

RESEARCH ARTICLE

Refining value-at-risk estimates using a Bayesian Markov-switching GJR-GARCH copula-EVT model

Marius Galabe Sampid, Haslifah M. Hasim*, Hongsheng Dai

Department of Mathematical Sciences, University of Essex, Colchester, United Kingdom

* hhashim@essex.ac.uk



Abstract

In this paper, we propose a model for forecasting Value-at-Risk (VaR) using a Bayesian Markov-switching GJR-GARCH(1,1) model with skewed Student's-*t* innovation, copula functions and extreme value theory. A Bayesian Markov-switching GJR-GARCH(1,1) model that identifies non-constant volatility over time and allows the GARCH parameters to vary over time following a Markov process, is combined with copula functions and EVT to formulate the Bayesian Markov-switching GJR-GARCH(1,1) copula-EVT VaR model, which is then used to forecast the level of risk on financial asset returns. We further propose a new method for threshold selection in EVT analysis, which we term the *hybrid* method. Empirical and back-testing results show that the proposed VaR models capture VaR reasonably well in periods of calm and in periods of crisis.

OPEN ACCESS

Citation: Sampid MG, Hasim HM, Dai H (2018) Refining value-at-risk estimates using a Bayesian Markov-switching GJR-GARCH copula-EVT model. PLoS ONE 13(6): e0198753. <https://doi.org/10.1371/journal.pone.0198753>

Editor: Cathy W.S. Chen, Feng Chia University, TAIWAN

Received: September 8, 2017

Accepted: May 24, 2018

Published: June 22, 2018

Copyright: © 2018 Sampid et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: The data employed in this analysis consist of 2870 daily observations of stock prices from 31 December 2004 to 31 December 2015. The stocks are actively traded on the London Stock Exchange and belong to the banking sector; they are: HSBC Holdings, Lloyds Banking Group, Barclays Plc., Royal Bank of Scotland Group, and Standard Chartered Plc. We refer to these banks as Bank 1, Bank 2, Bank 3, Bank 4, and Bank 5 in the manuscript. All relevant data are available from datastream; a database for financial and economic research data from Thomas Reuters, which is a third-party database, and not

Introduction

In recent decades, Value-at-Risk (VaR) has become a key tool for measuring market risk; it provides risk managers with a quantitative measure of the downside risk of a firm or investment portfolio during a given time frame. VaR attempts to summarise the total risk in a portfolio of asset or exposures to risk factors in a single number over a target horizon.

There are several methods to estimate VaR; the most commonly used by financial institutions are the variance-covariance, historical simulation and Monte Carlo simulation methods (see [1–3] and the references therein). Historical simulation relies on actual data and is based on the assumption that history will repeat itself; the VaR is estimated by running hypothetical portfolios from historical data [4]. The variance-covariance and Monte Carlo simulation methods assume that asset returns are independent and identically distributed, a major weakness in these VaR models.

Traditional VaR models assume asset returns in financial markets to be normally distributed; thus, changes in asset prices are independent of each other and exhibit constant volatility over time. This is not the case in real life i.e., financial asset returns are leptokurtic and

owned by the authors (<http://financial.thomsonreuters.com/en/products/tools-applications/trading-investment-tools/datastream-macroeconomic-analysis.html>). Other researchers will be able to access these data in the same manner as the authors.

Funding: The author(s) received no specific funding for this work.

Competing interests: The authors have declared that no competing interests exist.

heavy tailed with non-constant volatility [5, 6]. The normality assumption leads to inaccurate estimates in the tails of the distribution and hence of the probability of extreme events, which leads to underestimation of the likelihood of extreme tail losses. This is because the normal distribution has light tails, and VaR attempts to capture the behaviour of the portfolio return in the left tail. A model based on the normal distribution underestimates the frequency of outliers and hence the true VaR [2]. Additionally, the normality assumption implies volatility is constant over time, and recent price changes, which are based on current market information, will be assigned weights in equal proportion to older ones. If the dependence characteristics of the extreme realisations differ from all others in the sample, the consequences might be dire [7].

Non-normality for univariate models is associated with the dependence (i.e., correlation) structure between the asset returns. For multivariate models, non-normality is associated with the joint probability of the univariate models' marginal probabilities, i.e., the joint probability of large market movements, known as *tail dependence*. Because of the complexity of multivariate distributions, the VaR estimation of a portfolio of assets can be quite difficult. To avoid the normality assumption, extreme value theory (EVT) is often used to model the tail behaviour of asset returns. However, EVT also assumes extreme events to be independent and identically distributed, which might not hold in periods of severe crisis [8]. [9] suggests applying EVT to the noise variables of the return series, which are normally distributed, to obtain the q^{th} quantile used to estimate conditional, robust VaR estimates. By doing so, the problem of volatility clustering and other related effects, such as excess kurtosis, is accounted for. This approach was further investigated by [4]; they combined a GARCH(1,1) model as the underlying volatility model with EVT to estimate the VaR of the Tunisian stock market index and showed that the GARCH-EVT-based VaR approach appears to be more effective and realistic than traditional VaR estimation methods.

A study by [10] have shown that volatility predictions following econometric models that ignore regime changes and time varying parameters have several drawbacks. For example, they may fail to capture the dynamics of fluctuations in the time series data. Ignoring regime changes and time varying parameters in high-volatility periods causes significant upwards bias in estimating the GARCH parameters, which impairs volatility forecasts [11]. The Markov-switching GARCH model, first developed by [12] and later improved by [11, 13], helps address the issues since it allows the parameters of GARCH models to vary over time according to a latent discrete Markov process, which leads to volatility forecasts that can rapidly adapt to variations [14].

The problem of dependence can also be improved with the help of copula theory, which enables the construction of flexible multivariate distributions with different margins and dependence structures. This allows the joint distribution of the portfolio to be free from assumptions of normality and linear correlation. [15, 16] have demonstrated that VaR estimates obtained by combining GARCH models, EVT and copula functions are more accurate than those obtained using traditional VaR estimation methods or methods that combined copulas with conventionally employed empirical distributions.

In this paper, we combine the Bayesian Markov-switching GJR-GARCH(1,1) model with skewed Student's- t distribution, copula functions and the peaks over threshold (POT) method of EVT to estimate VaR in selected banks in the United Kingdom (UK) using actively traded stocks on the London Stock Exchange.

Methodology

Markov-switching GJR-GARCH model

Let r_t represent a time series, then a general Markov-switching GARCH specification can be represented as

$$r_t | (\Delta_t = k, \Omega_{t-1}) \sim D(0, h_{k,t}, \Theta_k), \tag{1}$$

$$r_t = \epsilon_t \left(h_{\Delta_t,t}^{\frac{1}{2}} \right), \tag{2}$$

where Δ_t is a Markov chain (a stochastic variable) defined on the parameter space $S = \{1, \dots, K\}$ that symbolises the model, $D(0, h_{k,t}, \Theta_k)$ is a continuous distribution with zero mean and conditional variance $h_{k,t}$, ϵ_t is the distribution of the noise variables, which assumes a skewed Student's- t distribution, Ω_{t-1} is the information set observed up to time $t - 1$, and Θ_k is a vector of the shape parameters.

We define a $K \times K$ transition probability matrix \mathbf{P} , with distinctive elements

$$p_{ij} = \mathbf{P}[\Delta_t = j | \Delta_{t-1} = i], \quad \forall i, j \in \{1, \dots, K\}, \quad 0 < p_{ij} < 1, \quad \sum_{j=1}^K p_{ij} = 1, \tag{3}$$

where p_{ij} is the probability of transition from state $\Delta_{t-1} = i$ to state $\Delta_t = j$. k represents each regime in the Markov chain. The conditional variance, $h_{k,t}$, for $k = 1, \dots, K$ are assumed to follow K -separate GARCH-type processes which evolve in parallel [11, 14]. The Markov switching GARCH models use a stochastic process to define the unknown states [17].

The reliability of a good VaR model depends on the type of volatility model which it incorporates. As discussed above, most financial asset returns are not independently and identically distributed; they exhibit fat tails, leverage effects, and volatility is not constant over time. Volatility reacts differently with large negative returns as compared to positive returns reflecting leverage effects [2]. GARCH models often fail to capture these movements. A good volatility estimator must be able to capture the true behaviour of risk factor returns, it should be easy to implement for a wide range of risk factors, and finally, it should be possible to extend the approach to portfolios with a number of different risk factors [3]. It is well known that traditional GARCH models cannot capture the asymmetric response of volatility. Several other extensions of GARCH models have since been developed as possible solutions to these drawbacks. The most common of these are the exponential generalised ARCH (EGARCH) model [18], the threshold GARCH (TGARCH) model [19], and the GJR-GARCH model [20]. The only significant, albeit minor, difference between TGARCH and GJR-GARCH models is that TGARCH uses standard deviation instead of variance in its specifications [21]. We employ the Markov-switching GARCH model of [11] to capture the differences in the variance dynamics of high and low volatility periods [14], and use the GJR-GARCH(1,1) model to capture the asymmetry response in the conditional volatility process, hence the Markov-switching GJR-GARCH(1,1) model (MS-GJR-GARCH(1,1)).

The conditional variance of a MS-GJR-GARCH model is defined as

$$h_{k,t} = \alpha_{0,k} + (\alpha_{1,k} + \alpha_{2,k} \mathbb{I}_{\{r_{t-1} < 0\}}) r_{t-1}^2 + \beta_k h_{k,t-1}, \quad k = 1, \dots, K, \tag{4}$$

where $\mathbb{I}_{\{\cdot\}}$ is an indicator function introduced to capture the leverage effect such that

$$\mathbb{I}_{\{r_{t-1} < 0\}} = \begin{cases} 1, & \text{if } r_{t-1} < 0, \\ 0, & \text{if } r_{t-1} \geq 0. \end{cases} \tag{5}$$

$\alpha_{2,k}$ controls the degree of asymmetry in the conditional volatility to the past shock in regime k [14]. Thus, $\alpha_{2,k} > 0$ indicates the presence of leverage effect which implies previous negative returns have higher influence on the volatility. The constraints $\alpha_{0,k} > 0$, $\alpha_{1,k} + \alpha_{2,k} \geq 0$ and $\beta_k \geq 0$ ensures a positive variance while covariance stationary is achieved by ensuring that

$$\alpha_{1,k} + \alpha_{2,k} E[\epsilon_{k,t}^2 \mathbb{I}_{\{\epsilon_{k,t} < 0\}}] + \beta_k < 1, \tag{6}$$

where $\mathbb{I}_{\{\cdot\}} = 1$ if the condition holds and 0 otherwise. Note that $E[\epsilon_{k,t}^2 \mathbb{I}_{\{\epsilon_{k,t} < 0\}}] = \frac{1}{2}$ when ϵ_k is symmetrically distributed.

For the conditional distribution of r_t in each regime of the Markov chain, we employ a skew and fat tail error probability distribution; the skewed Student's- t distribution. We use the skewed Student's- t distribution because it is able to account for the excess kurtosis in the conditional distribution that is common with financial time series processes [22]. Moreover, recent studies by [23, 24] have shown that skewed Student's- t errors distribution is a good choice, when compared to a range of existing alternatives. The probability density function (PDF) of a Student's- t distribution is defined as

$$f_s(\epsilon, \nu) = \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{(\nu-2)\pi}\Gamma(\frac{\nu}{2})} \left(1 + \frac{\epsilon^2}{\nu-2}\right)^{-\frac{\nu+1}{2}}, \quad \epsilon \in \mathbb{R}, \tag{7}$$

where the constraint on the degrees of freedom parameter $\nu > 2$ is imposed to guarantee that the second order moment exist, and $\Gamma(\cdot)$ is the Gamma function. Skewness is introduced by an additional parameter $\gamma_k > 0$ as defined in [25]; that is

$$p(\epsilon_k | \nu, \gamma_k) = \frac{2}{\gamma_k + \frac{1}{\gamma_k}} \left\{ f_s\left(\frac{\epsilon_k}{\gamma_k}\right) \mathbb{I}_{[0,\infty)}(\epsilon_k) + f_s(\gamma_k \epsilon_k) \mathbb{I}_{(-\infty,0)}(\epsilon_k) \right\}. \tag{8}$$

When $\gamma_k \neq 1$, the posterior distribution, $p(\epsilon_k | \nu, \gamma_k)$ loses symmetry (see [14, 25, 26] for more details on skewed Student's- t probability distribution).

We use Bayesian statistics to estimate the posterior distribution of the variance equation because the Bayesian estimation method provides reliable results even for finite samples. Moreover, it is usually straightforward when using the Bayesian estimation method, to obtain the posterior distribution of any non-linear function of the model parameter. By comparison, when using the classical maximum likelihood method, it is not easy to perform inferences on non-linear functions of the model parameters, while the convergence rate is slow and presents limitations when the residuals are heavy tailed. The constraints on the GARCH parameters to guarantee a positive variance can be incorporated via priors whereas the classical maximum likelihood method may impede some optimisation procedures [27, 28].

We define a vector of the risk factor returns as $\mathbf{r} = (r_1, \dots, r_T)'$, $\theta_k = (\alpha_{0,k}, \alpha_{1,k}, \alpha_{2,k}, \beta_k, P)'$, and a vector of the model parameters as $\Lambda = (\theta_1, \Theta_1, \dots, \theta_K, \Theta_K)$; $\Theta_K = (\nu_K, \gamma_K)$. Then, from Bayes theorem and prior distribution of the model parameters $p(\Lambda)$, we have

$$p_{ij} = \Pr[\Delta_t = j | \Delta_{t-1} = i] = \frac{f(r_t | \Delta_t = j, \Omega_{t-1}; \Lambda) \Pr(\Delta_t = j | \Omega_{t-1})}{\sum_{i=1}^k f(r_t | \Delta_t = i, \Omega_{t-1}; \Lambda) \Pr(\Delta_t = i | \Omega_{t-1})}, \tag{9}$$

where $f(r_t | \Delta_t = j, \Omega_{t-1}; \Lambda)$ is the conditional probability density of r_t at time t restrictive on Ω_{t-1} and regime j . Therefore we have

$$f(r_t | \Lambda, \Omega_{t-1}) = \sum_{i=1}^k \sum_{j=1}^k \Pr[\Delta_t = j | \Delta_{t-1} = i] f_D(r_t | \Delta_t = j, \Omega_{t-1}; \Lambda) \tag{10}$$

and a likelihood function

$$L(\Lambda|\mathbf{r}) = \prod_{t=1}^T f(r_t|\Lambda, \Omega_{t-1}). \tag{11}$$

The Metropolis Hasting (MH) algorithm of Markov Chain Monte Carlo (MCMC) is then employed to estimate the parameter values of the posterior distribution. As discussed in [22], because of the recursive nature of the variance equation, the prior density $p(\Lambda)$ and posterior density $p(\mathbf{r}|\Lambda)$ do not belong to the same distributional family and, consequently, cannot be expressed in close form. The MH algorithm allows draws to be generated from any density and whose normalising constant is unknown.

In the MH algorithm, Λ is a random variable with Markov chains generated as $(\Lambda^{[0]}, \dots, \Lambda^{[j]}, \dots)$ in a parameter space. As the number of realised chains reaches infinity, $p(\mathbf{r}|\Lambda)$ tends to a normalised probability distribution with a random variable $(\Lambda^{[j]})$ [29]. The chain converges to its stationary distribution and the optimal mean values of the posterior distribution parameters are realised. [22] summarises the MH algorithm as follows: (i) Initialise the iteration counter to $j = 1$ and set the initial value $\Lambda^{[0]}$. (ii) Move the chain to a new value Λ^* generated from a proposal density $q(\cdot|\Lambda^{[j-1]})$. (iii) Evaluate the acceptance probability of the move from $\Lambda^{[j-1]}$ to $\Lambda^{[*]}$ given by

$$\min \left\{ \frac{p(\Lambda^*|\mathbf{r})}{p(\Lambda^{[j-1]}|\mathbf{r})} \frac{q(\Lambda^{[j-1]}|\Lambda^*)}{q(\Lambda^*|\Lambda^{[j-1]})}, 1 \right\}.$$

If the move is accepted, set $\Lambda^{[j]} = \Lambda^*$; if not, set $\Lambda^{[j]} = \Lambda^{[j-1]}$ (i.e., the chain does not move). If chosen from a symmetric proposal density, i.e., $q(\Lambda^{[j]}|\Lambda^*) = q(\Lambda^*|\Lambda^{[j]})$, then the acceptance probability reduces to

$$\min \left\{ \frac{p(\Lambda^*|\mathbf{r})}{p(\Lambda^{[j]}|\mathbf{r})}, 1 \right\}.$$

(iv) Finally, change the counter from j to $j + 1$ and go back to step (ii) until convergence is reached. More details on MH algorithms can be found in [30–33].

Copula theory

Copula theory enables the construction of a flexible multivariate distribution with varying margins and dependence structures; it is free from assumptions of normality or linear correlation. In addition, copulas can easily capture the tail dependence of asset returns, i.e., the joint probability of large market movements.

Copula theory was first developed by [34] to describe the dependence structure between random variables. It was later introduced to the finance literature by [35, 36]. Consequently, [37] introduced the application of copula theory to financial asset returns, and [38] expanded the framework of copula theory with respect to the time-varying nature of financial dependence schemes. Copula theory has also been used in risk management to measure the VaR of portfolios, including both unconditional [39–41] and conditional distributions [42–44].

