Forecasting robust value-at-risk estimates: Evidence from UK banks

SAMPID MARIUS GALABE† and HASLIFAH M. HASIM *†

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

In this paper, we present a novel approach for forecasting Value-at-Risk (VaR) by combining a Bayesian GARCH(1,1) model with Student's-t distribution for the underlying volatility models, vine copula functions to model dependence, and peaks-over-threshold (POT) method of extreme value theory (EVT) to model the tail behaviour of asset returns. We further propose a new approach for threshold selection in extreme value analysis, which we call a *hybrid* method. The empirical results and back-testing analysis show that the model captures VaR quite well through periods of calmness and crisis; therefore, it is suitable for use as a measure of risk. Our results also suggest that with a correct implementation of the VaR model, Basel III is not needed.

Keywords: Value-at-risk; Risk management; Extreme value theory; GARCH; Volatility model; Vine copulas.

JEL Classification: C15, C58, G11, G32

1. Introduction

Over the past few decades, Value-at-Risk (VaR) has become an influential tool for measuring market risk as 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 returns in a single number over a target horizon.

There are many methods to estimate VaR, and the most common methods used by banks and financial institutions include the variance-covariance method, historical simulation and Monte Carlo simulation (Holton (2014), Jorion (2007), Malz (2011) and the references therein). Except for historical simulation, the conventional VaR models assume that asset returns in financial markets are normally distributed, which is a major weakness. Financial asset returns are leptokurtic and heavy tailed with nonconstant volatility (Berkowitz et al. 2011, Sheikh and Qiao 2010). Normality assumption in cases of nonnormality will lead to inaccurate estimates in the tails of the distribution and hence, the probability of extreme events, which leads to underestimation of the likelihood of extreme tail losses. To avoid this problem, researchers tend to use extreme value theory (EVT) 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 (Wong 2013). Sampid and Hasim (2018) proposed a multivariate copula-based volatility model for estimating VaR in the banking sector of some selected European countries by combining dynamic conditional correlation (DCC) multivariate GARCH (M-GARCH) volatility model and copula functions. Their results showed that the DCC M-GARCH copula-based approach captures VaR reasonably well based on

---

*Corresponding author. Email: hhashim@essex.ac.uk

the differences in the numbers of exceptions produced during different observation periods at the same confidence level.

Our work is motivated by the work of McNeil and Frey (2000) who suggest applying EVT to the noise variable of the return series, which is normally distributed, to obtain the qth 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 Soltane et al. (2012); they combined a GARCH(1,1) model as the underlying volatility model with EVT to estimate VaR and showed that the GARCH-EVT-based VaR approach appears to be more effective and realistic than the conventional VaR estimation methods. Hsu et al. (2012) and Bob (2013) also used combined GARCH-EVT and copula functions in estimating VaR. Their findings showed better performance compared to conventional VaR estimation methods, or methods that combined copulas with conventionally employed empirical distributions.

The Basel Committee on Banking and Supervision (BCBS) recognised VaR models as the official risk management tool for measuring market risk in the Basel II Accord. Basel requires banks to have appropriate risk management measures in place that include effective supervisory protocol and assets disclosure (Resti 2008). As such, banks must calculate minimum capital to face market risks based on the internal model approach (IMA), which can be determined by the banks' VaR estimates. In the aftermath of the 2008 global financial crisis and following the Basel III, the minimum capital requirements (MCR) were further increased by the introduction of a stress VaR metric (sVaR). Moreoever, cushion buffers were placed over the MCR to guarantee that banks have enough capital in periods of stress; these buffers provide pool of cash from which to draw when losses are incurred.

1.1. *The Basel Capital Accord*

Following the collapse of the Herstatt Bank in Cologne in 1974, the central bank governors of the Group of Ten countries established the Basel Committee of Banking and Supervision (BCBS), which drafted rules and regulations for the banking industry to avoid future major bank collapses (Goodhart 2011). In 1988, BCBS adopted the first Basel Accord (Basel I) by introducing minimum capital requirements (MCR) of 8% of risk-weighted assets (RWA) that must be held by banks. Basel I mainly addressed the problem of credit risk by raising deposits and lending to households and businesses (Carmassi and Micossi 2012). Interest rates and market risks were ignored as no capital requirements were defined. A banking system faced credit risk with insufficient capital to provide proper cushion capable of withstanding sudden losses in periods of huge distress.

Following criticisms from regulators and banks, Basel II was adopted in the late $1990s$ by amending Basel I to incorporate capital requirements to guard against market risks. Designed to encourage banks to engage in sensible risk-taking, Basel II allows banks to calculate MCR for market risk based on their internal 99% VaR models using the internal model approach (IMA), which is the maximum of the average within the last 60 days VaR or the most recent VaR:

$$MCR_t = max\left(\frac{k}{60}\sum_{i=1}^{60}VaR_{t-i}, VaR_{t-1}\right), \tag{1a}$$

$$k = \begin{cases} 3 & \text{if } T_1 \leq 4, \text{ green zone} \\ 3 + 0.2(T_1 - 4) & \text{if } 5 \leq T_1 \leq 9, \text{ yellow zone} \\ 4 & \text{if } T_1 \geq 10, \text{ red zone.} \end{cases} \tag{1b}$$

Basel II also introduced a back-testing procedure to validate the reliability of the banks' internal 99% VaR model from which the supervisory multiplier k is determined. Back-testing was designed to compare the subsequent VaR estimates with the actual returns and recording the number of days T_1 in which the realised losses exceeded the 99% VaR for a 250-days observation period. Basel also

requires a liquidation period of 60 days, which they believe is enough for a financial institution that is in trouble to raise funds. Eqn.(1a) is designed such that $VaR_{t-1} > \frac{k}{60} \sum_{i=1}^{60} VaR_{t-i}$ will occur only in periods of extreme crisis, such as a crash, and the internal VaR model must be validated by supervisors.

