Review of Quantitative Finance and Accounting
https://doi.org/10.1007/s11156-019-00860-1

ORIGINAL RESEARCH

Multiday expected shortfall under generalized *t* distributions: evidence from global stock market

Robina Iqbal[1] · Ghulam Sorwar[1] · Rose Baker[1] · Taufiq Choudhry[2]

Abstract
We apply seven alternative *t*-distributions to estimate the market risk measures Value at Risk (VaR) and its extension Expected Shortfall (ES). Of these seven, the twin *t*-distribution (*TT*) of Baker and Jackson (in Twin *t* distribution, University of Salford Manchester. https://arxiv.org/abs/1408.3237, 2014) and generalized asymmetric distribution (*GAT*) of Baker (in A new asymmetric generalization of the t-distribution, University of Salford Manchester. https://arxiv.org/abs/1606.05203, 2016) are applied for the first time to estimate market risk. We analytically estimate VaR and ES over 1-day horizon and extend this to multi-day horizon using Monte Carlo simulation. We find that taken together *TT* and *GAT* distributions provide the best back-testing results across individual confidence levels and horizons for majority of scenarios. Moreover, we find that with the lengthening of time horizon, *TT* and *GAT* models performs well, such that at the 10-day horizon, *GAT* provides the best back-testing results for all of the five indices and the *TT* model provides the second best results, irrespective period of study and confidence level.

Keywords Generalize *t* distribution · Asymmetric *t* distribution · Expected shortfall · EGARCH models · Multi-days ahead expected shortfall

JEL Classification C13 · C15 · C51 · C52 · C53 · C58 · G17

✉ Ghulam Sorwar
g.sorwar@salford.ac.uk

Robina Iqbal
r.iqbal@edu.salford.ac.uk

Rose Baker
r.baker@edu.salford.ac.uk

Taufiq Choudhry
t.choudhry@soton.ac.uk

1 Salford Business School, University of Salford, Manchester M5 4WT, UK

2 Southampton Business School, University of Southampton, Highfield Campus, Southampton SO17 1BJ, UK

Published online: 14 January 2020

1 Introduction

From its very beginnings in the 1980s Value-at-Risk (VaR) as a measure of market risk has received widespread acceptance both amongst industry and regulators on account of its ease of calculation and intuitive interpretation. In its most basic form, VaR provides the worst possible loss at a given confidence level over a specific horizon. The main drawback of VaR, other than, that it is a single number is that there is no one accepted way of calculating it. It is possible that the use of different models will lead to different VaRs and that this could be very costly to financial institutions. In that, if VaR is over estimated, then the institution is tying of capital which it could use elsewhere for a higher return; or if it under estimates, then the firm is severely exposed to market down turns as it has not set aside correct amount of capital. The financial crisis of 2007–2008 has illustrated the drawbacks in stark terms of the VaR methodology and this has resulted in debate amongst academics, regulators and market practitioners. As part of this debate, the related measure to VaR, the expected shortfall (ES) is now given more prominence under Basel III.

Given the underlying nature of equity returns, forecasting of volatility is critical to the success of VaR models Siu (2018) and Chiou et al. (2009) amongst others. The volatility clustering resulting from infrequent large jump has been modelled using GARCH type process of Bollerslev (1986). This basic GARCH model leads to the development of more advanced models such as EGARCH, NGARCH, which are explicitly able to incorporate the skewness and excess kurtosis that are observed in equity returns.

To calculate VaR and ES, GARCH models need to be enhanced with more complex distributions. One such approach has been the use of the family of *t*-distributions. The student *t*-distributions have played particularly significant role in financial research as models for the distribution of heavy-tailed phenomena such as financial markets data. However, student *t*-distribution that allows for heavy tails than the normal, but assumes that the distribution is symmetric around zero. Huang and Lin (2014) compare the forecasting performance of several VaR models. Lin et al (2006) use historical simulation to estimate portfolio VaR. Baixaali and Alvarez (2006) consider the impact of excess kurtosis on VaR. Angelidis et al. (2007) examine different weighing schemes for robust VaR estimation. Wong et al (2012) model tail risk beyond VaR. The comparison focuses on the difference between normal distribution and student *t*-distribution. Mogel and Auer (2018) imply student *t* and extreme value theory to compute Value at Risk and compare them with historical simulation other approaches. Their results suggest that historical simulation outperforms EVT-based approach.

The student's *t*-distribution can permit for kurtosis in the conditional distribution but not for skewness. Hansen (1994) was the first to propose a generalization of student's *t*-distribution that allowed modelling skewness in conditional distributions of financial returns.

In this study we compare the performance of seven different *t*-distributions. The first is the standardized *t*-distribution (*ST*) used by Bollerslev (1987). The second is the Twin *t*-distribution (*TT*) of Baker and Jackson (2014). This distribution is heavy-tailed like a *ST* distribution but closer to the normality at the central part of the curve. The third distribution is the Generalized *t*-distribution (*GAT*) of Baker (2016). This distribution generalizes the *t*-distribution through two types of skewness. Fourth and fifth distributions are the Asymmetric exponential power distribution (*AEP*) and its special case (*SEP*) of Zhu and Zinde-Walsh (2009). The sixth and seventh distributions are the Asymmetric student *t*-distribution (*AST*) and Skewed student *t*-distribution (*SST*) respectively of Zhu and Galbraith (2010).

Our analysis focuses on datasets of five major stock indices covering S&P500, FTSE100, NASDAQ100, NIKKEI225 and DAX30 for the period 1995–2014. Calculation of 1-day ahead ES follows a two-stage procedure. In the first step, an asymmetric GARCH-type volatility model is fitted to the historical data by maximum likelihood estimation. From this model, the so-called standardized residuals are extracted. The asymmetric GARCH-type model is used to calculate 1-step predictions of conditional mean and conditional standard deviation. In the second step, various long tail and asymmetric distributions are applied to the standardized residuals and calculate with estimated parameters of distributions. Finally, 1 day ahead conditional expected shortfall ES_{t+1} is calculated.

For the situation where the variance is time varying, going from 1-day-ahead to *h*-days-ahead expected shortfall is not so straightforward. As in the case of GARCH, scaling by the horizon *h* is not attainable as variance mean revert. Additionally, the returns over the next *h* days are not normally distributed. To overcome this difficulty in calculating VaR and ES we use Monte Carlo simulation to generate the returns *h*- ahead.

We find overall EGARCH (1,1) provides the best fit for volatility for the indices considered in this study. We find substantial evidence in the improvement of our results with the use of EGARCH(1,1) combined with *GAT* and EGARCH(1,1) combined with *TTD*. When we compare the *GAT* distribution proposed by Baker (2016) with *AST* distribution proposed by Zhu and Galbraith (2010) we find *GAT* outperforms *AST* by providing better fit to financial returns and more accurate forecast of the ES. As the empirical distribution of the financial returns has been reported to be asymmetric and shows a significant excess of kurtosis (Abad et al. 2014). The longer period ES forecasts is estimated using Monte Carlo Simulation with *GAT*, *AEPD*, *SEPD*, *AST*, *SST*, *ST* and *TT* as standardized distributions of returns for world's major five stock indices (S&P500, FTSE100, NASDAQ100, NIKKEI225 and DAX30).

The contribution of this paper is as follows. First, our study provides further support for the usefulness and superiority of fat tailed distributions especially asymmetric distributions in the major stock markets. Second, it proposes the use of fat tailed distribution to measure financial risk for a longer horizon. In contrast to the current literature that mainly focuses on the 1 day ahead ES, our approach considers the usefulness of fat tail distribution for calculation of ES beyond 1-day. To the best of our knowledge, our research is the first to consider two new distributions and compare them with other previous distributions for ES calculation.

The remainder of this paper is organized as follow: Sect. 2 addresses the methodological framework. Results are discussed in Sect. 3. Section 4 concludes the findings.

2 Methodological framework

Since its inception in the 1980s, VaR and its extension the ES have been the market risk measure of choice both for industry and regulators. To calculate market risk, we follow the risk measure of Dowd et al. (2008) and define M_φ as follows:

$$M_\varphi = \int_0^1 \varphi(p) q_p dp \tag{1}$$

where q_p is the p loss quintile, $\varphi(p)$ is a weighting function defined over the full range of cumulative probabilities $p \in [0, 1]$ and M_φ is the class of quantile-based risk measures.

As noted by Dowd et al. (2008) VaR and ES constitute two well-known members of this class. The VaR at confidence level α with R_t as the index return in period t and Ψ_{t-1} represents the information available at time $t-1$ is defined as follows:

$$VaR_{\alpha} = q_{\alpha}\left(R_t|\Psi_{t-1}\right) \tag{2}$$

Moreover, each individual risk measure is characterised by its individual weighting function $\varphi(p)$. The weighting function for VaR is a Dirac delta function that gives the outcome $(p = \alpha)$ an infinite weight and zero weight for every other outcome.

The ES at confidence level α is the average of the worst $1-\alpha$ losses, which is defined as follows:

$$ES_a = \frac{1}{1-a}\int_a^1 q_p dp \tag{3}$$

The weighting function for ES gives all tail quantiles the same weight of $1/1-a$ and the non-tail quantiles zero weight.