In multivariate settings, we use the following version of Sklar’s theorem as given by [41] for the purpose of VaR estimation:

Sklar’s theorem: Consider an n -dimensional joint distributional function $F(x)$, with uniform margins $F_1(x_1), \dots, F_n(x_n)$; $x = (x_1, \dots, x_n)$, with $-\infty \leq x_i \leq \infty$, then there exists a copula

$C: [0, 1]^n \rightarrow [0, 1]$ such that

$$F(x_1, \dots, x_n) = C(F_1(x_1), \dots, F_n(x_n)), \tag{12}$$

determined under absolute continuous margins as

$$C(u_1, \dots, u_n) = F(F_1^{-1}(u_1), \dots, F_n^{-1}(u_n)), \tag{13}$$

otherwise, C is uniquely determined on the range $R(F_1) \times \dots \times R(F_n)$. Equally, if C is a copula and F_1, \dots, F_n are univariate distribution functions, then Eq (12) is a joint distribution function with margins F_1, \dots, F_n [45].

The copula $C(u_1, \dots, u_n)$ has density $c(u_1, \dots, u_n)$ associated to it and defined as

$$c(u_1, \dots, u_n) = \frac{\partial_n C(u_1, \dots, u_n)}{\partial u_1, \dots, \partial u_n} \tag{14}$$

and is related to the density function F for continuous random variables denoted as f , by the canonical copula representation [16]

$$f(x_1, \dots, x_n) = c(F_1(x_1), \dots, F_n(x_n)) \prod_{i=1}^n f_i(x_i), \tag{15}$$

where f_i are the marginal densities that can be different from each other [41, 43, 45, 46].

[16, 47] discuss two commonly used families of copulas in financial applications: the elliptical and the Archimedean copulas.

Elliptical copulas are derived from the elliptical distribution by applying Sklar’s theorem. The most common are the Gaussian and the Student’s- t copulas, which are symmetric. Their dependence structure is determined by a standardised correlation or dispersion matrix because of the invariant property of copulas. Consider a symmetric positive definite matrix ρ with $diag(\rho) = (1, 1, \dots, 1)^T$; we can represent the multivariate Gaussian copula (MGC) as

$$C_\rho^{G_a} = \Phi_\rho(\Phi^{-1}(u_1), \dots, \Phi^{-1}(u_n)), \tag{16}$$

where Φ_ρ is the standardised multivariate normal distribution and Φ^{-1} is the inverse standard univariate normal distribution function of u with correlation matrix ρ . If the margins are normal, then the Gaussian copula will generate the standard Gaussian joint distribution function with density function

$$c_\rho^{G_a}(u_1, u_2, \dots, u_n) = \frac{1}{|\rho|^{1/2}} \exp\left(-\frac{1}{2} \zeta'(\rho^{-1} - \mathbf{I})\zeta\right), \tag{17}$$

where $\zeta = (\Phi^{-1}(u_1), \dots, \Phi^{-1}(u_n))'$ and \mathbf{I} is the identity matrix.

On the other hand, the multivariate Student’s- t copula (MTC) has the form

$$T_{\rho, \nu}(u_1, \dots, u_n) = t_{\rho, \nu}(t_\nu^{-1}(u_1), \dots, t_\nu^{-1}(u_n)) \tag{18}$$

with density function

$$c_{\rho, \nu}(u_1, \dots, u_n) = |\rho|^{-\frac{1}{2}} \frac{\Gamma(\frac{\nu+n}{2})}{\Gamma(\frac{\nu}{2})} \left(\frac{\Gamma(\frac{\nu}{2})}{\Gamma(\frac{\nu+1}{2})}\right)^n \frac{(1 + \frac{1}{\nu} \zeta' \rho^{-1} \zeta)^{-\frac{\nu+n}{2}}}{\prod_{j=1}^n (1 + \frac{z_j^2}{\nu})^{-\frac{\nu+1}{2}}}, \tag{19}$$

where $t_{\rho, \nu}$ is the standardised Student’s- t distribution with correlation matrix ρ and ν degrees of freedom.

Archimedean copulas are useful in risk management analysis because they capture asymmetric tail dependencies between financial asset returns. The most common are the Gumbel [48], Clayton [49] and Frank [50] copulas [51]. These copulas are built via a generator as

$$C(u_1, \dots, u_n) = \varphi^{-1}(\varphi(u_1) + \dots + \varphi(u_n)) \tag{20}$$

with density function

$$c(u_1, \dots, u_n) = \varphi^{-1}(\varphi(u_1) + \dots + \varphi(u_n)) \prod_{i=1}^n \varphi'(u_i), \tag{21}$$

where φ is the copula generator and φ^{-1} is completely monotonic on $[0, \infty]$. That is, φ must be infinitely differentiable with derivatives of ascending order and alternative sign such that $\varphi^{-1}(0) = 1$ and $\lim_{x \rightarrow +\infty} \varphi(x) = 0$ [47]. Thus, $\varphi'(u) < 0$ (i.e., φ is strictly decreasing) and $\varphi''(u) > 0$ (i.e., φ is strictly convex).

The Gumbel copula captures upper tail dependence, is limited to positive dependence, and has generator function $\varphi(u) = (-\ln(u))^\alpha$ and generator inverse $\varphi^{-1}(x) = \exp(-x^{\frac{1}{\alpha}})$. This will generate a Gumbel n -copula represented by

$$C(u_1, \dots, u_n) = \exp \left\{ - \left[\sum_{i=1}^n (-\ln u_i)^\alpha \right]^{\frac{1}{\alpha}} \right\} \quad \alpha > 1. \tag{22}$$

The generator function for the Clayton copula is given by $\varphi(u) = u^{-\alpha} - 1$ and generator inverse $\varphi^{-1}(x) = (x + 1)^{-\frac{1}{\alpha}}$, which yields a Clayton n -copula represented by

$$C(u_1, \dots, u_n) = \left[\sum_{i=1}^n u_i^{-\alpha} - n + 1 \right]^{-\frac{1}{\alpha}} \quad \alpha > 0. \tag{23}$$

Frank copula has generator function $\varphi(u) = \ln \left(\frac{\exp(-\alpha u) - 1}{\exp(-\alpha) - 1} \right)$ and generator inverse $\varphi^{-1}(x) = -\frac{1}{\alpha} \ln(1 + e^x(e^{-\alpha} - 1))$, which will result in a Frank n -copula represented by

$$C(u_1, \dots, u_n) = -\frac{1}{\alpha} \ln \left\{ 1 + \frac{\prod_{i=1}^n (e^{-\alpha u_i} - 1)}{(e^{-\alpha} - 1)^{n-1}} \right\} \quad \alpha > 0. \tag{24}$$

We follow [52] and employ Gaussian, Student's- t , Gumbel, Frank and Clayton copulas in this study.

Modelling dependence

The traditional way to measure the relationship between markets and risk factors is to look at their linear correlations, which depend both on the marginal and joint distributions of the risk factors. If there is a non-linear relationship (i.e., in the case of non-normality) the results might be misleading [47]. In this situation, non-parametric invariant measures that are not dependent on marginal probability distributions such as Kendall's τ or Spearman's ρ are more appropriate. Copulas measure a form of dependence between pairs of risk factors (i.e., asset returns) known as concordance using these invariant measures.

Two observations (x_i, y_i) and (x_j, y_j) from a vector (X, Y) of continuous random variables are concordant if $(x_i - x_j)(y_i - y_j) > 0$ and discordant if $(x_i - x_j)(y_i - y_j) < 0$. Large values of X

are paired with large values of Y and small values of X are paired with small values of Y as the proportion of concordant pairs in the sample increases. On the other hand, the proportion of concordant pairs decreases as large values of X are paired with small values of Y and small values of X are paired with large values of Y [53].

Consider n paired continuous observations (x_i, y_i) ranked from smallest to largest, with the smallest ranked 1, the second smallest ranked 2, and so on. Then, Kendall's τ is defined as the sum of the number of concordant pairs minus the sum of the number of discordant pairs divided by the total number of pairs, i.e., the probability of concordance minus the probability of discordance:

$$\tau_{X,Y} = \Pr[(x_i - x_j)(y_i - y_j) > 0] - \Pr[(x_i - x_j)(y_i - y_j) < 0] = \frac{C - D}{C + D}, \tag{25}$$

where C is the number of concordant pairs below a particular rank that are larger in value than that particular rank, and D is the number of discordant pairs below a particular rank that are smaller in value than that particular rank.

Spearman's ρ , on the other hand, is defined as the probability of concordance minus the probability of discordance of the pair of vectors (x_1, y_1) and (x_2, y_3) with the same margins. That is,

$$\rho_{X,Y} = 3(\Pr[(x_1 - x_2)(y_1 - y_3) > 0] - \Pr[(x_1 - x_2)(y_1 - y_3) < 0]).$$

The joint distribution function of (x_1, y_1) is $H(x, y)$, while the joint distribution function of (x_2, y_3) is $F(x)G(y)$ because x_2 and y_3 are independent [54]. Alternatively,

$$\rho_{X,Y} = 1 - \frac{6 \sum_{i=1}^n d_i^2}{n(n^2 - 1)},$$

where d is the difference between the ranked samples.

A study by [54] has shown that Kendall's τ and Spearman's ρ depend on the vectors (x_1, y_1) , (x_2, y_2) and $(x_1, y_1), (x_2, y_3)$, respectively, through their copulas C , and that the following relationship holds:

$$\tau_{X,Y} = 4 \int_0^1 \int_0^1 C(u, v) dC(u, v) - 1$$

and

$$\rho_{X,Y} = 12 \int_0^1 \int_0^1 C(u, v) dudv - 3.$$

Extreme value theory

EVT is a statistical approach for estimating extreme events with low frequency but high severity. This technique is widely used in financial risk management since empirical evidence from various studies [5, 6] show that in the majority of cases, financial asset return distributions are heavy-tailed, especially in times of financial instability.

There are two fundamental approaches for modeling extreme events with low frequency but high severity: the block maxima method and the POT method. The POT method is a commonly used method to model extreme events in financial time series data. On the other hand, the block maxima method is not commonly used for statistical inference on financial time series data for a few reasons: (i) The method does not make sufficient use of data as it uses only the sub-period maxima, (ii) the choice of sub-period length is not clearly defined, and (iii) the

method is unconditional and does not take into account the effects of other explanatory variables [55]. In this paper we use the POT method based on the generalised Pareto distribution (GPD). The POT method focuses on modeling the exceedances of the losses above a certain threshold ϑ and the time of occurrence. The threshold is selected such that there are enough data points to carry out a meaningful statistical analysis. Techniques for selecting the proper threshold are discussed below.

Let $\{x_i\}_{i=1}^T$ represent the loss variables of an asset return, then as $T \rightarrow \infty$, $\{x_i\}_{i=1}^T$ is assumed to be independent and identically distributed, and $(x - \mu)/\sigma$ follows a generalised extreme value (GEV) distribution:

$$F_{\xi, \mu, \sigma}(x) = \begin{cases} \exp[-(1 + \xi x)^{-1/\xi}] & \text{for } \xi \neq 0, \\ \exp[-e^{-x}] & \text{for } \xi = 0, \end{cases} \tag{26}$$

where ξ is the shape parameter and $1/\xi$ is the tail index of the GEV distribution. $x < -1/\xi$ if $\xi < 0$ and $x > -1/\xi$ if $\xi > 0$. Also, let the conditional distribution of the excesses over the threshold, i.e., $x_i - \vartheta = y | x_i > \vartheta$, then

$$\Pr(x - \vartheta \leq y | x > \vartheta) = \frac{\Pr(\vartheta \leq x \leq y + \vartheta)}{\Pr(x > \vartheta)} = \frac{\Pr(x \leq y + \vartheta) - \Pr(x \leq \vartheta)}{1 - \Pr(x \leq \vartheta)} \tag{27}$$

$$= \frac{F(y + \vartheta) - F(\vartheta)}{1 - F(\vartheta)} = F_{\vartheta}(y). \tag{28}$$

Again, as $T \rightarrow \infty$, $(y + \vartheta - \mu)/\sigma$ follows a GEV distribution; see Eq (26). Therefore,

$$\begin{aligned} \Pr(x - \vartheta \leq y | x > \vartheta) &= \frac{F(y + \vartheta) - F(\vartheta)}{1 - F(\vartheta)} \\ &= \frac{\exp\left[-\left(1 + \frac{\xi(y + \vartheta - \mu)}{\sigma}\right)^{-1/\xi}\right] - \exp\left[-\left(1 + \frac{\xi(\vartheta - \mu)}{\sigma}\right)^{-1/\xi}\right]}{1 - \exp\left[-\left(1 + \frac{\xi(\vartheta - \mu)}{\sigma}\right)^{-1/\xi}\right]} \\ &\approx 1 - \left(1 + \frac{\xi y}{\sigma + \xi(\vartheta - \mu)}\right)^{-1/\xi}, \end{aligned} \tag{29}$$

where $y > 0$ and $\sigma + \xi(\vartheta - \mu) > 0$. Let $\psi(\vartheta) = \sigma + \xi(\vartheta - \mu)$, then as $\vartheta \rightarrow \infty$, Eq (29) is approximated by the GPD

$$G_{\xi, \psi(\vartheta)}(y) = \begin{cases} 1 - \left[1 + \frac{\xi y}{\psi(\vartheta)}\right]^{-1/\xi} & \text{for } \xi \neq 0, \\ 1 - \exp[-y/\psi(\vartheta)] & \text{for } \xi = 0, \end{cases} \tag{30}$$

with shape parameter ξ and scale parameter $\psi(\vartheta)$, where $\psi(\vartheta) > 0$, $y \in [0, x - \vartheta]$ when $\xi \geq 0$, and $y \in \left[0, -\frac{\psi(\vartheta)}{\xi}\right]$ when $\xi < 0$. If $\xi = 0$, then Eq (30) becomes an exponential distribution with parameter $1/\sigma$ ([55]). Let $y = x - \vartheta$, then Eq (28) can be written as

$$\frac{F(y + \vartheta) - F(\vartheta)}{1 - F(\vartheta)} = \frac{F(x) - F(\vartheta)}{1 - F(\vartheta)} \approx G_{\xi, \psi(\vartheta)}(x - \vartheta) \tag{31}$$

$$\implies F(x) = F(\vartheta) + [1 - F(\vartheta)]G_{\xi, \psi(\vartheta)}(x - \vartheta). \tag{32}$$

The tail estimator for the underlying distribution $F(x|\xi, \psi(\vartheta))$ is constructed using an empirical estimate of $F(\vartheta)$, i.e., $\hat{F}(\vartheta) = (T - N_\vartheta)/T$ as

$$\hat{F}(x|\xi, \psi(\vartheta)) \approx \frac{T - N_\vartheta}{T} \left[1 + \frac{\hat{\xi}(x - \vartheta)}{\hat{\psi}(\vartheta)} \right]^{-1/\hat{\xi}}, \tag{33}$$

where N_ϑ is the number of observations above the threshold. We obtain the q^{th} quantile $F_q^{-1} = VaR_q$, by inverting Eq (33), for any given small upper tail probability p for VaR estimation as

$$VaR_q = \vartheta - \frac{\hat{\psi}(\vartheta)}{\hat{\xi}} \left\{ 1 - \left[\frac{T}{N_\vartheta} (1 - q) \right]^{-\hat{\xi}} \right\}, \tag{34}$$

where $q = 1 - p$ [4, 55, 56].

After deciding on the choice of ϑ , and assuming that the number of points above ϑ are independent and identically distributed, the parameters $\psi(\vartheta)$ and ξ can be estimated by means of maximum likelihood estimation with likelihood function

$$L(x_1, \dots, x_{N_\vartheta} | \xi, \sigma, \mu) = \prod_{i=1}^{N_\vartheta} f(x_i) \quad \text{for } x_i > \vartheta. \tag{35}$$

The choice of a threshold ϑ is an important step in the POT method because Eq (34) is dependent on ϑ and the number of points (i.e., exceedances) above ϑ since the parameters are estimated based on the exceedances. Thus, it is very important to find the proper threshold value. There is no clear-cut or wholly satisfactory method to determine a proper threshold that has been determined to date. [57] developed a semi-parametric estimator for the tails of the distribution that estimated the threshold of the bootstrap approximation of the mean square error (MSE) of the tail index and by minimising MSE through the choice of the threshold. [58] further used a two-step subsample bootstrap method to determine the threshold that minimised the asymptotic MSE. [59, 60] propose graphical tools to identify the proper threshold known as the Hill plot and the mean excess plot, respectively. In this paper, we use the mean excess plot and propose its extension, which we call a *hybrid* method as will be discussed later.