After the 2008 global financial crisis, some researchers pointed out that Basel II failed to provide proper cushions to guard against the banks' actual losses on their market risks. For example, the International Monetary Fund (IMF) published the Global Financial Stability Report (International Monetary Fund 2009) in April 2009, and revealed that during the 2008-9 financial crisis, the risk-weighted capital ratios were unable to distinguish between banks that were in distress or bailed out with taxpayer money and banks that could cope on their own (Carmassi and Micossi 2012). The Financial Services Authority of the United Kingdom stated that the assumption of a normal distribution with short-term observations could lead to a massive underestimation of the probability of extreme loss events (Turner et al. 2009). Banks that were bailed out or collapsed showed higher and improving solvency ratios in the months preceding their collapse (Rossignolo et al. 2012, Carmassi and Micossi 2012). However, McAleer et al. (2011) pointed out that Basel II was operational in Europe only starting in 2008, and the effects of the global financial crisis of 2008 could not be associated to any failings of Basel II because it was never implemented in the United States, which was the epicentre of the crisis.

In late 2010, the BCBS adopted a more practical framework by introducing Basel III that incorporated stricter measures to strengthen regulation, supervisions, risk management, transparency and disclosures in the banking sector. Under this framework, banks are still allowed to calculate their MCR for market risk using the IMA as required in Basel II; however, the BCBS introduces a stressed VaR (sVaR) metric that automatically increases MCR for market risk. In this version, the sVaR must be calculated from a data set from a continuous 12-month period of substantial financial stress (Rossignolo et al. 2012):

$$MCR_t = max\left(\frac{k}{60}\sum_{i=1}^{60} VaR_{t-i}, VaR_{t-1}\right) + max\left(\frac{k}{60}\sum_{i=1}^{60} sVaR_{t-i}, sVaR_{t-1}\right). \tag{2}$$

Evidence indicating that capital ratios higher than up to 2.5% of RWA was needed to correct internal model errors for market risk that resulted in underestimating losses, led Basel III to introduce a micro-prudential 2.5% conservation buffer (Carmassi and Micossi 2012), i.e. a cushion to protect MCR against falling below its minimum during financial distress or a crash. This buffer will lead to the restriction of dividend payouts once it falls below its minimum and reversed only when restored to its original value. Basel III also introduces a "macro-prudential countercyclical buffer" ranging from 0-2.5% and applied by the national authorities depending on the banks' credit-to-GDP ratio. Uylangco and Li (2016) assessed VaR models for Australian banks in a ten year period from 1 July 2001 to 30 June 2011. Their conclusion, based on results, supported the VaR methodology adopted under Basel II revision and the forthcoming Basel III proposal.

In this paper, we investigate the validity of VaR models, in line with Basel III, constructed using Bayesian-GARCH(1,1), vine copula functions, and EVT. We combine the Bayesian-GARCH(1,1) model with a Student-t distribution to model the volatility equation, vine copula functions to model dependence, and peaks-over-threshold (POT) method based on EVT to model the left tail, to estimate VaR in some selected banks in the United Kingdom (UK) using actively traded stocks on the London Stock Exchange. The motivation of Bayesian-GARCH(1,1) model is because Bayesian estimation method provides reliable results even for finite samples, and is usually straightforward to obtain the posterior distribution of any non-linear function of the model parameters whereas for the classical maximum likelihood method, it is not easy to perform inferences on non-linear function of the model parameters, 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

optimization procedures (Virbickaite et al. 2015, Hall and Yao 2003). The motivation of Student's-t distribution is because it is able to account for the the excess kurtosis in the conditional distribution common with financial time series processes (Ardia and Hoogerheide 2010).

The rest of the paper is structured as follows: In section 2, we discuss Bayesian statistics in relation to the GARCH(1,1) model with Student's-t distribution, the basic theory of copula and vine copulas, EVT and the POT method. In section 3, we present the data and discuss some important characteristics of the data before proceeding to empirical analysis. We propose a *hybrid* method as a new approach for the threshold selection in EVT; and the results of robust VaR estimates are presented in section 4. In section 5, we check the reliability of the VaR model by conducting back-testing, followed by the summary and conclusion in section 6.

2. Methodology

2.1. *Bayesian GARCH(1,1) model with Student's-t innovations*

The GARCH(1,1) model, first proposed by Bollerslev (1986), allows conditional variance to be dependent upon previous lags. The GARCH(1,1) model following the Student's-t distribution has the form

$$r_t = \mu_t + a_t, \quad a_t = \epsilon_t \left(\frac{v-2}{v} \omega_t h_t \right)^{1/2} \tag{3a}$$

$$h_t = \alpha_0 + \alpha_1 a_{t-1}^2 + \beta_1 h_{t-1}, \tag{3b}$$

$$\epsilon_t \overset{iid}{\sim} N(0,1); \quad \omega_t \overset{iid}{\sim} IG\left(\frac{v}{2}, \frac{v}{2} \right); \quad t = 1, \dots, T,$$

where r_t are the log-returns of daily stock prices and μ_t is the unconditional mean of the log-returns. IG and $N(0,1)$ symbolises the inverted gamma and standard normal distributions with probability density functions

$$f(\omega) = \frac{\omega^{-(\kappa+1)} e^{-\frac{1}{b}\omega}}{\Gamma(\kappa) b^{\kappa}} \quad \text{and} \quad \varphi(\epsilon) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\epsilon^2},$$

respectively, where κ is the shape parameter and b the scale parameter. The probability density function (PDF) of a standardized Student's-t distribution is defined as

$$f_s(\epsilon, \nu) = \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{(\nu-2)\pi}\Gamma(\frac{\nu}{2})} \left(1 + \frac{\epsilon^2}{\nu-2} \right)^{-\frac{\nu+1}{2}}, \quad \epsilon \in \mathbb{R}, \tag{4}$$

where $\Gamma(\cdot)$ is the Gamma function. The degrees of freedom parameter $v > 2$ guarantees finite conditional variance (Ardia 2008). Eqns.(3a) and (3b) are the mean and variance equations, respectively.