We define an asset's return process at time t as follows:

$$R_t = \mu_t + \sigma_t z_t \tag{4}$$

where σ_t is the conditional volatility, μ_t is the conditional mean of returns and z_t is an independent and identically distributed random variable that follows alternative t-distributions.

The key challenge in calculating VaR and other market risk measures is the modelling and estimation of the conditional volatility that incorporates the observed characteristics of share price and index returns such as volatility clustering, asymmetry and long memory. Since its introduction by Bollerslev (1986), the GARCH approach to modelling volatility has become popular, resulting in a wide range of alternative GARCH specifications being proposed.

2.1 VaR and ES calculation over single period

Following Christoffersen (2012) the calculation of VaR and ES follows a two-stage procedure:

1. A GARCH-type volatility model is fitted to the historical data by maximum likelihood estimation (ML). From this model, the so-called standardized residuals are extracted. The GARCH-type model is used to calculate 1-step predictions of conditional mean $\left(\mu_{t+1}\right)$ and conditional standard deviation $\left(\sigma_{t+1}\right)$.
2. Various long tail and asymmetric distributions are applied to the standardized residuals to calculate $F^{-1}(p)$ with estimated parameters of the distributions. Finally, the 1-day ahead conditional VaR^p_{t+1} and conditional ES^p_{t+1} are calculated based on the following formulae:

$$VaR^p_{t+1} = -\mu_{t+1} - \sigma_{t+1}F^{-1}(p) \tag{5}$$

$$ES^p_{t+1} = -E_t\left[R_{t+1}|R_{t+1} < -VaR^p_{t+1}\right] \tag{6}$$

2.1.1 Standardized t-distribution

Bollerslev (1987) used the standardized t-distribution with $v > 2$. The standardized t-distribution density with $v > 2$ is then:

$$f_t(z, v) = \frac{\Gamma(\frac{1}{2}(v+1))}{\Gamma(v/2)\sqrt{\pi(v-2)}} \left(1 + \frac{z^2}{v-2}\right)^{-\left(\frac{1+v}{2}\right)} \tag{7}$$

where $\Gamma(v) = \int_0^\infty e^{-x} x^{v-1} dx$ is the gamma function. v is the parameter that describe the thickness of tails. Corresponding conditional VaR_{t+1}^p with t_p^{-1} as the pth quantile of student t-distribution and conditional ES_{t+1}^p are:

$$\begin{aligned} VaR_{t+1}^p &= -\mu_{t+1} - \sigma_{t+1}\sqrt{\frac{v-2}{v}} t_p^{-1}(v) \\ ES_{t+1}^p &= -\mu_{t+1} - \sigma_{t+1} ES_{t(v)}(p) \end{aligned} \tag{8}$$

where

$$ES_{t(v)}(p) = \frac{C(v)}{p} \left[\left[1 + \frac{1}{v-2} t_p^{-1}(v)\right]^{\frac{1-v}{2}} \frac{v-2}{1-v} \right]$$

$$\text{with } C(v) = \frac{\Gamma((v+1)/2)}{\Gamma(v/2)\sqrt{\pi(v-2)}}$$

The main drawback of the student t-distribution is that it is symmetrical while financial time series can be skewed.

2.1.2 Twin t-distribution (TT)

Baker and Jackson (2014) applied Johnson's transformation to statistical modelling and construct a new long tailed distribution that is like the t-distribution. The t like distribution is useful for fitting data, as it is more normal in the body of the distribution but has the same power law tail behavior.

The probability density function is:

$$f(x|v) = \frac{2^{5/2}\Gamma(v/4 + 3/2)}{\sqrt{\pi v}\Gamma(v/4)(v+1)} \left(x^2/v + \sqrt{1 + \left(1 + \left(x^2/v\right)\right)^2}\right)^{-(v+1)/2} \tag{9}$$

As $v \to \infty$ the distribution becomes standard normal. The distribution function for $x > 0$ is:

$$F_{TT}(x) = \frac{1}{2} + \frac{2^{3/2} x (S+C)^{-(v+1)/2}}{\sqrt{v}(v+1)B(v/4, 3/2)} + \left(\frac{1}{2}\right) I\left(1 - (C(x) + S(x))^{-2}; 3/2, v/4\right) \tag{10}$$

where $S = \frac{x^2}{v}$, $C = \sqrt{1 + S^2}$, B is the beta function and I the regularized incomplete beta function.

Conditional VaR_{t+1}^p and ES_{t+1}^p of TT are:

$$
\begin{aligned}
VaR_{t+1}^{p} &= -\mu_{t+1} - \sigma_{t+1} VaR_{TT}(p|\nu) \\
ES_{t+1}^{p} &= -\mu_{t+1} - \sigma_{t+1} ES_{TT}(p|\nu)
\end{aligned} \tag{11}
$$

where

$$
VaR_{TT}(p|\nu) = F_{TT}^{-1}(p|\nu)
$$

F_{TT}^{-1} is the inverse of cdf F_{TT}.

$$
ES_{TT}(p|\nu) = -E_t\left[R_{t+1}|R_{t+1} < -VaR_{TT}(p|\nu)\right]
$$

2.1.3 Generalized asymmetric t-distribution (GAT)

A 6-parameter asymmetric fat-tailed distribution (*GAT*) is proposed by Baker (2016). The pdf of the *GAT* is:

$$
\begin{aligned}
&f_{GAT}(x|\mu, \phi, \alpha, r, c, \nu) \\
&= \frac{\alpha\left(1+r^2\right)}{r\phi} \frac{\left\{(cg((x-\mu)/\phi))^{\alpha r} + (cg((x-\mu)/\phi))^{-\alpha/r}\right\}^{-\nu/\alpha}}{B\left(\frac{\nu/a}{1+r^2}, \frac{r^2\nu/\alpha}{1+r^2}\right)} \left(1+((x-\mu)/\phi)^2\right)^{-1/2}
\end{aligned} \tag{12}
$$

where B is the beta function, $\nu > 0$ controls tail power, μ is a centre of location (not necessarily the mean), $\phi > 0$ is a measure of scale (but not the variance, which may not exist), $r > 0$ controls tail power asymmetry, $c > 0$ controls the scale asymmetry, and $\alpha > 0$ controls how early 'tail behaviour' is apparent.

The cdf of the *GAT* distribution is:

$$
F_{GAT}(x|\mu, \phi, \alpha, r, c, \nu) = B\left(\frac{\nu}{\alpha\left(1+r^2\right)}, \frac{\nu r^2}{\alpha\left(1+r^2\right)}; q(x)\right) \tag{13}
$$

where

$$
\begin{aligned}
&q(x) \\
&= \frac{1}{1 + c^{-\alpha\left(1+r^2\right)/r}\left\{\frac{(x-\mu)}{\phi} + \sqrt{1+\frac{(x-\mu)^2}{\phi^2}}\right\} - \alpha\left(1+r^2\right)/r}
\end{aligned}
$$

Conditional VaR_{t+1}^{p} and ES_{t+1}^{p} of *GAT* are:

$$
\begin{aligned}
VaR_{t+1}^{p} &= -\mu_{t+1} - \sigma_{t+1} VaR_{GAT}(p|\mu, \phi, \alpha, r, c, \nu) \\
ES_{t+1}^{p} &= -\mu_{t+1} - \sigma_{t+1} ES_{GAT}(p|\mu, \phi, \alpha, r, c, \nu)
\end{aligned} \tag{14}
$$

where

$$
VaR_{GAT}(p|\mu, \phi, \alpha, r, c, \nu) = F_{GAT}^{-1}(p|\mu, \phi, \alpha, r, c, \nu)
$$

and F_{GAT}^{-1} is the inverse of cdf F_{GAT}.

$$ES_{GAT}(p|\mu,\phi,\alpha,r,c,v) = -E_t\left[R_{t+1}|R_{t+1} < -VaR_{GAT}(p|\mu,\phi,\alpha,r,c,v)\right]$$

2.1.4 The asymmetric exponential power distribution (AEP)

The asymmetric exponential power distribution is proposed by Zhu and Zinde-Walsh (2009).

$$f_{AEP}(x|\beta) = \begin{cases} \left(\frac{\alpha}{\alpha^*}\right)\frac{1}{\sigma}K_{EP}(d_1)exp\left(-\frac{1}{d_1}\left|\frac{x-\mu}{2\alpha^*\sigma}\right|^{d_1}\right), & x \le \mu \\ \left(\frac{1-\alpha}{1-\alpha^*}\right)\frac{1}{\sigma}K_{EP}(d_2)exp\left(-\frac{1}{d_2}\left|\frac{x-\mu}{2(1-\alpha^*)\sigma}\right|^{d_2}\right), & x > \mu \end{cases} \tag{15}$$

where $\beta = (\alpha, d_1, d_2, \mu, \sigma)$ is parameter vector, $\mu \in R$ and $\sigma > 0$ is still location and scale parameters respectively,$\alpha \in (0,1)$ is skewness parameter.$d_1 > 0$ and $d_2 > 0$ are left and right tail parameters respectively, $K_{EP}(d)$ is the normalizing constant is:

$$K_{EP}(d) \equiv \frac{1}{\left[2d^{1/d}\Gamma\left(1+{}^1/_d\right)\right]}$$

and α^* is:

$$\alpha^* = \alpha K_{EP}(d_1)/\left[\alpha K_{EP}(d_1)+\right](1-\alpha)K_{EP}(d_2)$$

Note that:

$$\left(\frac{\alpha}{\alpha^*}\right)K_{EP}(d_1) = \left(\frac{1-\alpha}{1-\alpha^*}\right)K_{EP}(d_1) = \left[\alpha K_{EP}(d_1)+\right](1-\alpha)K_{EP}(d_2)$$

The *AEP* density function is still continuous at every point and unimodal with mode at μ. The parameter α^* in the *AEP* density provides scale adjustments respectively to the left and right parts of the density to ensure continuity of the density under changes of shape parameters (α, d, d_2).