A mean excess function of x over a certain threshold ϑ is defined as

$$e(\vartheta) = E(x - \vartheta | x > \vartheta) = \frac{\sigma + \xi\vartheta}{1 - \xi}. \tag{36}$$

A property of the GPD states that if the excess distribution of x given a threshold ϑ_0 be a GPD with shape parameter ξ and scale parameter $\psi(\vartheta_0)$, then for any random threshold $\vartheta > \vartheta_0$, the excess distribution over the threshold ϑ has a GPD with shape parameter ξ and scale parameter $\psi(\vartheta) = \psi(\vartheta_0) + \xi(\vartheta - \vartheta_0)$, where $0 < \xi < 1$ [55]. Then

$$e(\vartheta) = E(x - \vartheta | x > \vartheta) = \frac{\psi(\vartheta_0) + \xi(\vartheta - \vartheta_0)}{1 - \xi}, \tag{37}$$

which is a linear function of $\vartheta - \vartheta_0$ with slope $\xi/(1 - \xi)$ for $\vartheta > \vartheta_0$. From the ordered sample $\{x_i\}$, we calculate and plot the mean excess function, i.e., Eq (37) against each chosen ϑ_i for $\vartheta_i > \vartheta_0$. The threshold ϑ is then identified as the lowest point on the mean excess plot above which the graph appears to be approximately linear. However, the choice of ϑ from the mean excess plot is subjective [8, 55] and might differ from one bank to another using the same data because of different risk tolerances. Different ϑ values will give different estimates of the shape

and scale parameter. A very high threshold will result in too few data points in the left tail for any meaningful statistical analysis. In contrast, a very low threshold will result in a number of data points above the threshold lying close to the body of the sample data. This will result in a poor approximation because the GPD is a limiting distribution as $\vartheta \rightarrow \infty$; data beyond the threshold will deviate from the GPD since the GPD is not a good approximation for the body of the sample data [8, 56]. We propose a *hybrid* method for selecting a proper threshold value that will significantly diminish the possibility of different ϑ values with the same data.

Data

The data employed in this analysis consist of 2870 daily observations of stock prices actively traded on the London Stock Exchange. The stocks belong to the banking sector and of the top five banks in the UK, i.e., HSBC Holdings, Lloyds Banking Group, Barclays Plc., Royal Bank of Scotland Group, and Standard Chartered Plc. We refer to these banks as Bank 1, Bank 2, Bank 3, Bank 4, and Bank 5, respectively. The motivation for selection of these banks is to test the reliability of the proposed VaR models for the top UK banks in periods of distress. Therefore, our data covers the period from 31 December 2004 to 31 December 2015, covering the 2008 global financial crisis and the 2011 European financial crisis. All data are from DataStream.

In some literature the stability in financial systems is measured using a portfolio consisting of several banks (i.e., by considering the dependence among the banks), while other studies focus on individual banks. This paper considers both measures. Therefore, by using the stock prices for each bank, we calculate the log-return series and apply risk factor mappings to construct a simulated portfolio of returns for all banks as follows: Consider a portfolio consisting of N risk factors represented in vector form as $S_N = (s_{1t}, \dots, s_{Nt})$, the log-returns r_t , are calculated as

$$r_t = \left[\log \left(\frac{S_{1,t+\tau}}{S_{1,t}} \right), \dots, \log \left(\frac{S_{N,t+\tau}}{S_{N,t}} \right) \right] = (r_{1t}, \dots, r_{Nt}). \tag{38}$$

Let \mathbf{Inv} be the total amount invested in the portfolio, x_i be the fraction of the total investment invested in stock i , $r_{i,t}$ is the return of stock i at time t , then the weight applied to $r_{i,t}$ is the fraction of the portfolio invested in stock i calculated as $w_i = \frac{x_i}{\mathbf{Inv}}$. Since the stocks are all from banks of almost the same strength (i.e., the top five banks in the UK), we may assume equal weights. Therefore, the expected return on the portfolio at time t is given by

$$\bar{R}_{p,t} = E(R_{p,t}) = \sum_{i=1}^N w_i E(\mathbf{r}_{i,t}), \quad \sum_{i=1}^N w_i = 1, \tag{39}$$

which is a weighted average of the return on the individual stocks in the portfolio.

Fig 1 shows time plots of the log-return series and the portfolio; this shows evidence of volatility clustering in the return series. From the figure, we can also see the effects of the 2008 global financial crisis and the 2011 European financial crisis.

Table 1 presents summary statistics of the data. We see from the table that the log-return series for each bank and the portfolio are far from being normally distributed as indicated by their high excess kurtosis and skewness. Furthermore, Jarque-Bera normality tests, Ljung-Box tests on the squared residuals $a_{i,t}^2$; where $a_{i,t} = r_{i,t} - \mu_i$ (μ_i being the unconditional mean), and a Lagrange multiplier tests for autoregressive conditional heteroscedasticity (ARCH LM test) on the residuals $a_{i,t}$, as described in [55, 61, 62], are significant at 5% level.

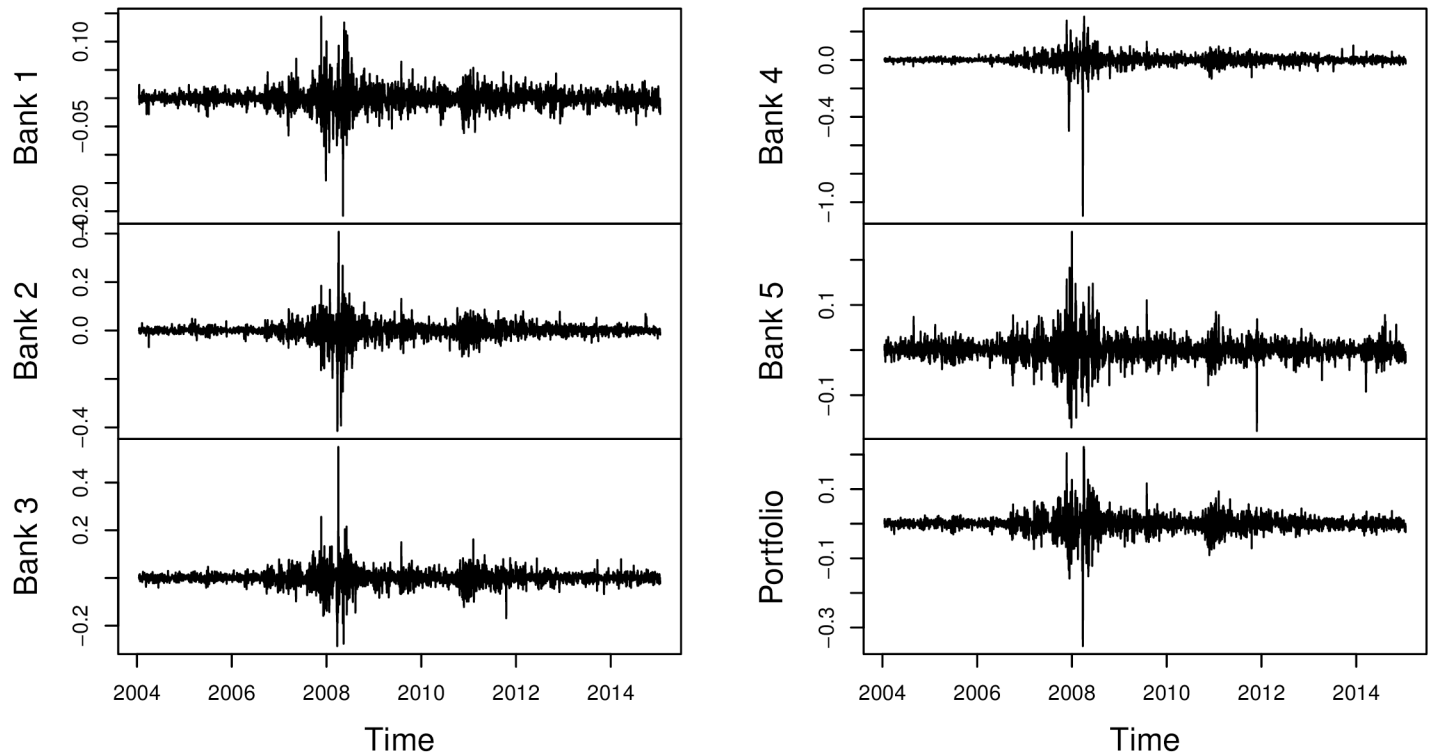


Fig 1. Time plots of the log-return series. Plots show the presence of volatility clustering in the log-return data.

<https://doi.org/10.1371/journal.pone.0198753.g001>

Table 1. Summary statistics of daily log-returns and portfolio return series.

	Bank 1	Bank 2	Bank 3	Bank 4	Bank 5	Portfolio
Mean	-0.0001	-0.0004	-0.0003	-0.0010	-0.0001	-0.0004
Variance	0.0003	0.0011	0.0010	0.0015	0.0006	0.0006
Std. deviation	0.0171	0.0328	0.0321	0.0388	0.0244	0.0239
Skewness	-0.3367	-1.0549	1.4387	-8.4013	0.3161	-0.7549
Excess kurtosis	16.9080	37.2754	40.2179	235.5263	13.0850	28.6129

<https://doi.org/10.1371/journal.pone.0198753.t001>

Results

Modelling the marginal distributions of volatility equations

As noted, the log-return series are leptokurtic and skewed. Thus, to capture the tail distribution and the dynamics of fluctuations in the time series data, we consider a single-state, $k = 1$ and two-state, $k = \{1, 2\}$ Markov Switching GARCH specifications. The underlying volatility model is a GJR-GARCH(1,1) model with skewed Student's- t distribution. Since we use just one variance specification (i.e., GJR-GARCH), the two-state Markov Switching GARCH is generated by setting the number of regimes in the conditional distribution to 2. For the single-state, the length of the variance specification is equal to the length of the conditional distribution, which is 1 (see [14]). Also note that the single-state Markov Switching GJR-GARCH(1,1)

Table 2. Parameter estimates following Bayesian GJR-GARCH(1,1) model with skewed Student's-t distribution.

	α_0	α_1	α_2	β_1	ν	γ
Bank 1	7.0531e-06(0.0000)	0.0508(0.0010)	0.1001(0.0000)	0.8488(0.0002)	5.8153 (0.0114)	1.0067(0.005)
Bank 2	1.4764e-6(0.0000)	0.0509(0.0000)	0.1001(0.0000)	0.8570(0.0001)	6.4085 (0.0133)	1.0014 (0.0005)
Bank 3	7.0777e-06(0.0000)	0.0511(0.0000)	0.1001(0.0000)	0.8716(0.0001)	6.1691 (0.0118)	1.0009 (0.0005)
Bank 4	9.4281e-06(0.0000)	0.0511(0.0000)	0.1002(0.0000)	0.8688(0.0001)	5.9166 (0.0111)	1.0160 (0.0005)
Bank 5	2.0683e-05(0.0000)	0.0508(0.0000)	0.1002(0.0000)	0.8321(0.0002)	6.3657 (0.0138)	1.0266 (0.0005)
Portfolio	5.4112e-06(0.0000)	0.0510(0.0000)	0.1002(0.0000)	0.8670(0.0001)	9.4379 (0.0298)	0.9936 (0.0005)

Note: Standard errors in parentheses.

<https://doi.org/10.1371/journal.pone.0198753.t002>

model corresponds to GJR-GARCH(1,1) model without regime change. Therefore, we simply refer to the single-state and two-state Markov Switching GJR-GARCH(1,1) models as GJR-GARCH(1,1) and MS-GJR-GARCH(1,1) models, respectively (see [14]). GARCH parameters are estimated using Bayesian statistics as follows: (i) We assign a prior distribution with initial hyperparameters and generate MCMC simulations of 40000 draws; (ii) if convergence is attained, we discard the first 20000 draws and select only the 10th draw from each chain such that auto-correlation between draws is reduced to almost zero. We merge the two chains together to obtain a sample data set of 2000 observations. (iii) If convergence is not attained, repeat (i) using parameter estimates from the previous draw as the hyperparameters to increase the chance of convergence. The mean value of each parameter with respect to its respective posterior distribution is the optimal parameter estimate of the Bayesian GJR-GARCH(1,1) and Bayesian MS-GJR-GARCH(1,1) models with skewed Student's-t distributions. Estimation results are presented in Tables 2 and 3 with standard errors in parenthesis. For MS-GJR-GARCH(1,1) model, the degrees of freedom parameter, ν is fixed across the regimes.

Table 3. Parameter estimates for two-state MS-GJR-GARCH(1,1) model with skewed Student's-t distribution.

	$k = 1$					
	$\alpha_{0,1}$	$\alpha_{1,1}$	$\alpha_{2,1}$	$\beta_{1,1}$	ν	γ_1
Bank 1	2.9335e-07(0.0000)	0.0270 (0.0010)	0.0121 (0.0004)	0.9612 (0.0005)	6.2679 (0.0159)	1.0380 (0.0009)
Bank 2	1.9132e-06(0.0000)	0.0302 (0.0004)	0.0811 (0.0014)	0.9208 (0.0011)	7.9120 (0.0270)	1.0115 (0.0010)
Bank 3	2.7159e-07(0.0000)	0.0109 (0.0002)	0.0253 (0.0003)	0.9729 (0.0002)	5.6388 (0.0142)	0.9516 (0.0007)
Bank 4	1.0034e-07(0.0000)	0.0367 (0.0002)	0.0030 (0.0001)	0.9595 (0.0002)	7.4127 (0.0220)	1.0146 (0.0009)
Bank 5	3.8283e-06(0.0000)	0.0338 (0.0008)	0.0932 (0.0032)	0.9083 (0.0027)	7.2792 (0.0186)	1.0462 (0.0011)
Portfolio	1.5441e-05(0.0000)	0.0341 (0.0004)	0.1694 (0.0040)	0.8665 (0.0025)	14.1506 (0.0621)	0.9954 (0.0013)
	$k = 2$					
	$\alpha_{0,2}$	$\alpha_{1,2}$	$\alpha_{2,2}$	$\beta_{1,2}$	ν	γ_2
Bank 1	1.0589e-05(0.0000)	0.0412 (0.0005)	0.1568 (0.0011)	0.8566 (0.0007)	6.2679 (0.0159)	0.9496 (0.0012)
Bank 2	1.8650e-05(0.0000)	0.0056 (0.0002)	0.2322 (0.0027)	0.8586 (0.0015)	7.9120 (0.0270)	0.9782 (0.0018)
Bank 3	1.5749e-05(0.0000)	0.0558 (0.0004)	0.0776 (0.0014)	0.9045 (0.0006)	5.6388 (0.0142)	1.1826 (0.0026)
Bank 4	6.8672e-05(0.0000)	0.0683 (0.0013)	0.8202 (0.0043)	0.4771 (0.0022)	7.4127 (0.0220)	1.0138 (0.0023)
Bank 5	1.8827e-05(0.0000)	0.1016 (0.0011)	0.3991 (0.0040)	0.6778 (0.0028)	7.2792 (0.0186)	1.0085 (0.0015)
Portfolio	2.2462e-06(0.0000)	0.0334 (0.0004)	0.2052 (0.0040)	0.8543 (0.0023)	14.1506 (0.0298)	0.9936 (0.0005)

Note: Standard errors in parentheses. Degrees of freedom parameter, ν is fixed across the regimes.

<https://doi.org/10.1371/journal.pone.0198753.t003>

Table 4. ARCH LM test on the standardised residuals and Ljung-Box test on the standardised squared residuals for $k = 1$. The null hypothesis of no ARCH effect or serial correlation is rejected at 5% significant level for Bank 4.

$k = 1$	ARCH LM test						Ljung-Box test				
	Bank 1	Bank 2	Bank 3	Bank 4	Bank 5		Bank 1	Bank 2	Bank 3	Bank 4	Bank 5
LM(5)	2.21	2.35	2.96	17.78	4.01	Q(5)	2.24	2.35	2.95	17.65	4.11
p -value	0.820	0.800	0.706	0.003	0.548	p -value	0.815	0.799	0.707	0.003	0.534
LM(10)	10.26	4.13	7.04	18.62	6.79	Q(10)	9.92	4.14	7.17	18.57	6.89
p -value	0.820	0.942	0.722	0.045	0.745	p -value	0.447	0.941	0.710	0.046	0.736
$k = \{1, 2\}$	ARCH LM test						Ljung-Box test				
	Bank 1	Bank 2	Bank 3	Bank 4	Bank 5		Bank 1	Bank 2	Bank 3	Bank 4	Bank 5
LM(5)	2.164	2.29	8.74	5.31	3.06	Q(5)	2.13	2.22	8.60	5.30	10.56
p -value	0.826	0.807	0.120	0.379	0.690	p -value	0.831	0.818	0.126	0.380	0.061
LM(10)	6.01	3.75	13.39	5.83	5.30	Q(10)	5.98	3.70	13.27	5.88	12.71
p -value	0.815	0.958	0.203	0.829	0.870	p -value	0.817	0.960	0.209	0.826	0.240

Note: For $k = 1$, we have a GJR-GARCH(1,1) model, and for $k = \{1, 2\}$, we have a MS-GJR-GARCH(1,1) model.

<https://doi.org/10.1371/journal.pone.0198753.t004>

Applying Eq (2), we then obtain a matrix Σ , which consists of the filtered marginal standardised residuals, $\{\epsilon_{i,t}\}_{t=1}^T$, of the overall process for the MS-GJR-GARCH(1,1) model and GJR-GARCH(1,1) model. That is

$$\Sigma_{i,t} = (r_{i,t}) \left(h_{\Delta_{i,t},t}^{-\frac{1}{2}} \right), \quad i = 1, \dots, N; \quad t = 1, \dots, T. \tag{40}$$

The ARCH LM test and Ljung-Box test on the standardised residuals and standardised squared residuals, respectively, for lags 5 and 10 are presented in Table 4. For the GJR-GARCH(1,1) model, there still exist some serial correlation in the standardised residuals of bank 4. For MS-GJR-GARCH(1,1) model, there is no evidence of an ARCH effect or serial correlations in the standardised residuals.