We use Bayesian statistics, following the procedures delineated in Ardia and Hoogerheide (2010) and Ardia (2015) to estimate the parameter values of the variance equation. Let $\boldsymbol{a} = (a_1, \dots, a_T)'$, $\omega = (\omega_1, \dots, \omega_T)'$ and $\alpha = (\alpha_0, \alpha_1)'$, a diagonal matrix is defined by

$$\Sigma = \Sigma(\psi, \omega) = diag\left\{ \left(\omega_t \frac{v-2}{v} h_t(\alpha, \beta_1) \right)_{t=1}^T \right\}, \tag{5}$$

where $\psi = (\alpha, \beta_1, v)$ for

$$h_t(\alpha, \beta_1) = \alpha_0 + \alpha_1 a_{t-1}^2 + \beta_1 h_{t-1}(\alpha, \beta_1). \tag{6}$$

Because the data samples are independent and drawn from a normal distribution, the likelihood function can be written as

$$L(\psi, \omega | \boldsymbol{a}) \propto (det\Sigma)^{-1/2} exp\left[-\frac{1}{2} \boldsymbol{a}' \Sigma^{-1} \boldsymbol{a} \right]. \tag{7}$$

The prior distribution of ω_t given v represented by

$$p(\omega | v) = \left(\frac{v}{2}\right)^{\frac{Tv}{2}} \left[\Gamma\left(\frac{v}{2}\right) \right]^{-T} \left(\prod_{t=1}^{T} \omega_t \right)^{-\frac{v}{2}-1} exp\left[-\frac{1}{2} \sum_{t=1}^{T} \frac{v}{\omega_t} \right], \tag{8}$$

where $\omega_t \overset{iid}{\sim} IG(\bullet)$ (e.g. Geweke (1993)). For the degrees of freedom, the prior distribution is a translated exponential distribution with parameters $\lambda > 0$ and $\delta \geq 2$ (e.g. Deschamps (2006)) represented as

$$p(v) = \lambda \exp[-\lambda(v - \delta)]1\{v > \delta\}. \tag{9}$$

Ardia and Hoogerheide (2010) and Deschamps (2006) point out two important considerations for the prior density $p(v)$: (i) It is useful to guarantee that $v \gg 2$ so the conditional variance will be finite. (ii) The error term can be assumed to be normally distributed when δ is chosen to be large. This is possible while still maintaining reasonably tight priors, which can lead to better convergence of the sampler. Assuming independence among the model parameters of the joint prior distribution (i.e. $p(\psi, \omega) = p(\alpha)p(\beta)p(\omega|v)p(v)$), the likelihood function of the model parameters is combined with the prior density to obtain the posterior density as

$$p(\psi, \omega | \boldsymbol{a}) \propto L(\psi, \omega | \boldsymbol{a})p(\psi, \omega). \tag{10}$$

We employ the Metropolis Hastings (MH) algorithm to estimate parameter values from the posterior distribution of the variance equation because the MH algorithm allows draws to be generated from any density, even if the normalising constant is unknown (Greenberg 2012).

In the MH algorithm, (ψ, ω) is a random variable with Markov chains constructed as $(\psi^{[0]}, \omega^{[0]}), \ldots, (\psi^{[j]}, \omega^{[j]}), \ldots$ in the parameter space. As the number of realised chains goes to infinity, $p(\psi, \omega | \boldsymbol{a})$ tends to a normalised probability distribution with a random variable $(\psi^{[j]}, \omega^{[j]})$ (Ardia and Hoogerheide 2010). The chain converges to its stationary distribution and the optimal mean values of the posterior distribution parameters are realised.

2.2. *Vine copulas*

The copula theory was first developed by Sklar (1959) to describe the dependence structure between random variables. It was later introduced to the finance literature by Embrechts and McNeil (1999), Li (2000) and Frey et al. (2001). Consequently, Embrechts et al. (2002) introduced the application of copula theory to financial asset returns, and Patton (2004) expanded the framework of copula theory with respect to the time-varying nature of financial dependence schemes. The copula theory has also been used in risk management to measure the VaR of portfolios, including both unconditional (Cherubini and Luciano 2001, Embrechts et al. 2001, Cherubini et al. 2004) and, recently, conditional distributions (Silva Filho et al. 2014, Huang et al. 2009, Fantazzini 2008).

Copula functions enable the construction of a flexible multivariate distribution with varying

margins and dependence structures; this allows the joint distribution of a portfolio to be free from assumptions of normality or linear correlation. Additionally, copulas can easily capture extreme dependencies such as tail dependence, while the normal distribution assumes no extreme dependencies.

In multivariate settings, we use the following version of Sklar's theorem (Sklar 1959) as given by Cherubini et al. (2004) for the purpose of VaR estimation:

THEOREM 2.1 *Sklar's theorem: Let $F_1(x_1), \ldots, F_n(x_n)$ be known marginal distribution functions. Then, for every $X = (x_1, \ldots, x_n) \in \Re^{\pm n}$, where $\Re^{\pm n} = [-\infty, +\infty]^n$.*

DEFINITION 1 An $n-$dimensional copula $C(u_1, u_2, u_3, \ldots, u_n)'$ is a distribution function on $\mathbf{I}^n$ with standard uniform marginal distributions (Tsay 2013).