The VaR and ES is computed analytically for the *AEP* distribution in Zhu and Galbraith (2011).

Conditional VaR^p_{t+1} conditional ES^p_{t+1} of *AEP* are:

$$\begin{aligned} VaR^p_{t+1} &= -\mu_{t+1} - \sigma_{t+1}VaR_{AEP}(p|\alpha,d_1,d_2) \\ ES^p_{t+1} &= -\mu_{t+1} - \sigma_{t+1}ES_{AEP}(p|\alpha,d_1,d_2) \end{aligned} \tag{16}$$

where

$$VaR_{AEP}(p|\alpha,d_1,d_2) = \begin{cases} -2\alpha^*\left[d_1Q^{-1}\left(\frac{p}{\alpha},\frac{1}{d_1}\right)\right]^{\frac{1}{d_1}}, & p \le \alpha \\ 2(1-\alpha^*)\left[d_2Q^{-1}\left(\frac{1-p}{1-\alpha},\frac{1}{d_2}\right)\right]^{1/d_2}, & p > \alpha \end{cases}$$

$Q(\alpha, x)$ denotes the regularized complementary incomplete gamma function:

$$Q(\alpha, x) = \int_x^{\infty} t^{\alpha-1} exp(-t)dt/\Gamma(\alpha)$$

Q^{-1} denotes the inverse of $Q(\alpha, x)$ and Γ is gamma function:

$$ES_{AEP}(p|\alpha, d_1, d_2) = -\frac{2\alpha^*}{p} \int_0^p \left[d_1 Q^{-1}\left(\frac{p}{\alpha}, \frac{1}{d_1} \right) \right]^{\frac{1}{d_1}} dp + \frac{2(1-\alpha^*)}{p} \int_0^p \left[d_2 Q^{-1}\left(\frac{1-p}{1-\alpha}, \frac{1}{d_2} \right) \right]^{\frac{1}{d_2}} dp \tag{17}$$

2.1.5 Skewed exponential power distribution (SEP)

Skewed is the special case of *AEP* proposed by Zhu and Zinde-Walsh (2009), if $d_2 = d_1 = d$ implying $\alpha = \alpha^*$ The *AEP* reduced to *SEP*:

$$f_{SEP}(x|\beta) = \begin{cases} \frac{1}{\sigma} K_{EP}(d) exp\left(-\frac{1}{d} \left| \frac{x-\mu}{2\alpha\sigma} \right|^d \right), & x \leq \mu \\ \frac{1}{\sigma} K_{EP}(d) exp\left(-\frac{1}{d} \left| \frac{x-\mu}{2\alpha\sigma} \right|^d \right), & x > \mu \end{cases} \tag{18}$$

The *SEP* density is skewed to the right for $\alpha < 1/2$ and to the left for $\alpha < 1/2$.
Conditional VaR^p_{t+1} and ES^p_{t+1} of *SEP* are:

$$\begin{aligned} VaR^p_{t+1} &= -\mu_{t+1} - \sigma_{t+1} VaR_{SEP}(p|\alpha, d) \\ ES^p_{t+1} &= -\mu_{t+1} - \sigma_{t+1} ES_{SEP}(p|\alpha, d) \end{aligned} \tag{19}$$

where

$$VaR_{SEP}(p|\alpha, d) = \begin{cases} -2\alpha^* \left[d_1 Q^{-1}\left(\frac{p}{\alpha}, \frac{1}{d} \right) \right]^{\frac{1}{d}}, & p \leq \alpha \\ 2(1-\alpha^*) \left[dQ^{-1}\left(\frac{1-p}{1-\alpha}, \frac{1}{d} \right) \right],^{1/d} & p > \alpha \end{cases}$$

$$ES_{SEP}(p|\alpha, d) = -\frac{2\alpha^*}{p} \int_0^p \left[dQ^{-1}\left(\frac{p}{\alpha}, \frac{1}{d} \right) \right]^{1/d} dp + \frac{2(1-\alpha^*)}{p} \int_0^p \left[dQ^{-1}\left(\frac{1-p}{1-\alpha}, \frac{1}{d} \right) \right]^{1/d} dp$$

2.1.6 Asymmetric student t-distribution (AST)

AST proposed by Zhu and Galbraith (2010) and density function is defined as:

$$f_{AST}(x|\beta) = \begin{cases} \left(\frac{\alpha}{\alpha^*} \right) K(v_1) \left[1 + \frac{1}{v_1} \left(\frac{x}{2\alpha^*} \right)^2 \right]^{-\frac{v_1+1}{2}}, & x \leq 0 \\ \left(\frac{1-\alpha}{1-\alpha^*} \right) K(v_2) \left[1 + \frac{1}{v_2} \left(\frac{x}{2\alpha^*} \right)^2 \right]^{-\frac{v_2+1}{2}}, & x > 0 \end{cases} \tag{20}$$

$\alpha \in (0,1)$ is skewness parameter.$v_1 > 0$ and $v_2 > 0$ are left and right tail parameters respectively.

$$K(v) \equiv \left(\Gamma(v+1)/2/\left[\sqrt{\pi v}(v/2)\right]\right)$$

where Γ (.) is gamma function and α^* is:

$$\alpha^* = \alpha\left(v_1\right)/\left[\alpha K\left(v_1\right)+\right](1-\alpha)K\left(v_2\right)$$

Denoting by μ and σ the location (centre) and scale parameters, respectively, the general form of the *AST* density is expressed as $\frac{1}{\sigma}f_{AST}\left(\frac{x-\mu}{\sigma};\alpha,v_1,v_2\right)$.

Note that

$$\left(\frac{\alpha}{\alpha^*}\right)K\left(v_1\right) = \left(\frac{1-\alpha}{1-\alpha^*}\right)K\left(v_2\right) = \left[\alpha K\left(v_1\right)+\right](1-\alpha)K_{EP}\left(v_2\right) \equiv \mathrm{B}$$

Conditional VaR^p_{t+1} and ES^p_{t+1} of *AST* are:

$$\begin{aligned} VaR^p_{t+1} &= -\mu_{t+1} - \sigma_{t+1}VaR_{AST}\left(p|\alpha,v_1,v_2\right) \\ ES^p_{t+1} &= -\mu_{t+1} - \sigma_{t+1}ES_{AST}\left(p|\alpha,v_1,v_2\right) \end{aligned} \tag{21}$$

where

$$\begin{aligned} &VaR_{AST}\left(p|\alpha,v_1,v_2\right) \\ &= 2\alpha^* S^{-1}_{v_1}\left(\frac{min(p,\alpha)}{2\alpha}\right) + 2(1-\alpha^*)S^{-1}_{v_2}\left(\frac{max(p,\alpha)+1-2\alpha}{2(1-\alpha)}\right) \end{aligned}$$

where $S_v(.)$ is the cumulative distribution function of the standard student *t*-distribution with ν degrees of freedom and S^{-1}_v is its inverse.

$$\begin{aligned} &ES_{AST}\left(p|\alpha,v_1,v_2\right) \\ &= -\frac{4B}{p}\left\{\frac{(\alpha^*)^2v_1}{v_1-1}\left(1+\frac{1}{v_1}\left[\frac{min(q-\mu,0)}{2\alpha^*}\right]^2\right)^{\frac{1-v_1}{2}} - \frac{(1-\alpha^*)^2v_2}{v_2-1}\left(1+\frac{1}{v_2}\left[\frac{min(q-\mu,0)}{2\alpha^*}\right]^2\right)^{\frac{1-v_2}{2}}\right\} \end{aligned}$$

where $q = VaR_{AST} \equiv F^{-1}_{AST}$.