Modelling dependence with copulas

We model the dependence structure among the stock returns using copula functions. Copula parameters are estimated by the canonical maximum likelihood (CML) method [41]. This entails the use of pseudo-observations of the standardised residuals to estimate the marginals. We then estimate the copula parameters by inversion of Kendall's τ , which is one of the most commonly used invariant measures and has been proven to provide more efficient ways of estimating correlations [63, 64]. The copula that fits the data best is selected by maximum likelihood estimation (MLE) method by maximising the likelihood function

$$\hat{\Psi}_2 = \text{ArgMax}_{\Psi_2} \sum_{t=1}^T \ln c(\hat{F}_1(X_{1t}), \dots, \hat{F}_n(X_{nt}); \Psi_2), \tag{41}$$

where $\hat{\Psi}_2$ are estimates of the copula parameters. The estimated copula parameters are reported in Table 5, along with their Akaike information criterion (AIC) values. For both models, Frank and Student's- t copulas are selected from each copula family based on the highest MLE values. From Table 5, the same copula types have been selected based on the AIC values (the copula with the smallest AIC value is preferred). Note that Gaussian copula gives a higher MLE value compared to the Archimedean copulas but also higher AIC value. Table 6

Table 5. Copula parameter estimates are based on inversion of Kendall's τ following CML estimation method.

GJR-GARCH(1,1)	Archimedean copulas			Elliptical copulas	
	Gumbel	Clayton	Frank	Gaussian	Student's- <i>t</i>
Kendall's τ	1.782 (0.023)	1.563 (0.046)	4.697 (0.042)	$\rho_G = \rho_\tau(\rho_{SE})$	$\rho_t = \rho_\tau(\rho_{SE})$
MLE	3226	2745	3250	3846	4108
AIC	-14.158	-13.835	-14.173	3.491	3.359
MS-GJR-GARCH(1,1)	Archimedean copulas			Elliptical copulas	
	Gumbel	Clayton	Frank	Gaussian	Student's- <i>t</i>
Kendall's τ	1.773 (0.023)	1.546 (0.046)	4.657 (0.041)	$\rho_G = \rho_\tau(\rho_{SE})$	$\rho_t = \rho_\tau(\rho_{SE})$
MLE	3163	2705	3206	3773	4013
AIC	-14.119	-13.806	-14.146	3.529	3.405

Note: Standard errors in parentheses. The best copula for modeling dependence among the risk factors is that with the highest MLE value or smallest AIC value (in bold).

<https://doi.org/10.1371/journal.pone.0198753.t005>

Table 6. Kendall's τ ; $\rho_\tau(\rho_{SE})$ for Gaussian and Student's-*t* copula parameter estimates.

		Bank 1	Bank 2	Bank 3	Bank 4	Bank 5
GJR-GARCH(1,1)	Bank 1	1				
	Bank 2	0.6230 (0.013)	1			
	Bank 3	0.5521 (0.015)	0.7054 (0.011)	1		
	Bank 4	0.5741 (0.014)	0.7262 (0.011)	0.7176 (0.011)	1	
	Bank 5	0.6383 (0.013)	0.6027 (0.014)	0.5437 (0.015)	0.5460 (0.015)	1
MS-GJR-GARCH(1,1)	Bank 1	1				
	Bank 2	0.6257 (0.013)	1			
	Bank 3	0.5544 (0.015)	0.7074 (0.011)	1		
	Bank 4	0.5779 (0.015)	0.7282 (0.011)	0.7225 (0.011)	1	
	Bank 5	0.6437 (0.012)	0.6075 (0.014)	0.5439 (0.015)	0.5483 (0.015)	1

Note: Standard errors in parentheses.

<https://doi.org/10.1371/journal.pone.0198753.t006>

shows the Kendall's τ for Gaussian and Student's-*t* copula parameter estimates. Thus, the Gaussian copula is not a good fit for the data. The analysis continues based on the selected copulas. Next, we specify the desired marginal distributions, which we set to Student's-*t* distribution, and using the estimated copula parameters, we generate 10000 simulations to obtain a new matrix of marginal standardised residuals

$$\hat{\Sigma} = \{\zeta_{i,j}\}, \quad j = 1, \dots, T, \quad i = 1, \dots, N, \tag{42}$$

which is free from assumptions of normality and linear correlations. To confirm this, we employ a multivariate ARCH test based on the Ljung-Box test statistics

$$Q_k(m) = T^2 \sum_{i=1}^m \frac{1}{T-i} \mathbf{b}'_i (\hat{\rho}_0^{-1} \otimes \hat{\rho}_0^{-1}) \mathbf{b}_i \approx \chi_{k^2}^2(m), \tag{43}$$

and its modification $Q_k^r(m)$, known as a robust test, on the log returns at 5% significance level, where m is the number of lags of cross-correlation matrices used in the tests, k is the dimension of $r_{i,t}$, T is the sample size, $\mathbf{b}_i = \text{vec}(\hat{\rho}'_i)$ with $\hat{\rho}_j$ being the lag- j cross-correlation matrix of

Table 7. Multivariate ARCH test on $\{\zeta_{i,j}\}$ shows no evidence of conditional heteroscedasticity.

	GJR-GARCH(1,1)		MS-GJR-GARCH(1,1)	
	Frank copula	$Q_k(10) = 11.413$ $p\text{-value} = 0.326$	$Q_k^r(10) = 267.925$ $p\text{-value} = 0.208$	$Q_k(10) = 5.288$ $p\text{-value} = 0.871$
t-copula	$Q_k(10) = 5.507$ $p\text{-value} = 0.855$	$Q_k^r(10) = 235.133$ $p\text{-value} = 0.742$	$Q_k(10) = 2.554$ $p\text{-value} = 0.990$	$Q_k^r(10) = 249.171$ $p\text{-value} = 0.503$

<https://doi.org/10.1371/journal.pone.0198753.t007>

$r_{i,t}^2$. The modification $Q_k^r(m)$ involves discarding those observations from the return series whose corresponding standardised residuals exceed 95th quantile in order to reduce the effect of heavy tails. The motivation for $Q_k^r(m)$ test is that $Q_k(m)$ may fare poorly in finite samples when the residuals of the time series, $r_{i,t}$, have heavy tails [45]. The tests show no evidence of conditional heteroscedasticity lags $m = 10$; Table 7.

We follow the approach by [9] and apply the POT method of EVT to each of the marginal distributions of $\{\zeta_{i,t}\}$ (i.e., Eq (42)) to obtain the q^{th} quantile, $Var(Z)_q$ of the noise variables for VaR estimation. Let $\{\chi_{i,\tau}\}$ be the negative variables of the marginal distributions of $\{\zeta_{i,t}\}$ such that $\{\chi_{i,\tau}\} \subseteq \{\zeta_{i,t}\}$. Then, from the ordered sample of $\{\chi_{i,\tau}\}$, we calculate and plot the mean excess function to help identify the threshold. As an example, Figs 2 and 3 are mean excess function plots for Bank 1 following the Bayesian GJR-GARCH(1,1) Student's-t and Frank copula models. The plots suggest us to select threshold values of about 1.3 and 1.4 for Figs 2 and 3, respectively, which are the lowest points on the graphs above which the graph appears to be

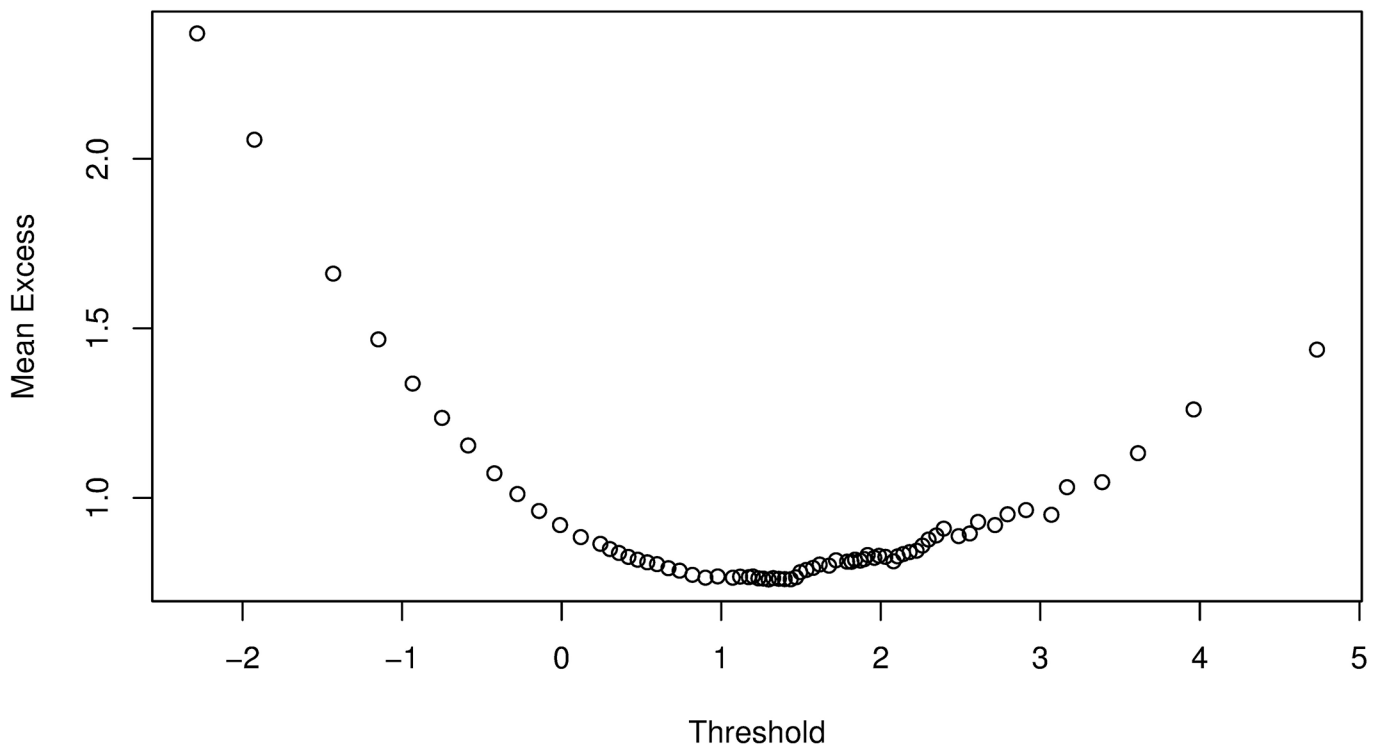


Fig 2. Mean excess function plot for Bank 1 following analysis with Bayesian GJR-GARCH(1,1) Student's-t copula model.

<https://doi.org/10.1371/journal.pone.0198753.g002>

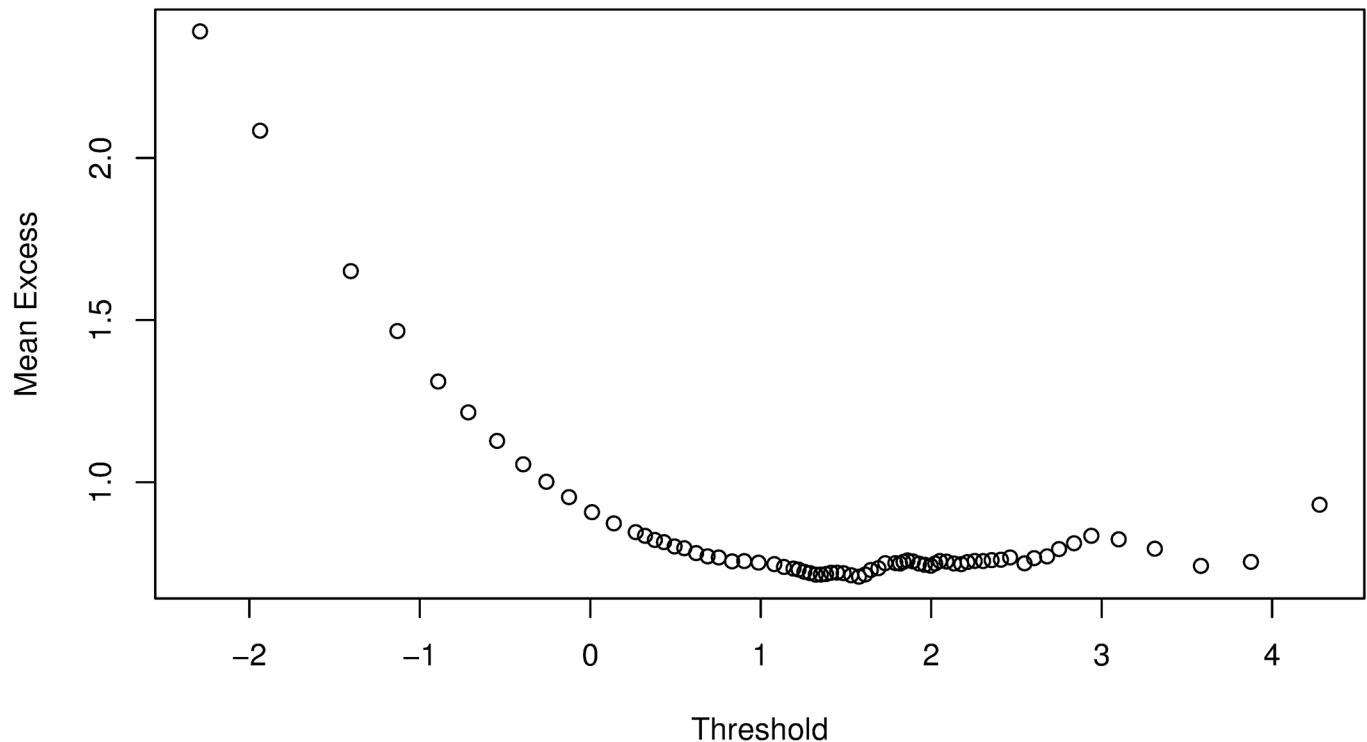


Fig 3. Mean excess function plot for Bank 1 following analysis with Bayesian GJR-GARCH(1,1) Frank copula model.

<https://doi.org/10.1371/journal.pone.0198753.g003>

approximately linear. However, if we select these points as the threshold values, we will have 1402 exceedances for Fig 2 and 1039 exceedances for Fig 3, which are too many compared to the size of the data (i.e., $T = 10000$). The number of exceedances thus lie towards the body of the data, which will inevitably result in a poor approximation of the GPD parameters and hence lead to inaccuracies in the VaR estimate. In addition, the threshold selection method is very subjective and will be different from one analyst to the other based on their preferences.

We propose an extension to the mean excess plot for threshold selection; the *hybrid* method. That is, from the mean excess plot, we identify the lowest point, making the graph appears approximately linear, a point ϑ_0 . We then insert a tangent line from ϑ_0 through the rest of the points ϑ_i , where $\vartheta_i > \vartheta_0$; see Fig 4. Since the tangent to a linear curve is the tangent itself and the mean excess function is a linear function of the threshold, we take an average of the set of points that touches the tangent line as the threshold value, a point ϑ^* . This point ϑ^* will lead to a better approximation of VaR estimates than ϑ_0 because the inference is restricted to the left tail. Apart from better approximation of VaR estimates, this method significantly reduces the probability of having different VaR estimates for the same data and also the probability of selecting a very low or very high threshold value. Let $\vartheta_i = \vartheta_1, \dots, \vartheta_{\hat{h}}$ be a set of points that touches the tangent line, then we obtain the value of ϑ^* as

$$\vartheta^* = \frac{1}{\hat{h}} \sum_{i=1}^{\hat{h}} \vartheta_i, \quad \vartheta_i \geq \vartheta_0, \tag{44}$$

where \hat{h} is the number of points in the set.