Sklar's theorem states that for every n-dimensional distribution of a random vector $X = (x_1, \ldots, x_n)$ with marginals $F_1(x_1), \ldots, F_n(x_n)$ can be written as

$$F(x_1, \ldots, x_n) = C(F_1(x_1), \ldots, F_n(x_n)), \tag{11}$$

where C is uniquely determined on $\mathbf{I}^n = [0, 1]^n$ for distributions F with absolutely continuous margins

$$C(u_1, \ldots, u_n) = F(F_1^{-1}(u_1), \ldots, F_n^{-1}(u_n)). \tag{12}$$

Each copula $C(u_1, \ldots, u_n)$ has a density $c(u_1, \ldots, u_n)$ related to it and defined as

$$c(u_1, \ldots, u_n) = \frac{\partial_n C(u_1, \ldots, u_n)}{\partial u_1, \ldots, \partial u_n}, \tag{13}$$

and the density function for the copula is

$$f(x_1, \ldots, x_n) = c(F_1(x_1), \ldots, F_n(x_n)) \prod_{i=1}^{n} f_i(x_i) \tag{14}$$

in $\mathbf{I}^n$ for a continuous random variable, where f_i are the marginal densities that can be different from each other (Ghalanos 2015, Tsay 2013, Huang et al. 2009, Cherubini et al. 2004).

The most commonly used families of copulas in financial applications are the elliptical and Archimedean copulas (Cherubini et al. 2011). Elliptical copulas are derived from multivariate elliptical distributions. The two most common elliptical copulas are the Gaussian and 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. Archimedean copulas are very useful in financial applications and risk management because they capture an asymmetric tail dependences between financial asset returns. The most commonly used Archimedean copulas in financial applications are the Gumbel (1960), Clayton (1978) and Frank (1979) copulas (Yan et al. 2007).

For higher dimensions, standard multivariate copulas can become inflexible and do not allow for different dependent structures between pairs of variables (Krämer and Schepsmeier 2011). In particular, the vine copulas, which are built from a set of bivariate copulas, provide more flexibility in modelling dependence for higher dimensional distribution that allows for different dependence structures between pairs of variables. For a bivariate case, we use the following version of Sklar's theorem (Cherubini et al. 2004, Krämer and Schepsmeier 2011, Tsay 2013, Ghalanos 2015):

DEFINITION 2 A $2-$dimensional copula $C(u, v)$ is a distribution function on $\mathbf{I}^2$ with standard

uniform margins.

Let F be a joint distribution function with margins F_1 and F_2, then a copula C such that:

$$\forall (x_1, x_2) \in \mathbf{I}^2; F(x_1, x_2) = C(F_1(x_1), F_2(x_2)). \tag{15}$$

If F_1 and F_2 are continuous, then C is unique; and for every (u, v),

$$C(u, v) = F(F_1^{-1}(u), F_2^{-1}(v)) \tag{16}$$

$$c(u, v) = \frac{\partial^2 C(u, v)}{\partial u, \partial v} \tag{17}$$

$$f(x_1, x_2) = c_{uv}(F_1(x_1), F_2(x_2)) f_1(x_1) . f_2(x_2). \tag{18}$$

Each copula has a density (Eqn.(17)), joint density (Eqn.(18)), and conditional densities

$$f(x_1|x_2) = c_{uv}(F_1(x_1), F_2(x_2)) f_1(x_1) \tag{19a}$$

$$f(x_2|x_1) = c_{uv}(F_1(x_1), F_2(x_2)) f_2(x_2), \tag{19b}$$

where c_{uv} is the pair-copula density for $F_1(x_1)$ and $F_2(x_2)$. For a $d-$dimensional vector $\mathbf{v}$, we have

$$f(x_1, \ldots, x_d) = c_{1\ldots d}(F_1(x_1), \ldots, F_d(x_d)) . f_1(x_1) \ldots f_d(x_d) \tag{20}$$

$$f(x|\mathbf{v}) = c_{x v_j | \mathbf{v}_{-j}}(F(x|\mathbf{v}_{-j}), F(v_j|\mathbf{v}_{-j})) . f(x|\mathbf{v}_{-j}), \tag{21}$$

where Eqn.(20) is the joint density function, and Eqn.(21) is a general formula for the pair-copula multiplied by their conditional marginal densities (Aas et al. 2009). Bedford and Cooke (2001) presented a tree diagram for selecting the possible pair-copula constructions referred to as the *regular vine structure*. This includes the *canonical vines* (C-vines), where each tree has a unique node connected to all other nodes, and the *drawable vines* (D-vines), where each tree is a path (Krämer and Schepsmeier 2011). In particular, for a D-vine, $f(x_1, \ldots, x_d)$ is given by

$$f(x_1, \ldots, x_d) = \prod_{k=1}^{d} f(x_k) \prod_{j=1}^{d-1} \prod_{i=1}^{d-j} c_{i,i+j|i+1,\ldots,i+j-1}$$
$$\times (F(x_i|x_{i+1}, \ldots, x_{i+j-1}), F(x_{i+j}|x_{i+1}, \ldots, x_{i+j-1})), \tag{22a}$$

and for a C-vine, the density is equal to

$$f(x_1, \ldots, x_d) = \prod_{k=1}^{d} f(x_k) \prod_{j=1}^{d-1} \prod_{i=1}^{d-j} c_{j,j+i|1,\ldots,j-1}$$
$$\times (F(x_j|x_1, \ldots, x_{j-1}), F(x_{j+i}|x_1, \ldots, x_{j-1})), \tag{23a}$$

where index j identifies the trees, and i denoted the edges in each tree. (Aas et al. 2009).

