtosis in financial return series. We hope that the results of this paper will be useful for practitioners and financial institutions.

References

Adamidis K, Loukas S (1998) A lifetime distribution with decreasing failure rate. Stat Probab Lett 39(1):35–42
Altun E, Tatlidil H, Özel G, Nadarajah S (2018) A new generalization of skew-T distribution with volatility models. J Stat Comput Simul 88(7):1252–1272
Altun E, Tatlıdil H, Özel G (2019) Conditional ASGT-GARCH approach to value-at-risk. Iran J Sci Technol Trans A Sci 43(1):239–247
Angelidis T, Benos A, Degiannakis S (2004) The use of GARCH models in VaR estimation. Stat Methodol 1(1–2):105–128
Bidram H, Nadarajah S (2016) A new lifetime model with decreasing, increasing, bathtub-shaped, and upside-down bathtub-shaped hazard rate function. Statistics 50(1):139–156
Bollerslev T (1986) Generalized autoregressive conditional heteroskedasticity. J Econom 31(3):307–327
Bollerslev T (1987) A conditionally heteroskedastic time series model for speculative prices and rates of return. Rev Econ Stat 69:542–547
Chen Q, Gerlach R, Lu Z (2012) Bayesian value-at-risk and expected shortfall forecasting via the asymmetric Laplace distribution. Comput Stat Data Anal 56(11):3498–3516
Christoffersen PF (1998) Evaluating interval forecasts. Int Econ Rev 39:841–862
Cont R (2001) Empirical properties of asset returns: stylized facts and statistical issues. Quant Finance 1:223–236
Dendramis Y, Spungin GE, Tzavalis E (2014) Forecasting VaR models under different volatility processes and distributions of return innovations. J Forecast 33(7):515–531
Engle RF (1982) Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. Econom J Econ Soc 50:987–1007
Engle RF, Manganelli S (2004) CAViaR: conditional autoregressive value at risk by regression quantiles. J Bus Econ Stat 22(4):367–381
Erdelyi A, Magnus W, Oberhettinger F, Tricomi FG (1953) Higher transcendental functions. McGraw-Hill, New York
Fernandez C, Steel MF (1998) On Bayesian modeling of fat tails and skewness. J Am Stat Assoc 93(441):359–371
Glosten LR, Jagannathan R, Runkle DE (1993) On the relation between the expected value and the volatility of the nominal excess return on stocks. J Finance 48(5):1779–1801
Hinkley DV, Revankar NS (1977) Estimation of the Pareto law from underreported data: a further analysis. J Econom 5(1):1–11
Kupiec PH (1995) Techniques for verifying the accuracy of risk measurement models. J Deriv 3(2):73–84
Lambert P, Laurent S (2001) Modelling financial time series using GARCH-type models and a skewed student density. Mimeo, Université de Liége, New York
Louzada F, Roman M, Cancho VG (2011) The complementary exponential geometric distribution: model, properties, and a comparison with its counterpart. Comput Stat Data Anal 55(8):2516–2524
Louzada F, Marchi V, Roman M (2014) The exponentiated exponential-geometric distribution: a distribution with decreasing, increasing and unimodal failure rate. Statistics 48(1):167–181
Louzada F, Ramos PL, Perdona GS (2016) Different estimation procedures for the parameters of the extended exponential geometric distribution for medical data. Comput Math Methods Med 2016:1–12
Lyu Y, Wang P, Wei Y, Ke R (2017) Forecasting the VaR of crude oil market: do alternative distributions help? Energy Econ 66:523–534
Nelson DB (1991) Conditional heteroskedasticity in asset returns: a new approach. Econom J Econom Soc 59:347–370
Sarma M, Thomas S, Shah A (2003) Selection of value-at-risk models. J Forecast 22(4):337–358

So MK, Philip LH (2006) Empirical analysis of GARCH models in value at risk estimation. J Int Financ Mark Inst Money 16(2):180–197