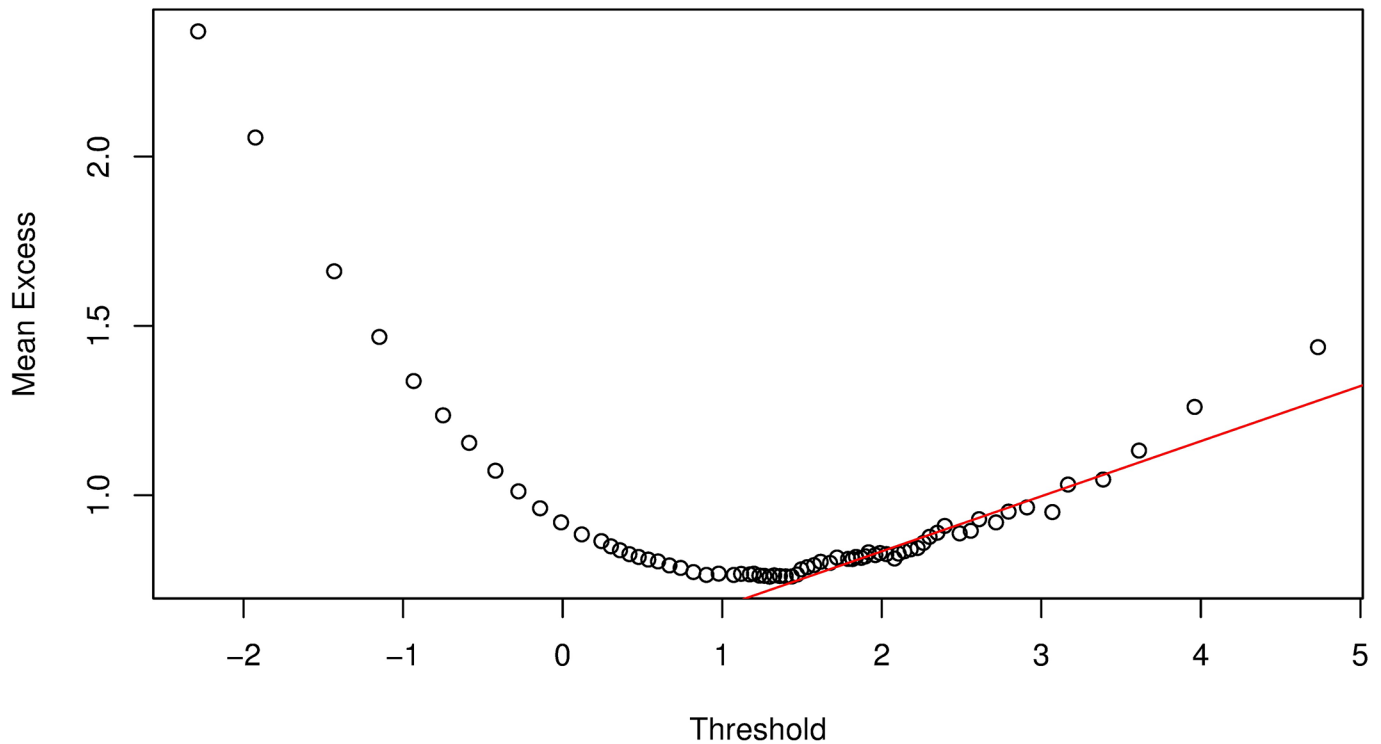


Fig 4. Mean excess function plot demonstrating the hybrid method for threshold selection.

<https://doi.org/10.1371/journal.pone.0198753.g004>

It can be seen in Fig 4 that the points touching the tangent line, i.e., ϑ_0 , are too compact and might lead us to miss some important points. A better way for selecting these points is by fitting a regression line

$$\hat{y} = b_0 + b_1x, \tag{45}$$

which is based on the least square method to the points $\{\vartheta_i\}_{i=1}^h$, where \hat{y} is the estimate of the dependent variable, and x is the independent variable with intercept b_0 and slope b_1 . In the presence of heteroscedasticity and outliers, it may be advantageous to consider fitting a robust regression line. Robust regression methods are not influenced by outliers, and are also very useful when there are problems with heteroscedasticity in the data set. This method is demonstrated for Bank 1 in Figs 5 and 6, which illustrates a comparison between simple linear and robust regression methods. It can be seen that the regression lines for standard regression models are affected by outliers in the left tails of the mean excess plots, hence a robust regression model is more reliable.

Following the above analysis, we obtained a threshold value of 2.2448 and 333 exceedances for Fig 5, and 2.6476 and 176 exceedances for Fig 6. The analysis are restricted to the tails and the data is sufficient to allow for reasonable statistical inferences with EVT. Tables 8 and 9 presents the POT parameter estimates and the forecast VaR estimates. The portfolio VaR estimates, $VaR_q^p(Z)$, based on the individual bank's VaR estimates and confidence level are also reported. $VaR_q^p(Z)$ is computed using the risk formula

$$VaR_q^p(Z) = \left(\sum_{i=1}^N w_i^2 VaR_{q,i}^2(Z) + 2w_i w_j \sum_{i<j} \rho_{ij} VaR_{q,i}(Z) VaR_{q,j}(Z) \right)^{\frac{1}{2}}, \quad \sum_{i=1}^N w_i = 1, \tag{46}$$

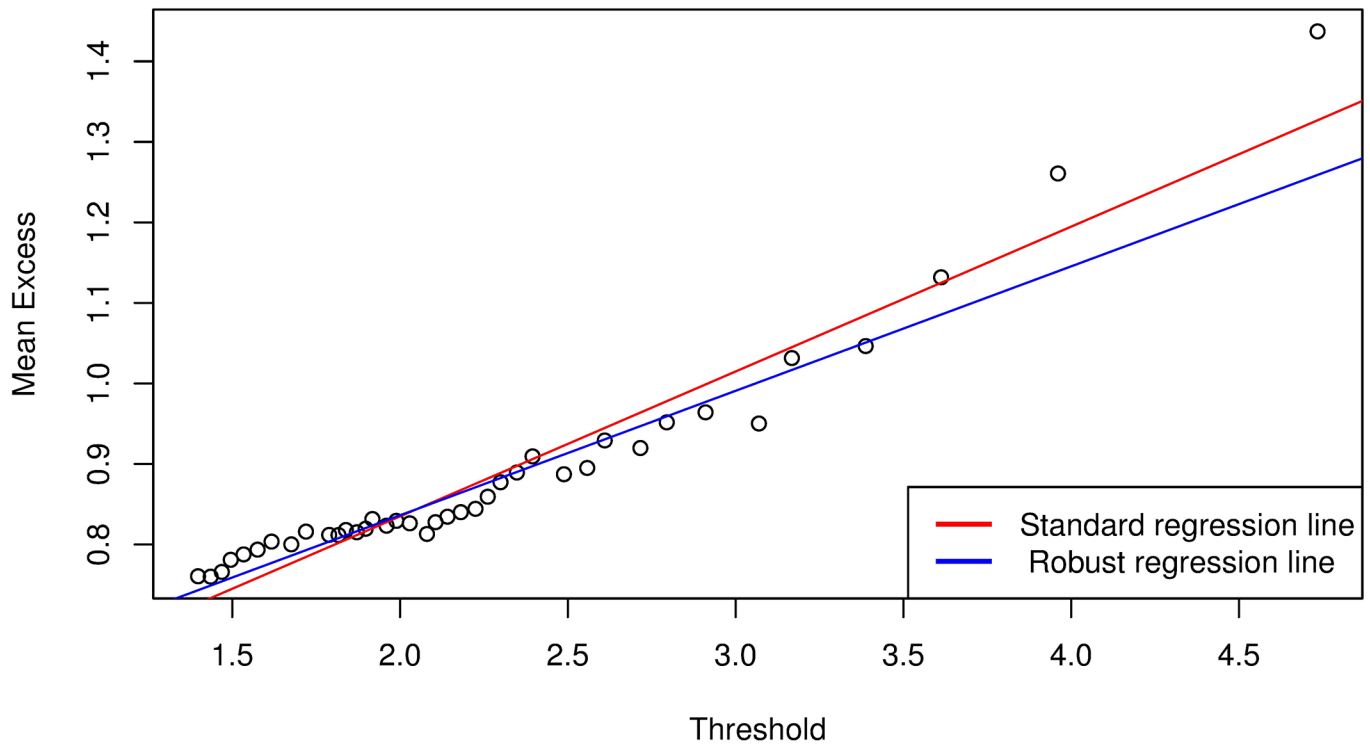


Fig 5. Mean excess function plot following analysis with Bayesian GJR-GARCH(1,1) Student's-*t* copula model for the number of exceedances above ϑ_0 . A reliable threshold is calculated by taking an average of the set of points that touches the robust regression line. The standard regression line is affected by outliers in the left tail.

<https://doi.org/10.1371/journal.pone.0198753.g005>

where $\rho_{i,j}$ is the Pearson cross-correlation coefficient between the returns of the i th and j th stocks. As noted, the overall risk measures are quite stable for both models and different thresholds indicating that the model has effectively captured the dynamics of fluctuations in the left tails of the return distributions. This claim may be validated through back-testing the model. We can also see the effect of diversification on the risk of the individual banks on the portfolio VaR. Employing Eq (46), the one step ahead VaR is then calculated as

$$VaR_{q,t}^p = VaR_q^p(Z)\hat{h}_{\Delta_r,t+1}^{\frac{1}{2}}. \tag{47}$$

Note that $\hat{h}_{\Delta_r,t+1}^{\frac{1}{2}}$ is the one-step-ahead conditional volatility forecast of the overall conditional variance for the portfolio at time $t + 1$ for state k , Δ_r is a Markov chain as defined in Eqs (1) and (2), but for \bar{R}_p is by Eq (39). That is, $\bar{R}_{p,i} | (\Delta_r = k, \Omega_{t-1})$ and the parameters are sampled from the posterior distribution using MH algorithm.

Figs 7–10 show time plots of profit and loss (P&L) of the portfolio return series and forecasts portfolio VaR estimates at 99% and 95% confidence levels. A visual observation of the plots suggests that the VaR models perform quite well in capturing the dynamics in the portfolio return series.

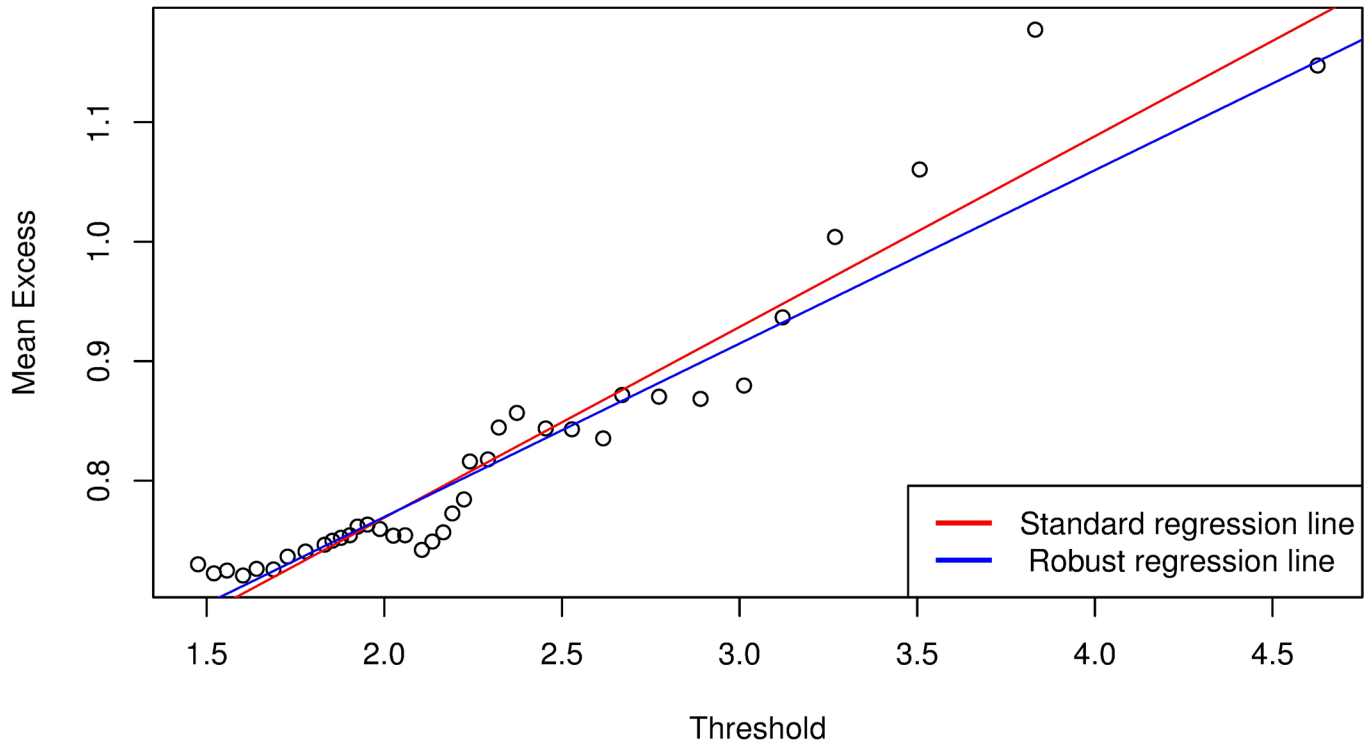


Fig 6. Mean excess function plot following analysis with Bayesian GJR-GARCH(1,1) Frank copula model for the number of exceedances above ϑ_0 .

<https://doi.org/10.1371/journal.pone.0198753.g006>

Table 8. POT parameter estimates, $VaR_q(Z)$ and $VaR_q^p(Z)$ following Bayesian GJR-GARCH(1,1) Frank and Student's- t copula-EVT models.

		Parameters						$VaR_q(Z)$	
		ξ	$\psi(\vartheta^*)$	ϑ^*	N_{ϑ^*}	μ	σ	99%	95%
Student's- t copula:	Bank 1	0.1660	0.7143	2.2448	333	0.3881	0.4061	3.1958	1.9640
	Bank 2	0.2239	0.6479	2.4624	218	0.7974	0.2750	3.0141	1.9716
	Bank 3	0.1838	0.6353	2.2321	356	0.6481	0.3441	3.1407	2.0229
	Bank 4	0.1273	0.7301	2.4687	271	0.3567	0.4612	3.2448	2.0385
	Bank 5	0.1465	0.7135	2.3293	287	0.3538	0.4242	3.1428	1.9489
$VaR_q^p(Z)$								2.5862	1.6282
Frank copula:	Bank 1	0.1019	0.7239	2.6476	176	0.2503	0.4796	3.0688	1.9305
	Bank 2	0.0497	0.7489	2.4331	235	-0.1297	0.6216	3.0867	1.8781
	Bank 3	0.0390	0.7892	2.6040	223	-0.1855	0.6804	3.2469	1.9767
	Bank 4	0.2073	0.6862	2.5407	217	0.7266	0.3102	3.1174	2.0147
	Bank 5	0.1062	0.6892	3.1337	105	0.6440	0.4249	3.1674	2.1425
$VaR_q^p(Z)$								2.5410	1.6459

Note: Estimations for a time horizon of 1 day at $q = (99\%, 95\%)$ confidence level. The risk measures are quite stable for different thresholds and copula functions indicating that the VaR models have successfully capture the dynamics of fluctuations in the left tails.

<https://doi.org/10.1371/journal.pone.0198753.t008>

Table 9. POT parameter estimates, $VaR_q(Z)$ and $VaR_q^p(Z)$ following Bayesian MS-GJR-GARCH(1,1) Frank and Student's- t copula-EVT models.

		Parameters						$VaR_q(Z)$	
		ξ	$\psi(\vartheta^*)$	ϑ^*	N_{ϑ^*}	μ	σ	99%	95%
Student's- t copula:	Bank 1	0.0194	0.8795	1.8425	565	-0.4893	0.8343	3.1862	1.9389
	Bank 2	0.0239	0.8003	1.9903	409	-0.3040	0.7455	2.9471	1.8512
	Bank 3	0.0462	0.7642	2.0585	453	-0.1450	0.6624	3.2542	1.9832
	Bank 4	0.0686	0.6675	2.3155	260	0.7943	0.5632	3.0091	1.8140
	Bank 5	0.0038	0.6891	1.9164	493	-0.1460	0.6814	3.0191	1.9067
$VaR_q^p(Z)$								2.5127	1.5471
Frank copula:	Bank 1	0.0868	0.8295	2.5454	201	-0.2030	0.5910	3.1457	1.9804
	Bank 2	0.0970	0.7780	1.9443	440	-0.7856	0.5132	2.9769	1.8480
	Bank 3	0.0819	0.7132	2.1074	358	0.0287	0.5430	3.0662	1.8724
	Bank 4	0.0806	0.6030	2.5029	200	0.4796	0.4399	2.9327	1.9703
	Bank 5	0.0924	0.6688	2.3509	230	0.2210	0.4719	2.9300	1.8498
$VaR_q^p(Z)$								2.4535	1.5513

Note: Estimations for a time horizon of 1 day at $q = (99\%, 95\%)$ confidence level. The risk measures are quite stable for different thresholds and copula functions indicating that the VaR models have successfully capture the dynamics of fluctuations in the left tails.