2.3. *Extreme value theory*

EVT is a statistical technique 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 (e.g. Sheikh and Qiao (2010), Berkowitz et al. (2011)) show that in the majority of cases,

financial asset return distributions are heavy-tailed, especially in times of financial instability.

There are two methods for modelling 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 (Tsay 2014). In this paper, we use the POT method based on the generalised Pareto distribution (GPD). The POT method focuses on modelling exceedances of losses above a certain threshold η 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 later in the subsequent section.

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

$$F_{\xi,\mu,\sigma}(x) = \begin{cases} exp[-(1 + \xi x)^{-1/\xi}] \text{ for } \xi \neq 0, \\ exp[-e^{-x}] \text{ for } \xi = 0, \end{cases} \tag{24}$$

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 $\xi > 0$. Taking into consideration the conditional distribution of the excesses over the threshold η, i.e. $x_i - \eta = y|x_i > \eta$, we have

$$Pr(x - \eta \leq y|x > \eta) = \frac{Pr(\eta \leq x \leq y + \eta)}{Pr(x > \eta)} = \frac{Pr(x \leq y + \eta) - Pr(x \leq \eta)}{1 - Pr(x \leq \eta)} \tag{25}$$

$$= \frac{F(y + \eta) - F(\eta)}{1 - F(\eta)} = F_\eta(y). \tag{26}$$

Again, as $T \to \infty$, $(y + \eta - \mu)/\sigma$ follows a GEV distribution (Eqn.24), therefore

$$Pr(x - \eta \leq y|x > \eta) = \frac{F(y + \eta) - F(\eta)}{1 - F(\eta)}$$

$$= \frac{exp\left[-\left(1 + \frac{\xi(y+\eta-\mu)}{\sigma}\right)^{-1/\xi}\right] - exp\left[-\left(1 + \frac{\xi(\eta-\mu)}{\sigma}\right)^{-1/\xi}\right]}{1 - exp\left[-\left(1 + \frac{\xi(\eta-\mu)}{\sigma}\right)^{-1/\xi}\right]}.$$

From the Taylor series for the expansion of $exp(-x)$, i.e,

$$e^{-x} = \sum_{n=0}^{\infty} \frac{(-x)^n}{n!} = \frac{(-x)^0}{0!} + \frac{(-x)^1}{1!} + \frac{(-x)^2}{2!} + \frac{(-x)^3}{3!} + \dots$$

$$= 1 - x + \frac{x^2}{2!} - \frac{x^3}{3!} + \dots$$

$$\approx 1 - x. \tag{27}$$

Therefore, we have

$$exp\left[-\left(1+\frac{\xi(y+\eta-\mu)}{\sigma}\right)^{-1/\xi}\right] \approx 1-\left(1+\frac{\xi(y+\eta-\mu)}{\sigma}\right)^{-1/\xi}$$

$$exp\left[-\left(1+\frac{\xi(\eta-\mu)}{\sigma}\right)^{-1/\xi}\right] \approx 1-\left(1+\frac{\xi(\eta-\mu)}{\sigma}\right)^{-1/\xi}.$$

Thus,

$$Pr(x-\eta \le y | x > \eta) = \frac{1-\left(1+\frac{\xi(y+\eta-\mu)}{\sigma}\right)^{-1/\xi} - \left[1-\left(1+\frac{\xi(\eta-\mu)}{\sigma}\right)^{-1/\xi}\right]}{1-\left[1-\left(1+\frac{\xi(\eta-\mu)}{\sigma}\right)^{-1/\xi}\right]}$$

$$= \frac{[\sigma+\xi(\eta-\mu)]^{-1/\xi} - [\sigma+\xi(y+\eta-\mu)]^{-1/\xi}}{[\sigma+\xi(\eta-\mu)]^{-1/\xi}}$$

$$= 1-\left(1+\frac{\xi y}{\sigma+\xi(\eta-\mu)}\right)^{-1/\xi}. \tag{28}$$

where $y > 0$ and $\sigma+\xi(\eta-\mu) > 0$. Let $\psi(\eta) = \sigma+\xi(\eta-\mu)$, then as $\eta \to \infty$, Eqn.(28) is approximated by the GPD

$$G_{\xi,\psi(\eta)}(y) = \begin{cases} 1-\left[1+\frac{\xi y}{\psi(\eta)}\right]^{-1/\xi} & \text{for } \xi \ne 0, \\ 1-exp[-y/\psi(\eta)] & \text{for } \xi = 0, \end{cases} \tag{29}$$

with shape parameter ξ and scale parameter $\psi(\eta)$, where $\psi(\eta) > 0, y \in [0, x-\eta]$ when $\xi \ge 0$, and $y \in [0, -\frac{\psi(\eta)}{\xi}]$ when $\xi < 0$. If $\xi = 0$, then Eqn.(29) becomes an exponential distribution with parameter $1/\sigma$ (Tsay 2014). Let $y = x - \eta$, then Eqn.(26) can be written as

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

$$\implies F(x) = F(\eta) + [1-F(\eta)]\,G_{\xi,\psi(\eta)}(x-\eta). \tag{30}$$

To construct the tail estimator for the distribution, $F(x)$, we need an estimate of $F(\eta)$. This is obtained using the empirical distribution of $F(\eta) = (T-N_\eta)/T$. Substituting $F(\eta)$ and $G_{\xi,\psi(\eta)}(x-\eta)$ in Eqn.(30), we have:

$$F(x|\xi,\psi(\eta)) = \frac{T-N_\eta}{T} + \left[1-\left(\frac{T-N_\eta}{T}\right)\right]\left[1-\left\{1+\frac{\xi(x-\eta)}{\psi(\eta)}\right\}^{-1/\xi}\right].$$

After some algebra, we have that:

$$F(x|\xi,\psi(\eta)) \approx \frac{T-N_\eta}{T}\left[1+\frac{\xi(x-\eta)}{\psi(\eta)}\right]^{-1/\xi} \tag{31}$$

where N_η is the number of observations above the threshold and T is the total number of observations.