2.1.7 Skewed student t-distribution (SST)

By letting $v_2 = v_1 = v$ and $\alpha^* = \alpha$ in *AST* by Zhu and Galbraith (2010), we obtain new parameterization of skewed student *t*-distribution (*SST*):

$$f_{SST}(x|\beta) = \begin{cases} \frac{1}{\sigma}K(v)\left[1+\frac{1}{v}\left(\frac{x-\mu}{2\alpha\sigma}\right)^2\right]^{-\frac{v+1}{2}}, & x \le \mu \\ \frac{1}{\sigma}K(v)\left[1+\frac{1}{v}\left(\frac{x}{2\alpha\sigma}\right)^2\right]^{-\frac{v+1}{2}}, & x > \mu \end{cases} \tag{22}$$

Conditional VaR^p_{t+1} and ES^p_{t+1} of *SST* are:

$$\begin{aligned} VaR_{t+1}^{p} &= -\mu_{t+1} - \sigma_{t+1} VaR_{SST}(p|\alpha, \nu) \\ ES_{t+1}^{p} &= -\mu_{t+1} - \sigma_{t+1} ES_{SST}(p|\alpha, \nu) \end{aligned} \tag{23}$$

where

$$\begin{aligned} &VaR_{SST}(p|\alpha, \nu) \\ &= 2\alpha^{*} S_{\nu}^{-1}\left(\frac{min(p,\alpha)}{2\alpha}\right) + 2(1-\alpha^{*}) S_{\nu}^{-1}\left(\frac{max(p,\alpha)+1-2\alpha}{2(1-\alpha)}\right) \end{aligned}$$

where $S_{\nu}(.)$ is the cumulative distribution function of the standard student t-distribution with ν degrees of freedom and S_{ν}^{-1} is its inverse.

$$ES_{SST}(p|\alpha, \nu) = -\frac{4B}{p}\left\{ \begin{array}{l} \frac{(\alpha^{*})^{2}\nu}{\nu-1}\left(1+\frac{1}{\nu}\left[\frac{min(q-\mu,0)}{2\alpha^{*}}\right]^{2}\right)^{\frac{1-\nu}{2}} \\ -\frac{(1-\alpha^{*})^{2}\nu}{\nu-1}\left(1+\frac{1}{\nu}\left[\frac{min(q-\mu,0)}{2\alpha^{*}}\right]^{2}\right)^{\frac{1-\nu}{2}} \end{array} \right\}$$

where $q = VaR_{SST} \equiv F_{SST}^{-1}$.

2.2 Term structure of risk: VaR and ES calculation over multi-period

To date majority of studies have focused on single day market risk estimation. Currently the most popular method is the square-root rule that is applied over short time horizons. If we consider a simple case of normal distribution with a constant variance σ_{PF}^{2}, per square–root rule, the VaR and ES for returns over the next h days calculated on day t, as:

$$\begin{aligned} VaR_{t+1,h}^{p} &= -\sqrt{h}\sigma_{PF}\Phi_{p}^{-1} = \sqrt{h} VaR_{t+1}^{p} \\ ES_{t+1,h}^{p} &= \sqrt{h}\sigma_{PH}\frac{\phi\left(\Phi_{p}^{-1}\right)}{p} = \sqrt{h} ES_{t+1}^{p} \end{aligned} \tag{24}$$

However, given the dynamic nature of variance, moving from one period ahead to multi period h-days ahead is not straightforward because scaling variance as modelled by GARCH processes is not mean reverting with the returns over the next h days are not normally distributed. This drawback means that Monte Carlo simulation needs to be used to calculate VaR and ES over multi-period horizon. We follow Christoffersen (2012) in simulating the index returns having first estimated the underlying GARCH model parameters. Further details on the simulation methodology can be found in Christoffersen (2012). Based on simulated returns over h-days $\left\{\check{R}_{i,t+1:t+h}\right\}_{i=1}^{MC}$, the VaR and ES over period h is:

$$\begin{aligned} VaR_{t+1:t+h}^{p} &= -Percentile\left\{\left\{\check{R}_{i,t+1:t+h}\right\}_{i=1}^{MC}, 100p\right\} \\ ES_{t+1:t+h}^{p} &= -\frac{1}{p.MC}\sum_{i=1}^{MC}\check{R}_{i,t+1:t+h}.1\left(\check{R}_{i,t+1:t+h} < -VaR_{t+1:t+h}^{p}\right) \end{aligned} \tag{25}$$

where $1(\cdot)$ takes the value 1 if the argument is true and zero otherwise and MC denotes the number of draws.

2.3 Back-testing risk models

2.3.1 Bootstrap test for the expected shortfall

To evaluate ES we first use McNeil and Frey (2000) test for zero unconditional mean. The test focuses on the discrepancy between the observed return and the ES forecast for the periods in which the return exceeds the VaR forecast, the assessment of ES forecasts is not independent of the VaR forecasts. McNeil and Frey (2000) defined residuals as:

$$R_{t+1} = \frac{X_{t+1} - ES_t^q\left(X_{t+1}\right)}{\sigma_{t+1}}$$
$$R_{t+1} = \frac{\mu_{t+1} + \sigma_{t+1}Z_{t+1} - \left(\mu_{t+1} + \sigma_{t+1}ES_t^q(Z)\right)}{\sigma_{t+1}} \tag{26}$$

$$R_{t+1} = Z_{t+1} - E_t\left[Z|Z\rangle z_q\right] \tag{27}$$

According to McNeil and Frey (2000) these residuals are *iid* and conditional on $X_{t+1} > x_q$ or equivalently $Z_{t+1} > z_q$ being the q-quantile of Z. Based on our stock price data and our estimates of expected shortfall, we can construct the corresponding residuals on days when violation occurs. McNeil and Frey (2000) call these residuals exceedance residuals and denote them by:

$$r = \left\{r_{t+1}; for\ t\ such\ that x_{t+1} > \hat{x}_q\right\} \tag{28}$$

where

$$r_{t+1} = \frac{x_{t+1} - ES_q^t\left(X_{t+1}\right)}{\hat{\sigma}_{t+1}}$$

Under the null hypothesis that we estimate μ_{t+1}, σ_{t+1} and the expected shortfall correctly, these residuals should behave like an iid sample from a random variable with mean zero and the alternative hypothesis is that the residuals have a mean greater than zero (McNeil and Frey, 2000).

2.3.2 MAE for back-testing ES

We evaluated the expected shortfall as measure of downside risk based on the mean absolute error defined as

$$MAE_j(q) = \frac{1}{N-1}\sum_{i=1}^{N-1}\left|R_{t+1} - ES_{t+i}(q)\right| \tag{29}$$

where $ES_{t+i}(q)$ is the expected shortfall as measure of downside risk and R_{t+1} are observed returns and N is the number of observations. The model with minimum MAE value is preferred to the other models.

Table 1 Summary descriptive statistics

	S&P500	FTSE100	NASDAQ100	NIKKEI225	DAX30
Mean	0.0002	0.0001	0.0004	− 0.0001	0.0003
Median	0.0001	0.0001	0.0005	0.0000	0.0007
Min	− 0.2283	− 0.0927	− 0.1111	− 0.1211	− 0.0887
S.D.	0.1096	0.0938	0.1720	0.1323	0.1080
Skewness	− 1.0212	− 0.1562	− 0.1083	− 0.3329	− 0.1238
Excess kurtosis	27.1950	5.9081	5.1532	6.1275	4.3434
J–B	38077.34**	6850.30**	5206.39**	7434.82**	3704.07**
ADF-Unit Root	− 23.54**	− 16.80**	− 15.70**	− 16.34**	− 16.01**
L-B(20)	145.55	129.88	132.97	63.02	64.02

**Significance at the 1% confidence level

3 Empirical results

3.1 Data and preliminary analysis

The data for this study comprises of five global stock indices, including S&P500, FTSE100, NASDAQ100 – comprising of non-American and non-financial top 100 companies on the NASDAQ exchange, NIKKEI225 and DAX30. All data is obtained from Datastream. For all the indices, the sample comprises of 18 years of daily observation from 1995 to 2013 with a total of 4698 daily observations. The continuously compounded returns are calculated as the logarithmic difference of daily closing price multiplied by 100.

The summary statistics are presented in Table 1. The value of skewness is negative for all return series, indicating an asymmetry in the distribution of return. A negatively skewed distribution or skewed to the left has a long-left tail. All our data series are characterized by many small gains and a few extreme losses. The kurtosis of our data set is greater than 3 and reflects fat tails. We reject the null hypothesis of the normal distribution as the p value for Jarque–Bera (1980) test is less than 0.05. Jarque–Bera test confirms that all return series have non-normal distributions. The Ljung–Box (1978) Q-statistics reported in Table 1 for both returns and squared returns for all data series also reject the null hypothesis of no autocorrelation through 20-lags at a 5% significance level.