<https://doi.org/10.1371/journal.pone.0198753.t009>

GJR-GARCH(1,1) Student's-t copula EVT model

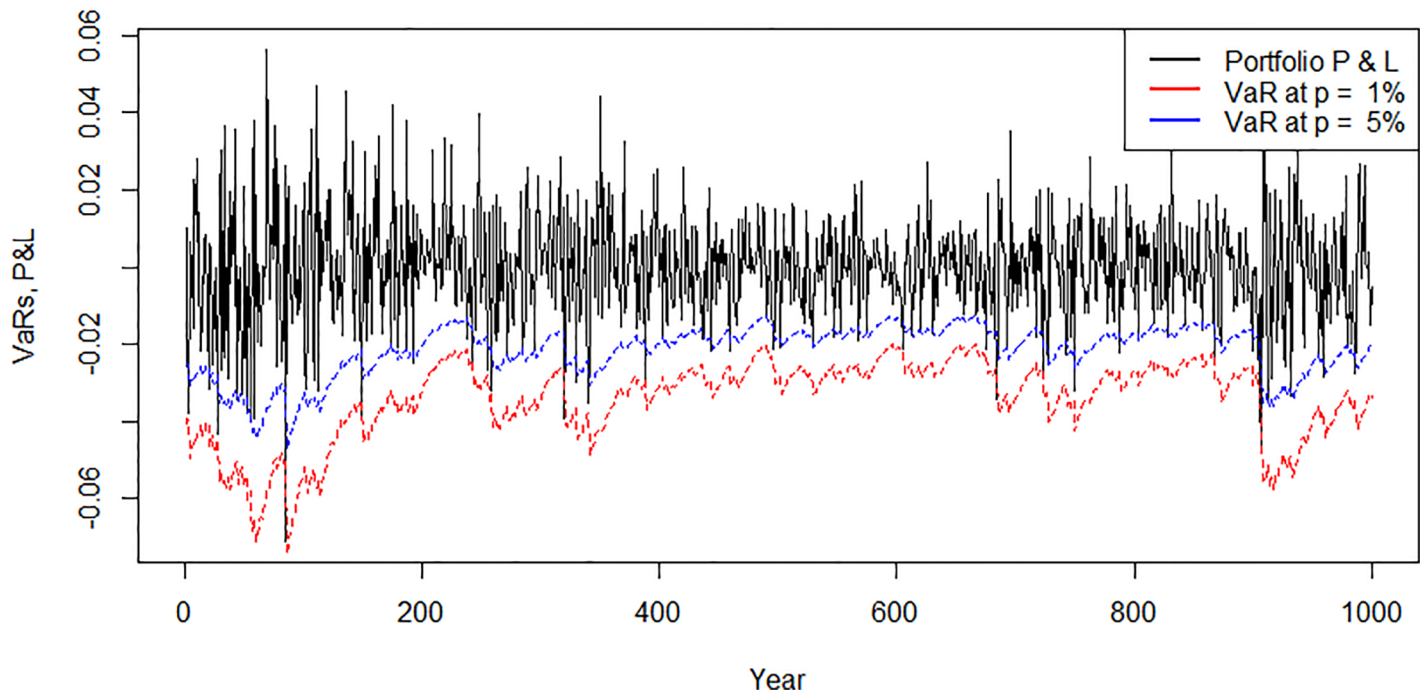


Fig 7. Forecasts daily VaR estimates and daily profit and loss (P&L) plots for an investment in a portfolio consisting of all banks following Bayesian GJR-GARCH(1,1) Student's- t copula EVT VaR model.

<https://doi.org/10.1371/journal.pone.0198753.g007>

GJR-GARCH(1,1) Frank copula EVT model

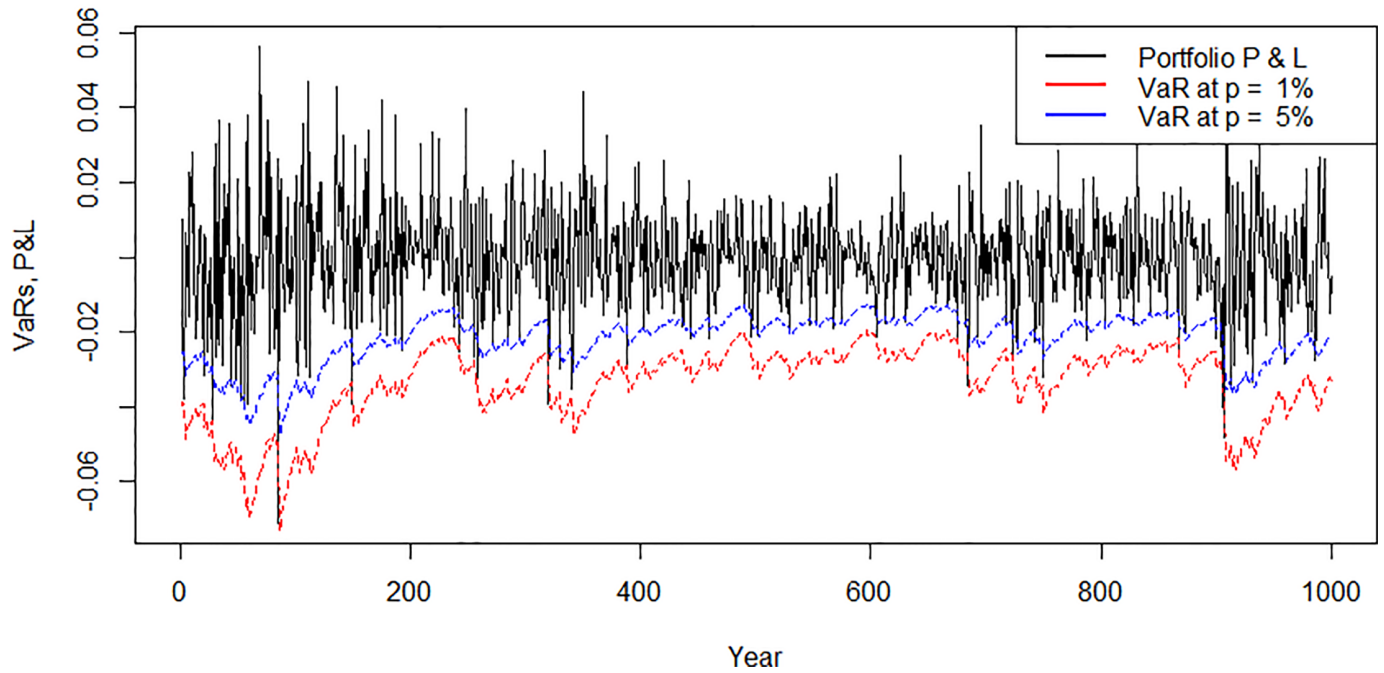


Fig 8. Forecasts daily VaR estimates and daily profit and loss (P&L) plots for an investment in a portfolio consisting of all banks following Bayesian GJR-GARCH(1,1) Frank copula EVT VaR model.

<https://doi.org/10.1371/journal.pone.0198753.g008>

MS-GJR-GARCH(1,1) Student's-t copula EVT model

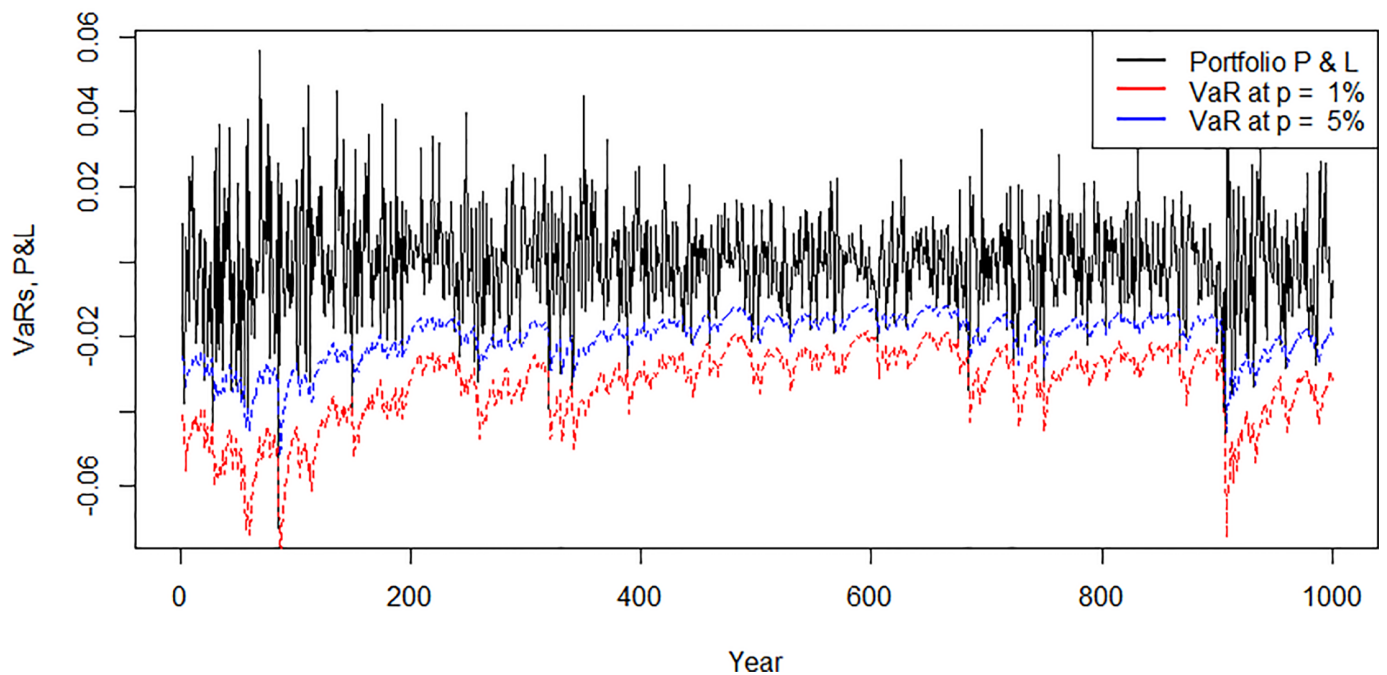


Fig 9. Forecasts daily VaR estimates and daily profit and loss (P&L) plots for an investment in a portfolio consisting of all banks following Bayesian MS-GJR-GARCH(1,1) Student's-t copula EVT VaR model.

<https://doi.org/10.1371/journal.pone.0198753.g009>

MS-GJR-GARCH(1,1) Frank copula EVT model

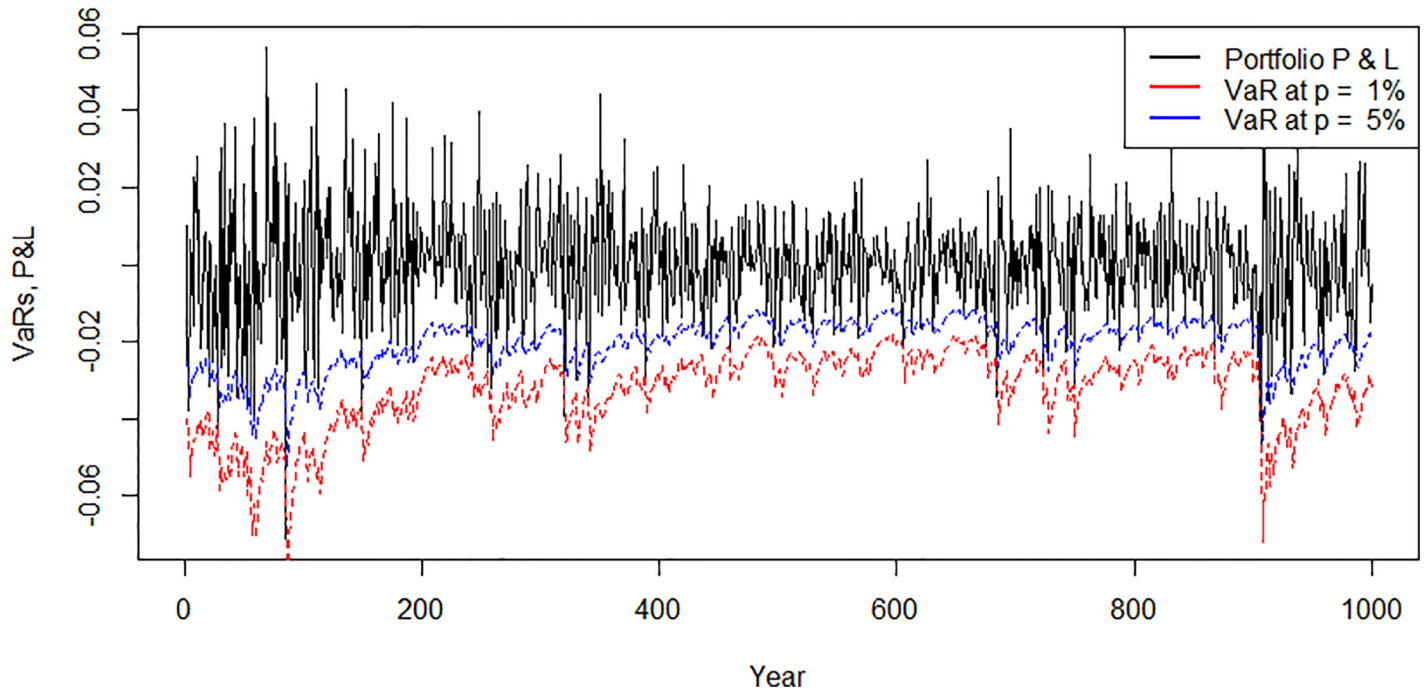


Fig 10. Forecasts daily VaR estimates and daily profit and loss (P&L) plots for an investment in a portfolio consisting of all banks following Bayesian MS-GJR-GARCH(1,1) Frank copula EVT VaR model.

<https://doi.org/10.1371/journal.pone.0198753.g010>

Model checking

The reliability of the VaR model is often tested by performing back-testing. This involves comparing the estimated VaRs for a given time horizon and observation period to the subsequent returns and recording the number of days in which the loss on the portfolio exceeds VaR. The number of days T_1 in which the loss on the portfolio exceeds VaR is recorded as the number of exceptions or failures. Too many exceptions imply that the VaR model underestimates the level of risk, and too few exceptions imply the model overestimates risk. For the VaR model to be accepted as a reliable risk measure, the number of exceptions produced for any given observation period should satisfy the unconditional coverage (UC) and independent (IND) properties. We define an indicator function on the exceptions at time t as

$$\mathbb{I}_t(1 - q) = \mathbb{I}_{\{L_t > VaR_{q,t}^p\}} = \begin{cases} 1, & \text{if } L_t > VaR_{q,t}^p \\ 0, & \text{otherwise,} \end{cases} \quad (48)$$

where the indicator function equal to 1 when the loss on the portfolio L_t exceed $VaR_{q,t}^p$, and 0 if otherwise; note that q is the choice of confidence level. For the UC property, $\Pr [\mathbb{I}_t(1 - q) = 1] \approx 1 - q, \forall t$; i.e., the number of exceptions should be reasonably close to $T_w(1 - q)\%$, depending on the choice of q , and should follow a binomial distribution. T_w is the size of the window over which back-testing is conducted. For the IND property, the exceptions produced on day $t - 1$ should be independent of exceptions produced on day t and evenly spread over time.

In this study, we use several back-testing methods to test the accuracy of the proposed VaR models. The most common are the Kupiec’s proportion of failures (POF) test for the UC [65], Christoffersen’s test for the UC and IND [66], Engle and Manganelli’s Dynamic Quantile (DQ) test [67], and Santos and Alves’ new class of independence test [68]. We also consider the Basel *traffic light* test proposed by the Basel Committee on Banking and Supervision (BCBS) [69].

Kupiec defined an approximate 95% confidence region whereby the number of exceptions produced by the VaR model must lie within this interval for it to be considered a reliable risk measurement model. The test is based on the likelihood ratio

$$LR_{UC} = -2 \ln \frac{q^{T_0}(1-q)^{T_1}}{\left(1 - \frac{T_1}{T_w}\right)^{T_0} \left(\frac{T_1}{T_w}\right)^{T_1}} \approx \chi_1^2, \tag{49}$$

where $T_0 = T_w - T_1$. Under the UC, the null hypothesis for LR_{POF} is $H_0 : E[\mathbb{I}_t(1-q)] = \frac{T_1}{T_w} = 1-q$ against $H_a : E[\mathbb{I}_t(1-q)] = \frac{T_1}{T_w} \neq (1-q)$. The VaR model is rejected if $LR_{POF} > \chi_1^2 = 3.841$.

A study by [66] extended Kupiec’s POF test to test the independence of conditional coverage. Under the null hypothesis that the number of exceptions produced are independent and evenly spread over time, $\pi_{01} = \pi_{11} = \pi$ with likelihood ratio

$$LR_{IND} = -2 \ln \frac{(1-\pi)^{(T_{00}+T_{10})} \pi^{(T_{01}+T_{11})}}{(1-\pi_{01})^{T_{00}} \pi_{01}^{T_{01}} (1-\pi_{11})^{T_{10}} \pi_{11}^{T_{11}}} \approx \chi_1^2, \tag{50}$$

where T_{ij} , with $i, j = 0(\text{noviolation}), 1(\text{violation})$, is the number of observed events with the j^{th} event following i^{th} , and π_{01}, π_{01} and π are estimates of the probabilities of T_{ij} [70]. The model is rejected for the independent property if $LR_{IND} > \chi_1^2 = 3.841$. Christoffersen conditional coverage test is a joint test of Kupiec’s POF test and the IND that test both properties of UC and IND instantaneously. The conditional coverage test has a likelihood ratio

$$LR_{CC} = LR_{POF} + LR_{IND} \approx \chi_2^2. \tag{51}$$

The hypothesis is $\Pr[\mathbb{I}_t(1-q) = 1 | \Omega_{t-1}] = 1-q, \forall t$ against $\Pr[\mathbb{I}_t(1-q) = 1 | \Omega_{t-1}] \neq 1-q, \forall t$, where Ω_{t-1} is the information available on day $t-1$. The model is rejected for the conditional coverage property if $LR_{CC} > \chi_2^2 = 5.99$.

The DQ test utilises the criterion that the number of exceptions produced on day t should be independent of the information available at day $t-1$. The function is defined as

$$Hit_t = \mathbb{I}(L_t < -VaR_{q,t}^p) - (1-q) = \begin{cases} q, & \text{if } L_t < VaR_{q,t}^p \\ -(1-q), & \text{otherwise.} \end{cases} \tag{52}$$

The Hit_t function assumes the value q when the loss on the portfolio at time t is less than $VaR_{q,t}^p$, and $-(1-q)$ otherwise. As explained in [67], clearly $E[Hit_t] = 0, E[Hit_t | \Omega_{t-1}] = 0$ and Hit_t must be uncorrelated with its own lagged values. The test statistics is given by

$$DQ = \frac{(Hit_t' X_t [X_t' X_t]^{-1} X_t' Hit_t)}{(1-q)q}, \tag{53}$$

where X_t is a vector containing all values of $Hit_t, VaR_{q,t}^p$ and its lags. Under the null hypothesis $E[Hit_t] = 0$ and $E[Hit_t | \Omega_{t-1}] = 0, Hit_t$ and X_t are orthogonal and Hit_t must be uncorrelated with its own lagged values [67, 71]. The DQ test is easy to perform, and does not depend on the

estimation procedure; all that is needed is a series of VaRs and the corresponding values of the portfolio returns [67]. In this study, we follow [67, 72, 73] to use a constant, four lagged values of Hit_t .