We obtain $VaR_q(Z)$, the qth quantile of $F_q^{-1} = VaR_q(Z)$, by simply inverting Eqn. (31) or solving for x for any given small upper tail probability p to obtain

$$VaR_q(Z) = \eta - \frac{\psi(\eta)}{\xi}\left\{1 - \left[\frac{T}{N_\eta}(1-q)\right]^{-\xi}\right\}, \tag{32}$$

where $q = 1 - p$ (see Tsay (2014), Soltane et al. (2012), Bhattacharyya and Ritolia (2008)).

Assuming that the number of points above the threshold η is independent and identically distributed, the parameters $\psi(\eta)$ and ξ can be estimated by means of maximum likelihood estimation with likelihood function

$$l(x_i, \ldots, x_{N_\eta}|\xi, \sigma, \mu) = \prod_{i=1}^{N_\eta} f(x_i) \quad \text{for} \quad x_i > \eta, \tag{33}$$

where N_η is the total number of points above η.

The choice of the threshold η is an important step in the POT method because Eqn.(32) is dependent on η and the number of points (i.e. exceedances) above η 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. Danielsson and de Vries (1998) developed a semi-parametric estimator for the tails of the distribution that estimated the threshold through the bootstrap approximation of the mean square error (MSE) of the tail index and by minimising MSE through the choice of the threshold. Danielsson et al. (2001) further used a two-step subsample bootstrap method to determine the threshold that minimised the asymptotic MSE. Hill et al. (1975) and Davison and Smith (1990) proposed 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 the *hybrid* method for threshold selection.

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

$$e(\eta) = E(x - \eta|x > \eta). \tag{34}$$

Following GPD, $e(\eta)$ is a linear function of η defined as

$$e(\eta) = \frac{\sigma + \xi\eta}{1 - \xi}, \tag{35}$$

where η is identified as the lowest point on the mean excess plot above which the graph appears to be approximately linear.

The choice of η from the mean excess plot is very subjective (Wong 2013, Tsay 2014), which might differ from one institution to another using the same data because of different risk tolerances. In this context, different η values will give different estimates of the shape and scale parameters, and hence different VaRs. 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 $\eta \to \infty$; data beyond the threshold will deviate from the GPD since the GPD is not a good approximation for the body of the sample data (Bhattacharyya and Ritolia 2008, Wong 2013). We propose a *hybrid* method for selecting a proper threshold value that will restrict inferences to the left tail of the distribution.

2.3.1. **The hybrid method for threshold selection.** Identify from the mean excess plot the lowest point above which the graph appears to be approximately linear and call this point η_0, then insert a tangent line from η_0 through the rest of the points η_i where $\eta_i > \eta_0$. 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 lie on the tangent line as the threshold value. We call this point η^*. η^* will lead to a better approximation of VaR estimates than η_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 on the same data and also the probability of selecting a very low or very high threshold value. Let $\eta_i = \eta_1, \ldots, \eta_\hbar$ be a set of points that lie on the straight line, then we obtain the value of η^* as

$$\eta^* = \frac{1}{\hbar} \sum_{i=1}^{\hbar} \eta_i, \quad \eta_i \geq \eta_0 \tag{36}$$

where $\hbar$ is the number of points in the set.

3. Empirical Analysis

3.1. *Data*

The data employed in this analysis consist of 2870 observations of daily stock prices actively traded on the London Stock Exchange. The stocks belong to the banking sector and of the top five banks in UK i.e, HSBC bank, LLOYDS Banking Group, Barclays bank, Royal Bank of Scotland, and Standard Chartered PLC bank. Our motivation for these banks is because we want to test the reliability of the VaR model in banks with relatively high ratings in periods of distress and also investigate the risk of collapse in banks with insufficient capital to provide proper cushions in periods of distress. If the top banks are not able to cope during financial distress, then the risk of collapse in the country's banking system is large. Therefore, our data covers the period from 31st of December 2004 to the 31st of December 2015, covering the 2008 global financial crisis and 2011 European financial crisis. For example, the United Kingdom (UK) economy was in recession for 15 months from the second quarter of 2008 to the second quarter of 2009 (DAmuri and Peri 2014, Jenkins et al. 2012, Lin et al. 2012). All data are from DataStream.

We use out-of-sample data of $m = T - n$ observations for back-testing; thus we have $n = 1869$, i.e. the 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 estimated for day $t = n + 1$ using data from day $t = 1$ to day $t = n$, VaR for day $t = n + 2$ is estimated using data from day $t = 2$ to day $t = n + 1$, and so on until the out-of-sample data are all used up.

Figure 1 shows the trends in the stock prices during the study period. The effects of the 2008 global financial crisis and the 2011 European financial crisis are clearly seen.
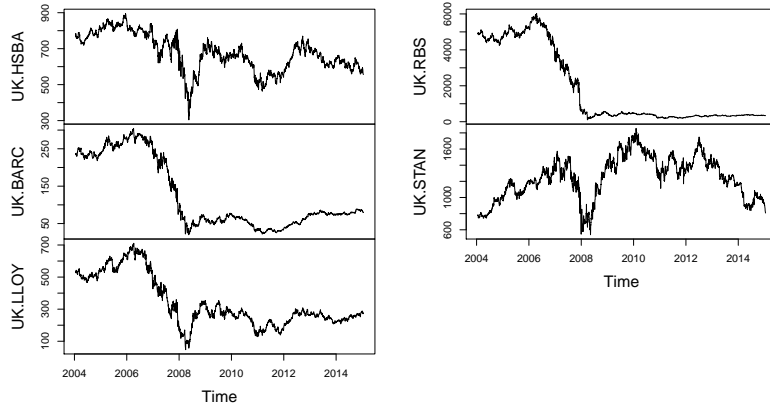
Figure 1. Trends in stock prices over the period from 31 December 2004 to 31 December 2015 showing an effect of the 2008 global financial crisis and the 2011 European financial crisis.