3.2 Parameter estimation of distributions of return

Specifically we estimate the parameters of the following seven models: Standardized *t*-distribution (*ST*), Twin *t*-distribution (*TT)* of Baker and Jackson (2014), Generalized asymmetric distribution (*GAT*) of Baker (2016), Asymmetric exponential power distribution (*AEP*) of Zhu and Zinde-Walsh (2009), Skewed exponential power distribution (*SEP*) and the special case of *AST*, the Skewed Student *t*-distribution, Asymmetric Student *t*- distribution (*AST*) of Zhu and Galbraith (2010). The estimation procedure is as follows:

- Given the specific ith t distribution with parameter $\Theta^{(i)}$for $1 \leq i \leq 7$, we identify the underlying GARCH process using the likelihood method. Then use the estimated

Table 2 Estimated parameters and goodness of fit tests for S&P500 for the period 1995–2013

Models	Estimated parameters						Goodness of fit tests		
							− Log L	AIC	BIC
ST	ν								
	6.8609						− 8686.1	**17378.1**	**17398.3**
	(0.4172)								
TT	ν						− 8720.0	**17442.0**	**17448.7**
	4.9625								
	(0.3334)								
GAT	μ	ϕ	α	r	c	v			
	0.0100	2.5691	1.6891	1.0721	1.3372	6.9217	− 8682.7	**17377.4**	**17417.8**
	(0.4825)	(0.4013)	(0.4361)	(0.1673)	(0.2238)	(0.6194)			
AEP	α	d_1	d_2						
	0.5467	1.4397	1.9396				− 8731.9	17469.9	17490.1
	(0.0063)	(0.0156)	(0.0015)						
SEP	α	d					− 8759.3	17522.6	17536.07
	0.5098	1.6356							
	(0.0042)	(0.0332)							
AST	α	v_1	v_2				− 8884.3	17774.5	17794.7
	0.416	3.9043	8.7875						
	(0.0326)	(4.3594)	(1.2350)						
SST	α	v							
	0.5140	18.032					− 8766.1	17536.3	17549.8
	(0.0054)	(2.285)							

Standard errors are reported in parenthesis. Log L is the maximum value of log likelihood function. AIC, Akaike information criterion; BIC, Bayesian information criterion; *ST,* Student *t*-distribution; *TT,* Twin *t*-distribution; *GAT,* Generalized asymmetric *t*-distribution; *AEP,* Exponential power distribution; *SEP,* Skewed exponential power distribution; *AST,* Asymmetric *t*-distribution; *SST,* Skewed student *t*-distribution. The lowest AIC and BIC are in bold

GARCH process to generate the conditional volatility ($\hat{\sigma}_t$) and the conditional mean ($\hat{\mu}$).

- We obtain standardized residuals $\hat{z}_t$ as the function of $\hat{\mu}_t$ and $\hat{\sigma}_t$.
- We define $F_{(i)}(.)$ as the CDF of i-th candidate t distribution, and $Q(.)$ as the quantile function of z_t. The desired $\Theta^{*}_{(i)}$ should satisfy

$$\Theta^{*}_{(i)} = \underset{\Theta^{*}_{(i)}}{\text{arg inf}} \left| Q\left(\hat{z}_t\right) - F^{-1}_{(i)}\left(\hat{z}_t\right); \Theta_{(i)} \right|, 1 \leq i \leq 7 \tag{30}$$

The parameter estimates are reported in Tables 2, 3, 4, 5 and 6 with the log-likelihood values, AIC and BIC values. The bold values of AIC and BIC criteria in all tables represent top three best models for the specific data set. Per AIC and BIC values in Table 2, the best-fitting models for S&P 500 data are the Generalized asymmetric *t*-distribution (*GAT*), Student *t*-distribution (*ST*) and Twin *t*-distribution (*TT*). All above

Table 3 Estimated parameters and goodness of fit tests for FTSE100 for the period 1995–2013

Models	Estimated parameters						Goodness of fit tests		
							– Log L	AIC	BIC
ST	ν								
	11.554						– 6625.2	**13256.4**	**13275.8**
	(1.741)								
TT	ν								
	6.6250						– 6632.6	13267.3	13273.7
	(0.7118)								
GAT	μ	ϕ	α	r	c	ν	– 6612.2	**13236.4**	**13275.1**
	2.2865	4.4041	2.3052	0.9841	1.6462	13.2794			
	(6.6407)	(7.4269)	(3.8638)	(1.0193)	(3.7233)	(13.7587)			
AEP	α	d_1	d_2						
	0.4851	1.5576	2.1301				– 6624.3	**13254.6**	**13273.9**
	(0.0060)	(0.0494)	(0.0704)						
SEP	α	d					– 6646.5	13297.1	13310.0
	0.5115	1.7938							
	(0.0045)	(0.0399)							
AST	α	v_1	v_2				– 6760.3	13526.7	13546.1
	0.5343	1.9972	1.0011						
	(4.2395)	(4.1308)	(2.4285)						
SST	α	v							
	0.5162	24.713					– 6644.4	13292.8	13305.7
	(0.0059)	(4.322)							

Standard errors are reported in parenthesis. Log L is the maximum value of log likelihood function. AIC, Akaike information criterion; BIC, Bayesian information criterion; *ST,* Student *t*-distribution; *TT,* Twin *t*-distribution; *GAT,* Generalized asymmetric *t*-distribution; *AEP,* Exponential power distribution; *SEP,* Skewed exponential power distribution; *AST,* Asymmetric *t*-distribution; *SST,* Skewed student *t*-distribution. The lowest AIC and BIC are in bold

models have lowest AIC and BIC while, asymmetric *t*-distribution (*AST*) and skewed exponential power distribution (*SEP*) have highest AIC and BIC value respectively.

Examining the values of AIC and BIC in Tables 3, 4, 5 and 6 for the indices FTSE100, NASDAQ100, NIKKEI225 and DAX30, we observe that the generalized asymmetric *t*-distribution (*GAT*),[1,2] Student *t*-distribution (*ST*) and Twin *t*-distribution (*TT*) are the top

[1] The flexibility of *GAT* distribution allows us to set $\alpha = 1$, leading to a 5-parameter distribution that turns out to fit returns data almost identically well as the *AST* distribution. On the other hand, by allowing α to deviate from the unity, we have a more general distribution that fits the data better. As α increases, the fatness of the tails decreases, while the power-law behaviour remains the same. In this study we allow α deviate from the unity to fit the data with fatter tails. *AST* and *GAT* distribution behaviour are compared by fixing α = 1.

[2] According to Baker (2016) the *AST* distribution by Zhu and Galbraith (2010) has discontinuity in the second derivative of the log-likelihood function, as a result the usual regularity conditions for maximum likelihood estimation are not satisfied and makes inference for parameter values difficult. This is a real problematic issue for estimation of standard error because it relies on the second derivative of the log-likelihood. On the other hand, *GAT* does not have the same inferential problems, as the log-likelihood function has no discontinuities in derivatives. When we compare *GAT* with Azzalini (2015) skew-normal distribution it reveals that the derivative of the log-likelihood with respect to the skewness parameter is zero when the parameter is zero (the skew-normal reduces to a normal distribution). *GAT* distribution does not have this

Table 4 Estimated parameters and goodness of fit tests for NASDAQ100 for the period 1995–2013

Models	Estimated parameters						Goodness of fit tests		
							– Log L	AIC	BIC
ST	ν						– 6610.8	**13227.7**	**13247.06**
	9.147								
	(1.173)								
TT	ν						– 6622.2	**13246.4**	**13252.9**
	5.9750								
	(0.5749)								
GAT	μ	ϕ	α	r	c	ν	– 6604.7	**13221.5**	**13260.2**
	0.1769	2.8841	1.1135	1.2002	0.8842	9.5866			
	(0.9549)	(3.4575)	(2.4286)	(0.5010)	(0.7110)	(0.9861)			
AEP	α	d_1	d_2				– 6623.0	13252.0	13271.4
	0.4879	1.5285	1.9742						
	(0.0061)	(0.0496)	(0.0661)						
SEP	α	d					– 6637.9	13279.8	13292.7
	0.5097	1.7201							
	(0.0046)	(0.0394)							
AST	α	v_1	v_2				– 6742.7	13491.5	13510.9
	0.4811	5.0000	4.9701						
	(0.0265)	(5.8300)	(9.0530)						
SST	α	v					– 6638.1	13280.2	13293.1
	0.5178	21.799							
	(0.0059)	(3.594)							

Standard errors are reported in parenthesis. Log L is the maximum value of log likelihood function. AIC, Akaike information criterion; BIC, Bayesian information criterion; *ST,* Student *t*-distribution; *TT,* Twin *t*-distribution; *GAT,* Generalized asymmetric *t*-distribution; *AEP,* Exponential power distribution; *SEP,* Skewed exponential power distribution; *AST,* Asymmetric *t*-distribution; *SST,* Skewed student *t*-distribution. The lowest AIC and BIC are in bold

three models respectively except for FTSE100 where *AEP* is the second best model. When we compare *GAT* and *AST* models, we find that *GAT* significantly outperforms *AST*. *AEP* model as an alternative to *AST* and *GAT* performs better that *AST* but under performs *GAT* . We see that the best fitting models for all our indices data sets are our two new distributions *GAT* and *TT* and Standardized Student *t*- distribution. Overall the *GAT* distribution is the best model, as it has many advantages over Standardized Student *t*-distribution. Standardized *t*-distribution does not support asymmetry. Neither of the Zhu–Zinde-Walsh (2009) Asymmetric exponential power distribution and Zhu and Galbraith (2010) asymmetric

Footnote 2 (continued)

problem (Baker 2016). The parameter r controls the asymmetry, with r = 1 for a symmetric distribution. We can also fit *GAT* distribution by setting α = 1 and r = 1 with only four parameters floated as many skew distributions require only μ, ϕ, v and c parameter so that skewness is modelled purely by having different probability mass in the two tails.