In Santos and Alves' new class of independence test [68], we first define the duration between two consecutive exceptions as $D_i = t_i - t_{i-1}$, where t_i denotes the time of exception number i ; and $t_0 = 0$ implies that D_1 is the time until the first exception. We denote a sequence of N durations by $\{D_i\}_{i=1}^N$, where the order statistics are $D_{1:N} \leq \dots \leq D_{N:N}$. The test statistics is defined as

$$T_{N,[N/2]} = \log 2 \frac{D_{N:N} - 1}{D_{[N/2]:N}} - \log N. \tag{54}$$

See [68] for more details on this test.

Finally, BCBS developed a set of requirements that the VaR model must satisfy for it to be considered a reliable risk measure and proposed the Basel *traffic light* test. That is, (i) VaR must be calculated with 99% confidence, (ii) back-testing must be done using a minimum of a one year observation period and must be tested over at least 250 days, (iii) regulators should be 95% confident that they are not erroneously rejecting a valid VaR model, and (iv) Basel specifies a one-tailed test—it is only interested in the underestimation of risk [74]. [2] summarises the acceptance region for the Basel *traffic light* approach to back-testing VaR models.

We use out-of-sample data of $m = T - n$ observations for back-testing; thus we have $n = 1869$ sample of the return observations for VaR estimation procedure containing the 2008 global financial crisis period, and $m = 1000$ of return observations for back-testing. VaR is then estimated following a rolling window approach. The out-of-sample data is further divided into blocks of 250, 500, and 1000 trading days to observe how the models behave for both longer and shorter observation periods. The division of out-of-sample data is also employed to meet the BCBS requirements. Table 10 presents the expected and observed number of exceptions produced following each model for a portfolio consisting of all five banks. At 99% confidence level and 250 trading days, the MS-GJR-GARCH(1,1) copula EVT VaR model registered 3 exceptions for a single-state and 0 exceptions for a two-state MS-GJR-GARCH (1,1) model. Thus, following Basel rules for back-testing, the VaR models passed the reliability test and are placed in the green zone. Back-testing results based on LR_{UC} , LR_{IND} , LR_{CC} , DQ ,

Table 10. Expected versus observed number of exceptions following Bayesian MS-GJR-GARCH(1,1) and GJR-GARCH(1,1) copula-EVT VaR model.

		250		500		1000	
		1%	5%	1%	5%	1%	5%
GJR-GARCH(1,1)	Expected exceptions	2.5	12.5	5	25	10	50
	Observed exceptions for t -copula	3	11	4	26	8	57
	Coverage rate	0.012	0.044	0.008	0.052	0.008	0.057
	Observed exceptions for Frank copula	3	11	5	24	9	55
MS-GJR-GARCH(1,1)	Coverage rate	0.012	0.044	0.010	0.048	0.009	0.055
	Observed exceptions for t -copula	0	6	0	15	0	33
	Coverage rate	0.000	0.024	0.000	0.030	0.000	0.033
	Observed exceptions for Frank copula	0	6	0	14	0	32
	Coverage rate	0.000	0.024	0.000	0.028	0.000	0.032

Note: Out-of-sample data is divided into blocks of 250, 500, and 1000 observation periods, time horizon of 1 day. The coverage rate $\frac{T_i}{T_w} \approx 1 - q$.

<https://doi.org/10.1371/journal.pone.0198753.t010>

Table 11. Back-testing results following Bayesian GJR-GARCH(1,1) Student's-*t* and Frank copula-EVT VaR models.

Student's- <i>t</i> copula			Back-test type				
Prob	Window	Exceptions	LR_{UC}	LR_{IND}	LR_{CC}	DQ	$T_{N,[N/2]}$
1%:	250	3	0.095 (0.758)	0.170 (0.680)	0.265 (0.876)	0.213 (0.999)	-0.072 (0.966)
	500	4	0.217 (0.641)	0.140 (0.708)	0.357 (0.837)	0.415 (0.998)	-0.616 (0.994)
	1000	8	0.434 (0.510)	0.268 (0.605)	0.702 (0.704)	1.057 (0.983)	1.386 (0.385)
5%:	250	11	0.197 (0.657)	2.475 (0.116)	2.672 (0.263)	2.057 (0.914)	2.108 (0.348)
	500	26	0.042 (0.838)	1.997 (0.158)	2.039 (0.361)	2.962 (0.814)	-0.254 (0.742)
	1000	57	0.989 (0.320)	3.145 (0.076)	4.134 (0.127)	7.494 (0.278)	-1.039 (0.939)
Frank copula			Back-test type				
Prob	Window	Exceptions	LR_{UC}	LR_{IND}	LR_{CC}	DQ	$T_{N,[N/2]}$
1%:	250	3	0.095 (0.758)	0.168 (0.682)	0.263 (0.877)	0.213 (0.999)	-0.072 (0.967)
	500	5	0.000 (1.000)	0.219 (0.640)	0.219 (0.896)	0.415 (0.998)	-0.839 (0.999)
	1000	9	0.105 (0.746)	0.340 (0.560)	0.445 (0.801)	1.066 (0.983)	0.537 (0.676)
5%:	250	11	0.197 (0.657)	1.887 (0.170)	2.084 (0.353)	2.057 (0.914)	2.426 (0.244)
	500	24	0.043 (0.836)	1.343 (0.247)	1.386 (0.500)	2.305 (0.890)	-0.174 (0.723)
	1000	55	0.510 (0.475)	2.531 (0.112)	3.041 (0.219)	6.335 (0.387)	-1.004 (0.934)

Note: *p*-values in parenthesis. Four lags were used for the *DQ* test.

<https://doi.org/10.1371/journal.pone.0198753.t011>

and $T_{N,[N/2]}$ tests are presented in Tables 11 and 12. For the *DQ* test, we use a lagged value of 4. In Tables 13 and 14 we present, as a benchmark for our VaR models, back-testing results for VaR models constructed using GJR-GARCH(1,1) and standard GARCH(1,1) (sGARCH(1,1)) volatility models with skewed Student's-*t* errors but without copula functions and EVT. It can be seen from the number of exceptions recorded that the MS-GJR-GARCH(1,1) copula-EVT VaR model and the benchmark VaR models do not underestimate risk but rather too conservative at 99% and 95% confidence levels and thus preferred by most financial institutions.

Table 12. Back-testing results following Bayesian MS-GJR-GARCH(1,1) Student's-*t* and Frank copula-EVT VaR models.

Student's- <i>t</i> copula			Back-test type				
Prob	Window	Exceptions	LR_{UC}	LR_{IND}	LR_{CC}	DQ	$T_{N,[N/2]}$
1%:	250	0	NaN	-	-	0.213 (0.999)	-
	500	0	NaN	-	-	0.415 (0.998)	-
	1000	0	NaN	-	-	1.057 (0.983)	-
5%:	250	6	4.369 (0.037)	0.641 (0.423)	5.010 (0.082)	1.527 (0.958)	3.069 (0.379)
	500	15	4.884 (0.027)	0.001 (0.975)	4.885 (0.087)	2.226 (0.898)	1.154 (0.474)
	1000	33	6.878 (0.009)	0.032 (0.858)	6.910 (0.032)	0.697 (0.995)	-0.793 (0.908)
Frank copula			Back-test type				
Prob	Window	Exceptions	LR_{UC}	LR_{IND}	LR_{CC}	DQ	$T_{N,[N/2]}$
1%:	250	0	NaN	-	-	0.213 (0.999)	-
	500	0	NaN	-	-	0.415 (0.998)	-
	1000	0	NaN	-	-	1.057 (0.983)	-
5%:	250	6	4.369 (0.037)	0.699 (0.403)	5.068 (0.079)	30.724 (0.000)	2.070 (0.307)
	500	14	6.018 (0.014)	0.032 (0.858)	6.050 (0.049)	13.819 (0.032)	1.223 (0.382)
	1000	32	7.777 (0.005)	0.007 (0.933)	7.784 (0.020)	7.321 (0.292)	-1.386 (0.979)

Note: *p*-values in parenthesis, NaN = Not a Number. Four lags were used for the *DQ* test.

<https://doi.org/10.1371/journal.pone.0198753.t012>

Table 13. Expected versus observed number of exceptions following sGARCH(1,1) and GJR-GARCH(1,1) models with skewed student's-t errors.

	250		500		1000	
	1%	5%	1%	5%	1%	5%
Expected exceptions	2.5	12.5	5	25	10	50
Observed exceptions sGARCH(1,1)	0	7	0	16	0	32
Coverage rate	0.000	0.028	0.000	0.032	0.000	0.032
Observed exceptions GJR-GARCH(1,1)	0	7	0	14	0	26
Coverage rate	0.000	0.028	0.000	0.028	0.000	0.026

Note: Out-of-sample data is divided into blocks of 250, 500, and 1000 observation periods, time horizon of 1 day. The coverage rate $\frac{T_1}{T_w} \approx 1 - q$.

<https://doi.org/10.1371/journal.pone.0198753.t013>

Table 14. Back-testing results following sGARCH(1,1) and GJR-GARCH(1,1) models with skewed student's-t errors.

sGARCH			Back-test type				
Prob	Window	Exceptions	LR_{UC}	LR_{IND}	LR_{CC}	DQ	$T_{N,[N/2]}$
1%:	250	0	NaN	-	-	37.933 (0.000)	-
	500	0	NaN	-	-	18.530 (0.005)	-
	1000	0	NaN	-	-	11.132 (0.084)	-
5%:	250	7	3.009 (0.083)	0.962 (0.327)	3.971 (0.137)	3.888 (0.692)	1.916 (0.448)
	500	16	3.888 (0.049)	0.011 (0.916)	3.899 (0.142)	2.409 (0.879)	-0.069 (0.708)
	1000	32	7.777 (0.005)	0.007 (0.933)	7.784 (0.020)	5.486 (0.483)	0.416 (0.541)
GJR-GARCH(1,1)			Back-test type				
Prob	Window	Exceptions	LR_{UC}	LR_{IND}	LR_{CC}	DQ	$T_{N,[N/2]}$
1%:	250	0	NaN	-	-	23.077 (0.001)	-
	500	0	NaN	-	-	48.651 (0.000)	-
	1000	0	NaN	-	-	95.686 (0.000)	-
5%:	250	7	3.009 (0.083)	0.962 (0.327)	3.971 (0.137)	57.759 (0.000)	4.119 (0.228)
	500	14	6.018 (0.014)	0.032 (0.858)	6.050 (0.049)	85.830 (0.000)	-0.773 (0.908)
	1000	26	14.597 (0.000)	0.263 (0.608)	14.860 (0.001)	169.533 (0.000)	-0.975 (0.928)

Note: *p*-values in parenthesis, NaN = Not a Number. Four lags were used for the *DQ* test.

<https://doi.org/10.1371/journal.pone.0198753.t014>

Most financial institutions will prefer VaR models with zero or very few exceptions as they routinely produce plots of *P&L* that show no violation of their 99% confidence VaR models over long periods stating that this supports their risk models. “The amount of economic capital banks currently hold is in excess of their regulatory capital. As a result, banks may prefer to report higher VaR numbers to avoid the possibility regulatory intrusion” [2]. Kupiec’s UC test (LR_{UC}) will reject VaR models that produce 0 exceptions. The GJR-GARCH(1,1) copula-EVT VaR model captures VaR quite well in periods of calm and in periods of crisis for short and long observation periods. It does not overestimate or underestimate the level of risk on the portfolio and is considered reliable as a measure of risk. Performance evaluation for rejection or acceptance of the VaR models based on 5% significance level are presented in Table 15. Clearly, the MS-GJR-GARCH(1,1) and GJR-GARCH(1,1) copula-EVT VaR models incorporating the *hybrid* method for threshold selection perform better than the standard GARCH (1,1) and GJR-GARCH(1,1) VaR models.

Table 15. Performance Evaluation based on 5% significance level.

$P = 1\%$		Back-test type				
VaR model	Window	LR_{UC}	LR_{IND}	LR_{CC}	DQ	$T_{N,[N/2]}$
GJR-GARCH(1,1) Student's- t copula-EVT	250	A (0.758)	A (0.680)	A (0.876)	A (0.999)	A (0.966)
	500	A (0.641)	A (0.708)	A (0.837)	A (0.998)	A (0.994)
	1000	A (0.510)	A (0.605)	A (0.704)	A (0.983)	A (0.385)
GJR-GARCH(1,1) Frank copula-EVT	250	A (0.758)	A (0.682)	A (0.877)	A (0.999)	A (0.967)
	500	A (1.000)	A (0.640)	A (0.896)	A (0.998)	A (0.999)
	1000	A (0.320)	A (0.076)	A (0.127)	A (0.278)	A (0.939)
MS-GJR-GARCH (1,1) Student's- t copula-EVT	250	R (NaN)	R (-)	R (-)	A (0.999)	R (-)
	500	R (NaN)	R (-)	R (-)	A (0.998)	R (-)
	1000	R (NaN)	R (-)	R (-)	A (0.983)	R (-)
MS-GJR-GARCH(1,1) Frank copula-EVT	250	R (NaN)	R (-)	R (-)	A (0.999)	R (-)
	500	R (NaN)	R (-)	R (-)	A (0.998)	R (-)
	1000	R (NaN)	R (-)	R (-)	A (0.983)	R (-)
sGARCH(1,1)	250	R (NaN)	R (-)	R (-)	R (0.000)	R (-)
	500	R (NaN)	R (-)	R (-)	R (0.005)	R (-)
	1000	R (NaN)	R (-)	R (-)	A (0.084)	R (-)
GJR-GARCH(1,1)	250	R (NaN)	R (-)	R (-)	R (0.001)	R (-)
	500	R (NaN)	R (-)	R (-)	R (0.000)	R (-)
	1000	R (NaN)	R (-)	R (-)	R (0.000)	R (-)
$P = 5\%$		Back-test type				
VaR model	Window	LR_{UC}	LR_{IND}	LR_{CC}	DQ	$T_{N,[N/2]}$
GJR-GARCH(1,1) Student's- t copula-EVT	250	A(0.657)	A(0.116)	A(0.263)	A(0.914)	A(0.348)
	500	A (0.838)	A (0.158)	A (0.361)	A (0.814)	A (0.742)
	1000	A (0.320)	A (0.076)	A (0.127)	A (0.278)	A (0.939)
GJR-GARCH(1,1) Frank copula-EVT	250	A (0.657)	A (0.170)	A (0.353)	A (0.914)	A (0.244)
	500	A (0.836)	A (0.247)	A (0.500)	A (0.890)	A (0.723)
	1000	A (0.475)	A (0.112)	A (0.219)	A (0.387)	A (0.934)
MS-GJR-GARCH(1,1) Student's- t copula-EVT	250	R (0.037)	A (0.423)	A (0.082)	A (0.958)	A (0.379)
	500	R (0.027)	A (0.975)	A (0.087)	A (0.898)	A (0.474)
	1000	R (0.009)	A (0.858)	R (0.032)	A (0.995)	A (0.908)
MS-GJR-GARCH(1,1) Frank copula-EVT	250	R (0.037)	A (0.403)	A (0.079)	R (0.000)	A (0.307)
	500	R (0.014)	A (0.858)	R (0.049)	R (0.032)	A (0.382)
	1000	R (0.005)	A (0.933)	R (0.020)	A (0.292)	A (0.979)
sGARCH(1,1)	250	A (0.083)	A (0.327)	A (0.137)	A (0.692)	A (0.448)
	500	R (0.049)	A (0.916)	A (0.142)	A (0.879)	A (0.708)
	1000	R (0.005)	A (0.933)	R (0.020)	A (0.483)	A (0.541)
GJR-GARCH(1,1)	250	A (0.083)	A (0.327)	A (0.137)	R (0.000)	A (0.228)
	500	R (0.014)	A (0.858)	R (0.049)	R (0.000)	A (0.908)
	1000	R (0.000)	A (0.608)	R (0.001)	R (0.000)	A (0.928)

Note: A = Accept, R = Reject, p -values in parenthesis. NaN = Not a Number; as a result of zero exceptions.