Some literature measures the stability in the financial systems using a portfolio of several banks; i.e. considering the dependence among the banks, while others focus on individual banks. This paper considers both measures. By using the stock prices for each bank, we calculate the simple return series and apply risk factor mappings to construct a simulated portfolio returns consisting of all banks involved. Since the stocks are all from banks of almost the same strength and ratings (i.e, the top five banks in the UK), we assume equal weights. That is, consider a portfolio consisting of N stocks represented in vector form as $S_N = (s_{1t}, \ldots, s_{Nt})$, the simple returns, R_t, are calculated as

$$R_{i,t} = (R_{1,t}, \ldots, R_{N,t}) = \left[\left(\frac{P_{1,t} - P_{t-\tau}}{P_{t-\tau}} \right), \ldots, \left(\frac{P_{N,t} - P_{N,t-\tau}}{P_{N,t-\tau}} \right) \right]. \tag{37}$$

and the expected return on the portfolio at time t is given by

$$\bar{R}_{p,t} = E\left(R_{p,t} \right) = \sum_{i=1}^{N} w_i R_{i,t}, \quad \sum_{i=1}^{N} w_i = 1. \tag{38}$$

which is a weighted average of the individual stocks in the portfolio, where w_i is the weight applied to stock i and $P_{i,t}$ is the price of stock i at time t. Since simple returns are not additive across time, we convert the simple returns for each stock and the portfolio to log-returns for further analysis. This is important in case we would want to calculate the portfolio VaR for different time horizons. The simple returns are converted to log-returns as

$$r_{i,t} = log(R_{i,t} + 1) \tag{39}$$

Figure 2 shows time plots of the log-return series and the portfolio; this shows evidence of volatility clustering in the return series. 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[1] as indicated by the Jarque-Bera normality test in addition to their high excess kurtosis and skewness.

---

[1]A normal distribution is symmetric with excess skewness and kurtosis equals to zero.

Furthermore, the Ljung-Box test on the squared residuals of the mean equation $\left\{a_{i,t}^2\right\}$; where $a_{i,t} = r_{i,t} - \mu_i$, and a Lagrange multiplier test for *Autoregressive Conditional Heteroscedasticity* (ARCH LM test) on the residuals of the mean equation $\{a_{i,t}\}$, as described by Tsay (2005, 2014), rejects the null hypothesis of no ARCH effect or serial correlation in the log-return series.
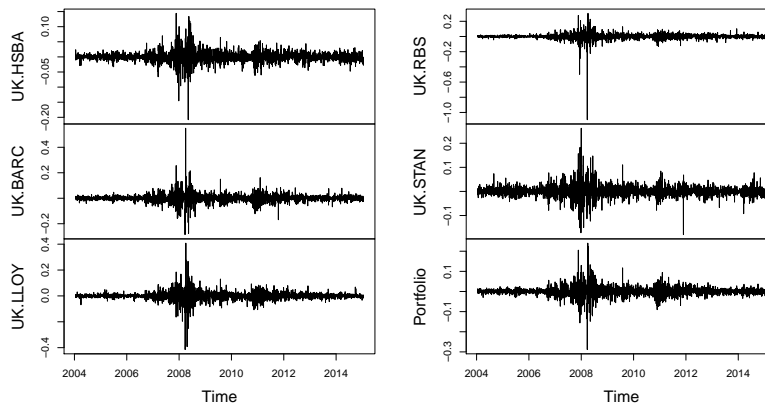

Figure 2. Time plots of the log-return series and the portfolio indicating the presence of volatility clustering in the data.

Table 1. Summary statistics of daily log-return series.

	UK.HSBA	UK.BARC	UK.LLOY	UK.RBS	UK.STAN	Portfolio
Mean	-0.0001	-0.0004	-0.0003	-0.0010	-0.0001	-0.0002
Variance	0.0003	0.0011	0.0010	0.0015	0.0006	0.0006
Stdev	0.0171	0.0328	0.0321	0.0388	0.0244	0.0236
Skewness	-0.3367	-1.0549	1.4387	-8.4013	0.3161	-0.0350
Excess Kurtosis	16.9080	37.2754	40.2179	235.5263	13.0850	21.3981
SE	0.0003	0.0006	0.0006	0.0007	0.0005	0.0004
Jarque-Bera						
X-squared	34284.9000	166881.0700	194636.6700	6674449.8600	20550.5600	54823.4918
Probability	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
ARCH Test						
Chi-squared	698.1700	933.2600	449.4000	45.1000	818.4000	544.2900
p-value	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Ljung-Box Test						
X-squared	2475.7000	2552.4000	948.1000	58.8670	3727.3000	132.1500
p-value	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Note: Stdev = Standard deviation, SE = Standard error.