Table 5 Estimated parameters and goodness of fit tests for NIKKEI225 for the period 1995–2013

Models	Estimated parameters						Goodness of fit tests		
							− Log L	AIC	BIC
ST	ν								
	7.6881						− 6580.7	**13167.5**	**13186.8**
	(0.8443)								
TT	ν						− 6593.7	**13189.5**	**13195.9**
	5.4500								
	(0.4582)								
GAT	μ	ϕ	α	r	c	ν	− 6577.9	**13167.8**	**13206.5**
	3.0281	9.5427	1.1226	1.0327	1.3563	18.3375			
	(8.8281)	(7.0586)	(4.7099)	(0.2828)	(1.7572)	(6.4887)			
AEP	α	d_1	d_2				− 6609.6	13225.3	13244.6
	0.4881	1.5004	1.8328						
	(0.0062)	(0.0497)	(0.0622)						
SEP	α	d					− 6618.7	13241.4	13254.3
	0.5051	1.6457							
	(0.0048)	(0.0393)							
AST	α	v_1	v_2				− 6701.4	13408.9	13428.2
	0.5348	7.0107	2.9356						
	(0.6859)	(1.7620)	(8.2456)						
SST	α	v					− 6609.1	13222.2	13235.1
	0.5337	18.648							
	(0.006)	(2.744)							

Standard errors are reported in parenthesis. Log L is the maximum value of log likelihood function. AIC, Akaike information criterion; BIC, Bayesian information criterion; *ST,* Student *t*-distribution; *TT,* Twin *t*-distribution; *GAT,* Generalized asymmetric *t*-distribution; *AEP,* Exponential power distribution; *SEP,* Skewed exponential power distribution; *AST,* Asymmetric *t*-distribution; *SST,* Skewed student *t*-distribution. The lowest AIC and BIC are in bold

t-distribution provide the best fits to the models. Per AIC and BIC, the new Twin *t*-distribution also performs better than Asymmetric *t*- distribution and Exponential power distribution for all data sets. To summarize based on the AIC and BIC criteria *GAT*, *TT* and *ST* provide a better fit than *AST* and *AEP* and their skewed versions.

3.3 One-day ahead expected shortfall back-testing

To test the validity of the different *t*-distributional assumptions on the five different indices, the VaR and ES at different confidence levels ranging from 5% to 0.5% are calculated. As the existing literature has demonstrated VaR has serious drawbacks, we therefore only evaluate different ES models to assess market risk. As stated earlier the competing risk models are *ST*, *TT*, *GAT*, *AEP*, *SEP*, *AST* and *SST*, with the *TT* and *GAT* used for the first time in the literature to calculate market risk as calculated by VaR or ES.

To evaluate the ES forecasts, we first use McNeil and Frey's (2000) bootstrap test. The high *p* values given by this test speak in favour of a model, while low *p* values speak against a model. The results indicate that *AST*, *AEP* and *GAT* have highest *p* values which

Table 6 Estimated parameters and goodness of fit tests for DAX30 for the period 1995–2013

Models	Estimated parameters						Goodness of fit tests		
							– Log L	AIC	BIC
ST	ν								
	9.254						– 6605.1	**13216.2**	**13235.5**
	(1.1715)								
TT	ν						– 6619.7	**13241.5**	**13247.9**
	5.9500								
	(0.5544)								
GAT	μ	ϕ	α	r	c	ν	– 6594.4	**13200.8**	**13239.5**
	1.3257	2.1184	0.8694	0.6905	2.7242	8.9509			
	(2.0776)	(0.3957)	(1.7480)	(0.1939)	(1.4251)	(9.0387)			
AEP	α	d_1	d_2				– 6620.0	13246.0	13265.4
	0.4890	1.5178	2.0030						
	(0.0060)	(0.0497)	(0.0626)						
SEP	α	d					– 6638.3	13280.8	13293.7
	0.5126	1.7396							
	(0.0046)	(0.0392)							
AST	α	v_1	v_2				– 6739.6	13485.2	13504.5
	0.5380	2.9423	7.0104						
	(0.9160)	(9.7721)	(4.9073)						
SST	α	v							
	0.5267	23.074					– 6631.9	13267.9	13280.8
	(0.005)	(4.020)							

Standard errors are reported in parenthesis. Log L is the maximum value of log likelihood function. AIC, Akaike information criterion; BIC, Bayesian information criterion; *ST,* Student *t*-distribution; *TT,* Twin *t*-distribution; *GAT,* Generalized asymmetric *t*-distribution; *AEP,* Exponential power distribution; *SEP,* Skewed exponential power distribution; *AST,* Asymmetric *t*-distribution; *SST,* Skewed student *t*-distribution. The lowest AIC and BIC are in bold

is significantly higher then of 0.01. As indicated by McNeil and Frey's (2000) that an assumption of normality always fails the test with p values in all cases much less than 0.01. In our case p values for all distribution is much higher than of 0.01 in almost all the cases. The results, which are shown in Table 7, and clearly provide insight for the use of asymmetric distributions for the risk analysis purpose. However, this test provides little information into the relative performance of the methods. This motivates the use of an additional approach to evaluating ES forecast accuracy.

We compare ES by using MAE that calculates the difference between the actual and the expected losses when a violation occurs. The small value of calculated mean absolute error and the mean squared errors appear small enough to suggest that the best fitting models are reasonable.

Table 7 contains the performance results for all the models and indices, with each of the panels containing the results for each of the indices across the seven models. From Panel A, we see that for the S&P500, at the 5% level *GAT, TT* and *SST* provide the best fit. At the 2.5% level, again results indicate that *GAT*, *TT* and *SST* models providing the best fit.

However, at the 1% or 0.5% level *SST* model provides the best fit by outperforming the other models. At 1% and 0.5% *SEP* is the second best model.

Panel B contains the results for the FTSE100 index, up to 2.5% confidence level *GAT, ST* and *TT* model outperforms all other models. However, at the 1% and 0.5% level *GAT, TT* are the best performers. Panel C comprises of the NASDAQ back-testing results where we find that at 5% *GAT* provides the best fit, while *SEP* and *AEP* are second and third best models. At 2.5%, 1% and 0.5% again *GAT*, *SEP* and *AEP* are the outperformers, however *SEP* outperforms *GAT* and *AEP* marginally. Moreover, *AST* performs better then *TT*, *SST* and *ST*. In Panel D we present the results for NIKKIE225 and note that our proposed *GAT* model outperforms all other models at 5%. At 2.5% only *SEP* outperforms *GAT*. However, at 1% and 2.5% both *SEP* and *AEP* perform better than of *GAT*. Panel E indicate the results for DAX30 *GAT, TT* and *SEP* are the best performing models at 5% and 2.5%. At 1% *TT, SEP* and *SST* are the best performers. AT 0.5% *SEP, SST* and *AEP* perform better than *GAT* and *TT*.

To summarize our key results:

(i) *GAT* model and *TT* models are in the top three models at 5% and 2.5% confidence level in almost all cases.
(ii) *AST* model have highest values of MAE for almost all datasets and significance levels except NASDAQ in panel C.
(iii) The skewed version of *AST* model (*SST*) is amongst the models with the highest MAE values except S&P500 in panel A, where it is third best model after *GAT* and *TT*.
(iv) *AEP* model as alternative to asymmetric distributions performs better than the *AST*, but *GAT* model clearly outperforms *AEP* in most of the cases.
(v) The skewed version of *AEP* model (*SEP*) performs better than of the skewed version of *AST* model (*SST*) in most of the cases. For NASDAQ, NIKKIE300 and DAX225 it is among the top three models.
(vi) The results of MAE indicate different model ranking for the same confidence level. However, for most of the cases *GAT* remain in the top three models.
(vii) These results give us a strong indication that new parameterization of generalized asymmetric distribution provides valuable improvement in the results. When we compare ES back-testing for two asymmetric *t*-distributions, MAE of *GAT* are significantly lower than that of *AST*. These results indicate strong implication for further research for use of asymmetric *t*-distribution as ES measure.

Based on the ES back-tests conducted, we conclude that the *GAT* model by Baker (2016) outperforms the competing *AST* by Zhu and Galbraith (2010) model by a significant margin. As an alternative to asymmetric *t*-distribution *AEP* model also underperforms *GAT* model.[3]

[3] To further test the robustness of our results, we created subsample for the whole period excluding the three financial crisis period and subsample for each of the financial crisis periods. We found that the performance of the models were independent of the sample period, i.e. *GAT* distribution was overall the best performer regardless of the sample period. However, magnitude of the risk measures VaR and ES decreased when we excluded the crisis periods from our sample and correspondingly they increased during each of the financial crisis periods.

3.4 Multi-period horizon ES back-testing

Table 8 contains the back-testing results across 4 days and 10 days for each of the indices across all of the models. Regarding the results for the 5-day horizon, we find that for the S&P500, *SST* model outperforms all other models up to 1% level. However, at the 0.5% level *GAT* model outperforms all other models. *SST* and *GAT* are among the top three outperformers for all the significance level while *TT* is among top two best model for 2.5% and 1% significance level.