<https://doi.org/10.1371/journal.pone.0198753.t015>

Conclusion

In recent decades, VaR has become the most common risk measure used by financial institutions to assess market risk of financial assets. Since VaR models often focus on the behavior of asset returns in the left tail, it is important that the models are calibrated such that they do not underestimate or overestimate the proportion of outliers, as this will have significant effects on the allocation of economic capital for investments. Due to the *extremistan* [75] nature of financial asset returns and volatility, the real tail risk of a financial asset is not stable as time passes, and the maximum loss is difficult to predict. Therefore, to implement a reliable VaR model, the time horizon and type of volatility model used is very important. We constructed our VaR models by combining a single-state and two-state Bayesian MS-GJR-GARCH(1,1) models with skewed Student's- t distributions for the underlying volatility model, copula functions to model dependence, and EVT to model the left tail. The single-state Bayesian MS-GJR-GARCH(1,1) is a GJR-GARCH(1,1) model without regime change, hence the names Bayesian GJR-GARCH(1,1) copula-EVT VaR model for the single-state MS-GJR-GARCH(1,1) and Bayesian MS-GJR-GARCH(1,1) copula-EVT VaR model for the two-state MS-GJR-GARCH(1,1).

We use as a benchmark, VaR models constructed using GJR-GARCH(1,1) and sGARCH(1,1) volatility models with skewed Student's- t distributions, but without copula functions and EVT to compare the performance of our VaR models. Back-testing results show that the GJR-GARCH(1,1) copula-EVT VaR model is much reliable than the MS-GJR-GARCH(1,1) copula EVT VaR model and the benchmark VaR models. Back-testing results further indicates that the GJR-GARCH(1,1) copula EVT VaR model does not overestimate or underestimate the level of risk on the portfolio whereas the two-state MS-GJR-GARCH(1,1) copula EVT VaR model and the benchmark VaR models seems to overestimate the level of risk.

It is also important to draw attention to the fact that Eq (34) is a point estimate with an error band that becomes larger as we move to more extreme quantiles. It is concerned only with the number of exceedances above a certain threshold and is not affected by data outside the tail of the distribution [8]. This can be problematic in some cases due to limited data points in the tail, which can inhibit proper analysis.

Eq (34) also depends on the threshold and the number of points (i.e., exceedances) above the threshold because the parameters are estimated based on the exceedances. Thus, it is logical to say that the reliability of Eq (34) rests solely on the choice of thresholds, which is subjective. The proposed *hybrid* method for the threshold addresses this issue and diminishes the possibility of selecting a less suitable threshold value. This method is reliable and can be implemented with other conditional multivariate volatility models providing positive-definite volatility matrices.

Acknowledgments

The authors thank the Academic Editor and the anonymous reviewers for their constructive comments and suggestions that have significantly improved this paper. We thank Dr. Spyridon Vrontos for having taken the time to critically read our manuscript.

Author Contributions

Conceptualization: Haslifah M. Hasim, Hongsheng Dai.

Data curation: Marius Galabe Sampid, Haslifah M. Hasim.

Formal analysis: Marius Galabe Sampid.

Investigation: Marius Galabe Sampid, Haslifah M. Hasim.

Methodology: Marius Galabe Sampid, Haslifah M. Hasim, Hongsheng Dai.

Project administration: Haslifah M. Hasim.

Resources: Marius Galabe Sampid.

Software: Marius Galabe Sampid.

Supervision: Haslifah M. Hasim, Hongsheng Dai.

Validation: Marius Galabe Sampid, Haslifah M. Hasim, Hongsheng Dai.

Visualization: Marius Galabe Sampid, Haslifah M. Hasim, Hongsheng Dai.

Writing – original draft: Marius Galabe Sampid.

Writing – review & editing: Marius Galabe Sampid, Haslifah M. Hasim.

References

1. Holton GA. Value-at-risk: Theory and practice. Vol. 2. New York: Academic Press; 2003.
2. Jorion P. Value at risk: The new benchmark for managing financial risk. Vol. 3. New York: McGraw-Hill; 2007.
3. Malz AM. Financial risk management: Models, history, and institutions. Vol. 538. New Jersey: John Wiley & Sons; 2011.
4. Soltane HB, Karaa A, Bellalah M. Conditional VaR using GARCH-EVT approach: Forecasting volatility in Tunisian financial market. *Journal of Computations and Modelling*. 2012; 2(2):95–115.
5. Berkowitz J, Christoffersen P, Pelletier D. Evaluating value-at-risk models with desk-level data. *Management Science*. 2011; 57(12):2213–2227. <https://doi.org/10.1287/mnsc.1080.0964>
6. Sheikh AZ, Qiao H. Non-normality of market returns: A framework for asset allocation decision making. *Journal of Alternative Investments*. 2010; 12(3):8–35. <https://doi.org/10.3905/JAI.2010.12.3.008>
7. Poon SH, Rockinger M, Tawn J. Modelling extreme-value dependence in international stock markets. *Statistica Sinica*. 2003; p. 929–953.
8. Wong MC. Bubble value at risk: A countercyclical risk management approach. Singapore: John Wiley & Sons; 2013.
9. McNeil AJ, Frey R. Estimation of tail-related risk measures for heteroscedastic financial time series: An extreme value approach. *Journal of Empirical Finance*. 2000; 7(3):271–300. [https://doi.org/10.1016/S0927-5398\(00\)00012-8](https://doi.org/10.1016/S0927-5398(00)00012-8)
10. Bauwens L, De Backer B, Dufays A. A Bayesian method of change-point estimation with recurrent regimes: Application to GARCH models. *Journal of Empirical Finance*. 2014; 29:207–229. <https://doi.org/10.1016/j.jempfin.2014.06.008>
11. Haas M, Mittnik S, Paoletta MS. A new approach to Markov-switching GARCH models. *Journal of Financial Econometrics*. 2004; 2(4):493–530. <https://doi.org/10.1093/jfinec/nbh020>
12. Gray SF. Modeling the conditional distribution of interest rates as a regime-switching process. *Journal of Financial Economics*. 1996; 42(1):27–62. [https://doi.org/10.1016/0304-405X\(96\)00875-6](https://doi.org/10.1016/0304-405X(96)00875-6)
13. Klaassen F. Improving GARCH volatility forecasts with regime-switching GARCH. In: *Advances in Markov-Switching Models*. Springer; 2002. p. 223–254.
14. Ardia D, Bluteau K, Boudt K, Trottier DA. Markov-switching GARCH models in R; 2017: Available from: <https://cran.r-project.org/package=MSGARCH>
15. Hsu CP, Huang CW, Chiou WJP. Effectiveness of copula-extreme value theory in estimating value-at-risk: Empirical evidence from Asian emerging markets. *Review of Quantitative Finance and Accounting*. 2012; 39(4):447–468. <https://doi.org/10.1007/s11156-011-0261-0>
16. Bob NK. Value at risk estimation. A GARCH-EVT-copula approach. *Mathematiska institutionen*. 2013; p. 1–41.
17. Chen CWS, So MKP, Lin EMH. Volatility forecasting with double Markov switching GARCH models. *Journal of Forecasting*. 2009; 28(8):681–697. <https://doi.org/10.1002/for.1119>
18. Nelson DB. Conditional heteroskedasticity in asset returns: A new approach. *Econometrica: Journal of the Econometric Society*. 1991; p. 347–370. <https://doi.org/10.2307/2938260>

19. Zakoian JM. Threshold heteroskedastic models. *Journal of Economic Dynamics and Control*. 1994; 18(5):931–955. [https://doi.org/10.1016/0165-1889\(94\)90039-6](https://doi.org/10.1016/0165-1889(94)90039-6)
20. Glosten LR, Jagannathan R, Runkle DE. On the relation between the expected value and the volatility of the nominal excess return on stocks. *The Journal of Finance*. 1993; 48(5):1779–1801. <https://doi.org/10.1111/j.1540-6261.1993.tb05128.x>
21. Ali G. EGARCH, GJR-GARCH, TGARCH, AVGARCH, NGARCH, IGARCH and APARCH models for pathogens at marine recreational sites. *Journal of Statistical and Econometric Methods*. 2013; 2(3):57–73.
22. Ardia D. *Financial risk management with Bayesian estimation of GARCH models: Theory and applications*. Berlin: Springer-Verlag; 2008.
23. Chen CWS, Weng M, Watanabe T. Bayesian forecasting of Value-at-Risk based on variant smooth transition heteroskedastic models. *Statistics and Its Interface*. 2017; 10(3):451–470. <https://doi.org/10.4310/SII.2017.v10.n3.a9>
24. Chen CWS, Gerlach R, Lin EM, Lee W. Bayesian forecasting for financial risk management, pre and post the global financial crisis. *Journal of Forecasting*. 2012; 31(8):661–687. <https://doi.org/10.1002/for.1237>
25. Fernández C, Steel MF. On Bayesian modeling of fat tails and skewness. *Journal of the American Statistical Association*. 1998; 93(441):359–371. <https://doi.org/10.1080/01621459.1998.10474117>
26. Trottier DA, Ardia D. Moments of standardized Fernandez–Steel skewed distributions: Applications to the estimation of GARCH-type models. *Finance Research Letters*. 2016; 18:311–316. <https://doi.org/10.1016/j.frl.2016.05.006>
27. Virbickaite A, Ausín MC, Galeano P. Bayesian inference methods for univariate and multivariate GARCH models: A survey. *Journal of Economic Surveys*. 2015; 29(1):76–96. <https://doi.org/10.1111/joes.12046>
28. Hall P, Yao Q. Inference in ARCH and GARCH models with heavy-tailed errors. *Econometrica*. 2003; 71(1):285–317. <https://doi.org/10.1111/1468-0262.00396>
29. David D, Hoogerheide L. Bayesian estimation of the GARCH (1, 1) model with student-t innovations. *The R Journal*. 2010; 2(2):41–47.
30. Greenberg E. *Introduction to Bayesian econometrics*. New York: Cambridge University Press; 2013.
31. Tierney L. Markov chains for exploring posterior distributions. *Annals of Statistics*. 1994; p. 1701–1728.
32. Roberts GO, Smith AF. Simple conditions for the convergence of the Gibbs sampler and Metropolis-Hastings algorithms. *Stochastic Processes and their Applications*. 1994; 49(2):207–216. [https://doi.org/10.1016/0304-4149\(94\)90134-1](https://doi.org/10.1016/0304-4149(94)90134-1)
33. Casella G, George EI. Explaining the Gibbs sampler. *The American Statistician*. 1992; 46(3):167–174. <https://doi.org/10.1080/00031305.1992.10475878>
34. Sklar M. Fonctions de repartition à n dimensions et leurs marges. *Publications de l'Institut de Statistique de l'Université de Paris*. 1959; 8:229–231.
35. Frey R, McNeil AJ. Dependent defaults in models of portfolio credit risk. *Journal of Risk*. 2003; 6:59–92. <https://doi.org/10.21314/JOR.2003.089>
36. Li DX. On default correlation: A copula function approach. *The Journal of Fixed Income*. 2000; 9(4):43–54. <https://doi.org/10.3905/jfi.2000.319253>
37. Embrechts P, McNeil A, Straumann D. Correlation and dependence in risk management: Properties and pitfalls. *Risk management: Value at risk and beyond*. 2002;176223.
38. Patton AJ. On the out-of-sample importance of skewness and asymmetric dependence for asset allocation. *Journal of Financial Econometrics*. 2004; 2(1):130–168. <https://doi.org/10.1093/jfinec/nbh006>
39. Cherubini U, Luciano E. Value-at-risk trade-off and capital allocation with copulas. *Economic Notes*. 2001; 30(2):235–256. <https://doi.org/10.1111/j.0391-5026.2001.00055.x>
40. Embrechts P, Lindskog F, McNeil A. Modelling dependence with copulas. *Rapport technique, Département de mathématiques, Institut Fédéral de Technologie de Zurich, Zurich*. 2001.
41. Cherubini U, Luciano E, Vecchiato W. *Copula methods in finance*. Chichester: John Wiley & Sons; 2004.
42. Silva Filho OC, Ziegelmann FA, Dueker MJ. Assessing dependence between financial market indexes using conditional time-varying copulas: Applications to Value at Risk (VaR). *Quantitative Finance*. 2014; 14(12):2155–2170. <https://doi.org/10.1080/14697688.2012.739726>
43. Huang JJ, Lee KJ, Liang H, Lin WF. Estimating value at risk of portfolio by conditional copula-GARCH method. *Insurance: Mathematics and Economics*. 2009; 45(3):315–324.

44. Fantazzini D. Dynamic copula modelling for value at risk. *Frontiers in Finance and Economics*. 2008; 5(2):72–108.
45. Tsay RS. *Multivariate Time Series Analysis: With R and Financial Applications*. New Jersey: John Wiley & Sons; 2013.
46. Ghalanos A. The rmgarch models: Background and properties. Available from: <http://cran-r-project.org/web/packages/rmgarch/index.html>
47. Cherubini U, Mulinacci S, Gobbi F, Romagnoli S. *Dynamic copula methods in finance*. Vol. 625. Chichester: John Wiley & Sons; 2011.
48. Gumbel EJ. Bivariate exponential distributions. *Journal of the American Statistical Association*. 1960; 55(292):698–707. <https://doi.org/10.1080/01621459.1960.10483368>
49. Clayton DG. A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence. *Biometrika*. 1978; 65(1):141–151. <https://doi.org/10.1093/biomet/65.1.141>
50. Frank MJ. On the simultaneous associativity of $F(x, y)$ and $x + y - F(x, y)$. *Aequationes mathematicae*. 1979; 19(1):194–226. <https://doi.org/10.1007/BF02189866>
51. Yan J. Enjoy the joy of copulas: With a package copula. *Journal of Statistical Software*. 2007; 21(4):1–21. <https://doi.org/10.18637/jss.v021.i04>
52. Breymann W, Dias A, Embrechts P. *Dependence structures for multivariate high-frequency data in finance*. ETHZ, Zürich, Switzerland. 2003;
53. Alexander C. *Market Risk Analysis, Practical Financial Econometrics*. Vol. II. Chichester: John Wiley & Sons; 2008.
54. Nelsen RB. *An introduction to copulas*. Lecture Notes in Statistics. New York: Springer Science & Business Media; 2007.
55. Tsay RS. *An introduction to analysis of financial data with R*. New Jersey: John Wiley & Sons; 2014.
56. Bhattacharyya M, Ritolia G. Conditional VaR using EVT—Towards a planned margin scheme. *International Review of Financial Analysis*. 2008; 17(2):382–395. <https://doi.org/10.1016/j.irfa.2006.08.004>
57. Danielsson J, De Vries CG. Beyond the sample: Extreme quantile and probability estimation. Tinbergen Institute Discussion Paper; 1998.
58. Danielsson J, de Haan L, Peng L, de Vries CG. Using a bootstrap method to choose the sample fraction in tail index estimation. *Journal of Multivariate Analysis*. 2001; 76(2):226–248. <https://doi.org/10.1006/jmva.2000.1903>
59. Hill BM. A simple general approach to inference about the tail of a distribution. *Annals of Statistics*. 1975; 3(5):1163–1174. <https://doi.org/10.1214/aos/1176343247>
60. Davison AC, Smith RL. Models for exceedances over high thresholds. *Journal of the Royal Statistical Society Series B (Methodological)*. 1990; p. 393–442.
61. Tsay RS. *Analysis of financial time series*. 3rd ed. New York: John Wiley & Sons; 2005.
62. Graves S. *FinTS: Companion to Tsay (2005) analysis of financial time series*; 2014. Available from: <http://CRAN.R-project.org/package=FinTS>
63. Howell DC. *Statistical methods for psychology*. Belmont: Cengage Learning; 2012.
64. Croux C, Dehon C. Influence functions of the Spearman and Kendall correlation measures. *Statistical Methods and Applications*. 2010; 19(4):497–515. <https://doi.org/10.1007/s10260-010-0142-z>
65. Kupiec PH. Techniques for verifying the accuracy of risk measurement models. *The Journal of Derivatives*. 1995; 3(2):73–84. <https://doi.org/10.3905/jod.1995.407942>
66. Christoffersen PF. Evaluating interval forecasts. *International Economic Review*. 1998; p. 841–862.
67. Engle RF, Manganelli S. CAViaR: Conditional autoregressive value at risk by regression quantiles. *Journal of Business & Economic Statistics*. 2004; 22(4):367–381. <https://doi.org/10.1198/073500104000000370>
68. Santos PA, Alves MF. A new class of independence tests for interval forecasts evaluation. *Computational Statistics and Data Analysis*. 2012; 56(11):3366–3380. <https://doi.org/10.1016/j.csda.2010.10.002>
69. Goodhart C. *The Basel Committee on Banking Supervision: A history of the early years 1974–1997*. Cambridge: Cambridge University Press; 2011.
70. Argyropoulos, Christos and Panopoulou, Ekaterini. Measuring the market risk of freight rates: A forecast combination approach. *Journal of Forecasting*, 2017.

71. Gaglianone WP, Lima LR, Linton O, Smith DR. Evaluating value-at-risk models via quantile regression. *Journal of Business & Economic Statistics*. 2011; 29(1):150–160. <https://doi.org/10.1198/jbes.2010.07318>
72. Kuester K, Mittnik S, Paolella MS. Value-at-risk prediction: A comparison of alternative strategies. *Journal of Financial Econometrics*. 2006; 4(1):53–89. <https://doi.org/10.1093/jfinec/nbj002>
73. Taylor JW. Using exponentially weighted quantile regression to estimate value at risk and expected shortfall. *Journal of Financial Econometrics*. 2007; 6(3):382–406. <https://doi.org/10.1093/jfinec/nbn007>
74. Resti A. Pillar II in the New Basel Accord: The challenge of economic capital. *Risk books*; 2008.
75. Taleb NN. *The black swan: The impact of the highly improbable*. London: Macat International Limited; 2017.