From Panel B, we see that for the FTSE100 index, up to 2.5% level, *SST* is the best performer, followed by *GAT* at 1% and 0.5%. Looking at Panel C and D, we see that for the NASDAQ100 and NIKKIE225 at all significance level *GAT*, *SST* and *TT* are the best performers. At 5% and 2.5% *SST* outperforms *GAT* and *TT*. However, at 1% and 0.5% *GAT* has smallest MAE than of *SST* and *TT*. In panel C the results for DAX30 indicates that *GAT*, *TT* and *AEP* are the top three models for all significance levels.

From Table 8, with regards to the 10-day horizon, results are straight cut, for the S&P500, FTSE100, NIKKEI225 and DAX30, *GAT* provides the best result across all confidence levels with the *TT* providing the best fit for the remaining index NASDAQ100. Our results for the predicted ES for 5-days and 10-days can be summarized as follows:

(i) At the 5-day horizon results are mixed with the *SST* being the best performer up to 1% level in majority of the cases. However, at 0.5% confidence level *GAT* is the best performer. Overall, *GAT* remains in the top three models based on the lowest MAE value.
(ii) When we increase the number of horizons to 10-days, MAE values clearly suggest *GAT* as the best model for almost all data sets.
(iii) Both *AEP* and *SEP* perform very poorly to forecast ES for both 5-days and 10-days horizon at various significance levels.
(iv) AST model has highest MAE value in most of the case for both 5-days and 10-days horizon.
(v) We can infer that results of ES models are not similar across different time horizons. However, the satisfactory predictions of the *GAT* are in accordance with the findings of 1-day ahead ES evaluation. Again, like 1-day ahead *GAT* model out performs *AST* model and give clear implication for the use of *GAT* distribution for risk forecasting.

4 Concluding remarks

The recent crisis has highlighted the weaknesses of VaR as a market measure of risk. This has resulted in the related superior measure ES being given more prominence under Basel III (Basel Committee on Banking and Supervision 2013, 2017). Previous studies have focused on VaR and more specifically on a single day VaR. This study has sought to complement earlier studies by expanding market risk measures to ES over multi-day horizon using seven different models that incorporate the observed empirical characteristics of equity returns as noted by Kellner and Rosch (2016) who recommends that only models which allow for heavy tailed and/or skewness can accurately estimate both VaR and ES.

In this study we make a number of contributions. First, we found that when seven different models based on alternative *t*-distributions were fitted to the standardized

Table 7 Back-testing results for 1-day ahead ES for international indices

p	5%		2.5%		1%		0.5%	
Models	MAE	BOOT	MAE	BOOT	MAE	BOOT	MAE	BOOT
Panel A: S&P500								
ST	0.0268	**1.0000**	0.0318	**1.0000**	0.0387	**0.9990**	0.0443	**0.9990**
TT	**0.0226**	0.4082	**0.0267**	0.4940	**0.0318**	0.4707	0.0392	0.4104
GAT	**0.0209**	**0.9764**	**0.0263**	**0.9772**	0.0341	**0.9810**	**0.0381**	**0.9721**
AEP	0.0291	0.7958	0.0323	0.2823	0.0395	0.0505	0.0437	0.2588
SEP	0.0240	0.0084	0.0275	0.0050	**0.0319**	0.0041	**0.0350**	0.0006
AST	0.0271	**1.0000**	0.0314	**1.0000**	0.0413	**0.9855**	0.0503	**0.9972**
SST	**0.0235**	0.0186	**0.0269**	0.0225	**0.0313**	0.0151	**0.0346**	0.0036
Panel B: FTSE100								
ST	**0.0222**	0.0021	**0.0228**	0.0000	0.0382	0.9958	0.0482	0.0000
TT	**0.0229**	0.1615	**0.0265**	0.7734	**0.0315**	0.5839	**0.0371**	0.0001
GAT	**0.0219**	**0.9799**	**0.0225**	**0.9836**	**0.0301**	**0.9827**	**0.0361**	**0.9809**
AEP	0.0254	0.9857	0.0293	0.6481	0.0342	0.2827	0.0377	0.2416
SEP	0.0254	1.0000	0.0292	0.5912	0.0338	0.1768	0.0371	0.1302
AST	0.0341	**0.9999**	0.0417	**0.9912**	0.0540	**0.9753**	0.0685	**1.0000**
SST								
Panel C: NASDAQ								
ST	0.0425	**0.9999**	0.0497	**1.0000**	0.0594	**0.9995**	0.0671	**0.7128**
TT	0.0374	0.8130	0.0436	0.6915	0.0523	0.4006	0.0603	0.0070
GAT	**0.0261**	**0.9964**	**0.0326**	**0.9759**	**0.0429**	**0.3016**	**0.0443**	**0.1038**
AEP	**0.0276**	0.8934	**0.0312**	0.5983	**0.0358**	0.0134	**0.0393**	0.0174
SEP	**0.0266**	0.5614	**0.0297**	0.3145	**0.0336**	0.0411	**0.0365**	0.0013
AST	0.0369	**1.0000**	0.0467	**0.9968**	0.0541	**0.9284**	0.0638	**0.9126**
SST	0.0399	0.5655	0.0456	0.3511	0.0528	0.0585	0.0581	0.0025
Panel D: NIKKEI225								
ST	0.0358	**1.0000**	0.0423	**1.0000**	0.0512	**0.9995**	0.0583	**0.9867**
TT	0.0308	0.1255	0.0362	0.5216	0.0439	0.4487	0.0531	0.0039
GAT	**0.0253**	**1.0000**	**0.0297**	**0.9210**	**0.0378**	**0.4307**	**0.0402**	**0.2413**
AEP	**0.0264**	0.7185	**0.0303**	0.2174	**0.0354**	0.1622	**0.0390**	0.0274
SEP	**0.0255**	0.1595	**0.0290**	0.2141	**0.0334**	0.1091	**0.0366**	0.0529
AST	0.0390	**0.9987**	0.0466	**0.9151**	0.0545	**0.3904**	0.0633	**0.2836**
SST	0.0340	0.3944	0.0389	0.1068	0.0453	0.1877	0.0500	0.0260
Panel E: DAX30								
ST	0.0362	**0.9999**	0.0426	**0.9999**	0.0515	**0.9886**	0.0586	**0.9663**
TT	**0.0300**	0.7413	**0.0350**	0.7144	**0.0420**	0.6281	0.0531	0.0001
GAT	**0.0301**	**0.9993**	**0.0368**	**0.8573**	0.0454	**0.5553**	0.0543	**0.6219**
AEP	0.0336	0.7116	0.0390	0.2139	0.0456	0.1506	**0.0504**	0.0482
SEP	**0.0319**	0.0676	**0.0364**	0.0856	**0.0420**	0.0480	**0.0459**	0.0441
AST	0.0589	**1.0000**	0.0768	**0.9979**	0.1072	**0.8631**	0.1369	**0.9980**
SST	0.0331	0.1378	0.0377	0.1126	**0.0436**	0.1101	**0.0479**	0.0508

MAE: mean absolute error and BOOT: *p* values from the bootstrap test for ES at 5%, 2.5%, 1% and 0.5% significance level. Student *t*-distribution: *ST*; twin *t*-distribution: *TT*; generalized asymmetric *t*-distribution: *GAT*; exponential power distribution: *AEP*; skewed exponential power distribution: *SEP*; asymmetric *t*-distribution: *AST*; skewed student *t*-distribution: *SST*. The lowest three MAE and highest three BOOT are in bold

Table 8 Back-testing results for 5-days and 10-days ahead ES using MAE for international indices

p	5%		2.5%		1%		0.5%	
Models	5-days	10-days	5-days	10-days	5-days	10-days	5-days	10-days
Panel A: S&P500								
ST	0.4386	**1.4184**	0.3608	**1.2326**	0.2875	**1.0457**	**0.2229**	**0.8713**
TT	0.4798	1.6000	**0.3115**	1.3892	**0.2388**	1.1644	0.2388	0.9543
GAT	**0.4074**	**1.1559**	**0.3214**	**1.0450**	**0.2434**	**1.0178**	**0.1833**	**0.7805**
AEP	0.7028	3.3821	0.5920	3.0459	0.4850	2.7110	0.3919	2.420
SEP	**0.4078**	2.1537	0.3559	2.0103	0.3006	1.8525	0.2499	1.7088
AST	0.4997	2.5772	0.4000	1.6412	0.3065	1.1067	0.2314	0.9054
SST	**0.3336**	**1.2182**	**0.2875**	**1.0899**	**0.2374**	**0.9418**	**0.1895**	**0.7906**
Panel B: FTSE100								
ST	**0.4258**	**2.3477**	**0.3469**	**1.8844**	**0.2720**	**1.4823**	**0.2101**	**1.1696**
TT	0.5809	3.2106	0.4563	2.5578	0.3442	1.9671	0.2596	1.5075
GAT	**0.4181**	**2.1396**	**0.3358**	**1.7178**	**0.2576**	**1.3564**	**0.1934**	**1.0674**
AEP	1.0883	17.110	0.8433	14.124	0.6390	11.3122	0.4893	9.0765
SEP	0.4320	4.4874	0.3736	3.9710	0.3129	3.4674	0.2585	3.0121
AST	0.6559	3.3531	0.5034	2.6126	0.3688	1.5345	0.2679	1.4835
SST	**0.3832**	**2.2210**	**0.3230**	**1.8350**	**0.2603**	**1.4734**	**0.2036**	**1.1718**
Panel C: NASDAQ								
ST	0.5056	3.0715	0.4116	2.5191	0.3234	1.9894	0.2453	1.5736
TT	**0.4493**	**2.3542**	**0.3546**	**1.9630**	**0.2696**	**1.5951**	**0.2036**	**1.2832**
GAT	**0.4924**	**2.4510**	**0.4039**	**2.0485**	**0.3014**	**1.6451**	**0.2230**	**1.3153**
AEP	0.5264	5.7652	0.4423	5.0583	0.3650	4.3713	0.2972	3.7823
SEP	0.5077	5.4814	0.4386	4.8805	0.3679	4.2782	0.3022	3.7496
AST	0.7335	4.0120	0.5699	3.2298	0.4225	2.5249	0.3086	1.9786
SST	**0.4377**	**2.7354**	**0.3693**	**2.2738**	**0.2964**	**1.8404**	**0.2293**	**1.4733**
Panel D: NIKKEI225								
ST	0.5845	2.8188	0.4564	2.3142	0.3494	1.8499	0.2631	1.4663
TT	**0.4808**	**2.3296**	**0.3687**	**1.9159**	**0.2781**	**1.5409**	**0.2096**	**1.2404**
GAT	**0.4788**	**2.3282**	**0.3739**	**1.8985**	**0.2831**	**1.5133**	**0.2118**	**1.2024**
AEP	0.5299	5.4371	0.4474	4.7553	0.3702	4.1006	0.3033	3.5433
SEP	0.5501	5.2936	0.4632	4.6467	0.3799	4.0389	0.3076	3.5203
AST	0.6989	3.5864	0.5458	2.8978	0.4081	2.2801	0.3012	1.7957
SST	**0.4210**	**2.4453**	**0.3574**	**2.0426**	**0.2890**	**1.6623**	**0.2257**	**1.3387**
Panel E: DAX30								
ST	0.7406	**6.9067**	0.5486	**5.4817**	**0.3917**	**4.1715**	**0.2739**	**3.1206**
TT	**0.6306**	**5.4223**	**0.4433**	**4.2996**	**0.3072**	**3.3353**	**0.2121**	**2.5555**
GAT	**0.6473**	**5.4831**	**0.4628**	**4.3249**	**0.3195**	**3.2380**	**0.2170**	**2.4076**
AEP	**0.6673**	14.1015	**0.5414**	12.0997	0.4218	10.1782	0.3262	8.5443
SEP	0.7166	13.8252	0.5662	11.9087	0.4330	10.0599	0.3292	8.4900
AST	0.9186	9.0201	0.6815	7.0409	0.4738	5.2647	0.3221	3.9203
SST	0.7166	9.3688	0.5662	7.8868	0.4330	6.5368	0.3292	5.3564

MAE: mean absolute error test for ES at 5%, 2.5%, 1% and 0.5% significance level. *ST,* student *t*-distribution; *TT,* twin *t*-distribution; *GAT,* generalized asymmetric *t*-distribution; *AEP,* exponential power distribution; *SEP,* skewed exponential power distribution; *AST,* asymmetric *t*-distribution; *SST,* skewed student *t*-distribution. The lowest three MAE are in bold residuals, we found that our two new proposed models Generalized asymmetric *t*-distribution (*GAT*) of Baker (2016) and Double *t*-distribution (*TT*) of Baker and Jackson (2014) provided the best fit, with *GAT* model being overall the best model. Moreover, surprisingly the Standard *t*-distribution outperformed many of the more complex *t*-distributions.

Second, the performance of the ES models are dependent on the market and the confidence level, particularly so at the 1-day and 5-day horizons. This result would indicate that for short horizons, risk managers and regulators should use a variety of models and check the accuracy of each model specific to each index and constantly re-assess the validity of each model. For longer horizons we find that our new proposed models *GAT* outperformed all the models considered in this study. This would indicate that for longer horizons, risk managers should focus on a single model, rather than a number of alternative models.

Third, complex models do not always lead to best fits or back-testing results. For example, in many cases the Standardized *t*-distribution outperforms the more complex Asymmetric exponential power distribution (*AEP*) of Zhu and Zinde-Walsh (2009). These findings are further reinforced by the outperformance of by our simpler *GAT* and *TT* distributions across different horizons, confidence levels and markets.

Finally the backtesting results indicates a wide variation of ES values across different models and indices. Given that the VaR and ES values form the basis of regulatory capital allocation, it is imperative that the most accurate model with the lowest estimated VaR and ES are used by both regulators and managers as the wrong model may mean either capital is not efficiently used or insufficient capital is set aside. In this regard, our *GAT* model provides a reliable alternative to many of the existing models in that it is overall the best performing model across different confidence levels, different horizons and different indices.

References

Abad P, Benito S, Lopez C (2014) A comprehensive review of value at risk methodologies. Span Rev Financ Econ 12:15–32

Angelidis T, Benos A, Degiannakis S (2007) A robust VaR model under different time periods and weighting schemes. Rev Quant Finance Account 28:187–201

Azzalini A (2015) Skew-normal family of distributions. Wiley StatsRef: Statistics Reference Online. Wiley, Hoboken

Baixaali JS, Alvarez S (2006) Evaluating effects of excess kurtosis on VaR estimates: evidence for international stock indices. Rev Quant Finance Account 27:27–46

Baker RD (2016) A new asymmetric generalization of the t-distribution. University of Salford Manchester. https://arxiv.org/abs/1606.05203

Baker RD, Jackson D (2014) Twin *t* distribution. University of Salford Manchester. https://arxiv.org/abs/1408.3237

Basel Committee on Banking Supervision (2013) Basel III: a global regulatory framework for more resilient banks and banking systems. www.bis.org. Accessed 21 Aug 2014

Basel Committee on Banking Supervision (2017) Revisions to the Basel III market risk framework Basel, Switzerland: BIS. www.bis.org. Accessed 10 Feb 2018

Bollerslev T (1986) Generalized autoregressive conditional heteroskedasticity. J Econom 31:307–327
Bollerslev T (1987) A conditional heteroskedastic time series model for speculative prices and rates of return. Rev Econ Stat 69:542–547
Chiou WJP, Lee CL, Lee CF (2009) Variation in stock returns risks: an international comparison. Revi Pac Basin Financ Mark Polic 12(2):245–266
Christoffersen P (2012) Elements of financial risk management, 2nd edn. Academic Press, San Diego, CA
Dowd K, Cotter J, Sorwar G (2008) Spectral risk measures: properties and limitations. J Financ Serv Res 34:61–75
Hansen BE (1994) Autoregressive conditional density estimation. Int Econ Rev 35:705–730
Huang YC, Lin B (2014) Value-at-risk analysis for Taiwan stock index futures: fat tails and conditional asymmetries in return innovations. Rev Quant Finance Account 22:79–95
Jarque CM, Bera AK (1980) Efficient tests for normality, homoscedasticity and serial independence of regression residuals. Econ Lett 6(3):255–259
Kellner R, Rosch D (2016) Quantifying market risk with value-at-risk or expected shortfall?—consequences for capital requirement and model risk. J Econ Dyn Control 68:45–63
Lin CH, Chien CCC, Chen SW (2006) Incorporating the time-varying tail fatness into the historical simulation method for portfolio value-at-risk. Revi Pac Basin Financ Mark Polic 9(2):257–274
Ljung GM, Box GEP (1978) On a measure of a lack of fit in time series models. Biometrika 65(2):297–303
McNeil AJ, Frey R (2000) Estimation of tail-related risk measures for heteroscedastic financial time series: an extreme value approach. J Empir Finance 7:271–300
Mogel B, Auer BR (2018) How accurate are modern value-at-risk estimators derived from extreme value theory? Rev Quant Finance Account 50:79–103
Siu YW (2018) Volatility forecast by volatility index and its use as a risk management tool under Value-at-Risk approach. Revi Pac Basin Financ Mark Polic 21(2):p1850010
Wong WK, Fan C, Zeng Y (2012) Capturing tail risk beyond VaR. Revi Pac Basin Financ Mark Polic 15(3):1–25
Zhu D, Galbraith JW (2010) A generalized asymmetric student *t*-distribution with application to financial econometrics. J Econom 157:297–305
Zhu D, Galbraith JW (2011) Modelling and forecasting expected shortfall with a generalized asymmetric student *t* and asymmetric exponential power distribution. J Empir Finance 18:765–778
Zhu D, Zinde-Walsh V (2009) Properties and estimation of asymmetric exponential power distribution. J Econom 148:86–99