3.2. *Modelling the marginal distributions of volatility equations*

As explained earlier, the data in the log-return series are leptokurtic. Thus, to capture the tail distribution and the dynamics of fluctuations in the time series data, we fit a GARCH(1,1) model with Student's-t distribution to the time series data to accommodate the heavy tails and estimate the GARCH parameters using Bayesian statistics as follows: (i) We assign a prior distribution with initial hyperparameters following GARCH specifications (i.e. $\alpha_0 > 0, \alpha_1, \beta_1 \geq 0$ and $\alpha_1 + \beta_1 < 1$) and generate two Markov chain Monte Carlo (MCMC) of 100000 draws each; (ii) if

convergence is attained, we discard the first 50000 draws and select only the $50th$ 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-GARCH(1,1) model with a Student's-t distribution. We test for convergence of the sampler with the help of a diagnostic test by Gelman and Rubin (1992) (i.e. for example, Figures 3, 4, 5, and 6 in the case of the portfolio return series). Estimation results are presented in Table 2 with standard errors in parenthesis.

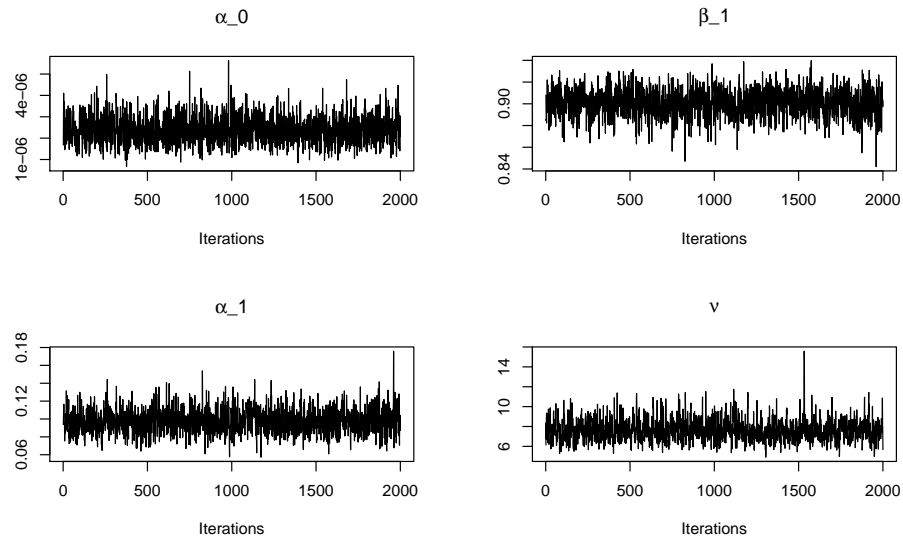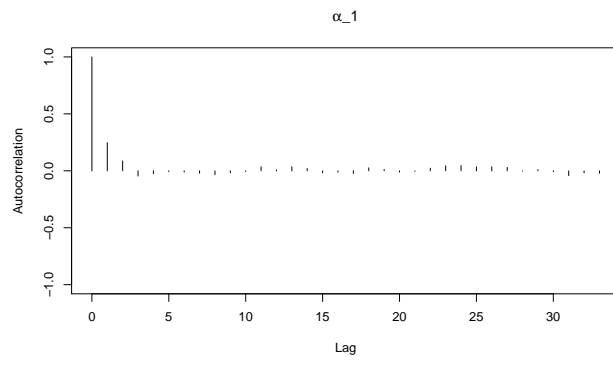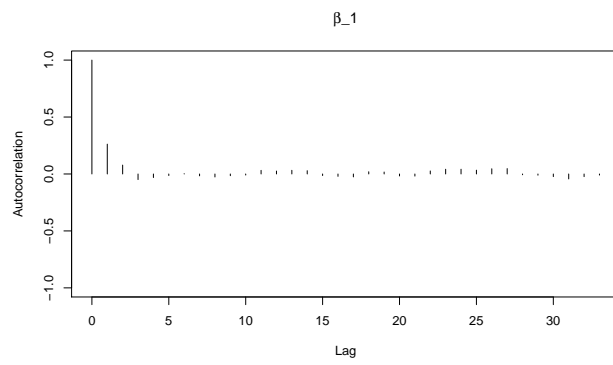Figure 3. Trace plots of 2000 iterations against the values of the draws of the parameters at each iteration after merging the two chains. The plots show no evidence against convergence.
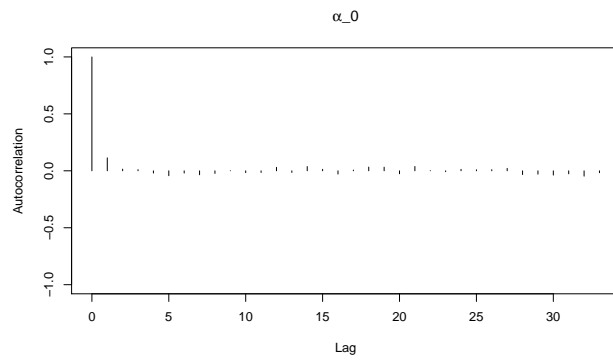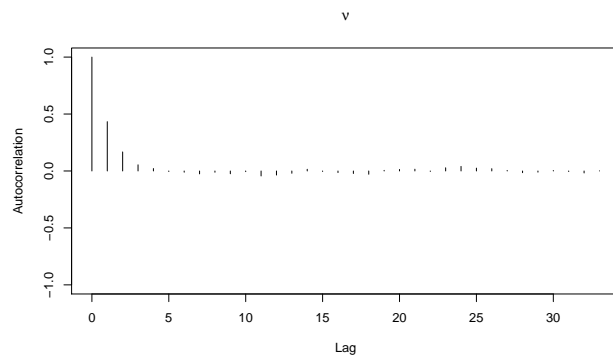
(a)

(b)

Figure 4. Autocorrelation plots of 2000 samples for α_1 and β_1 after merging the two chains.

α_0

(a)

ν

(b)

Figure 5. Autocorrelation plots of 2000 samples for α_0 and ν after merging the two chains.
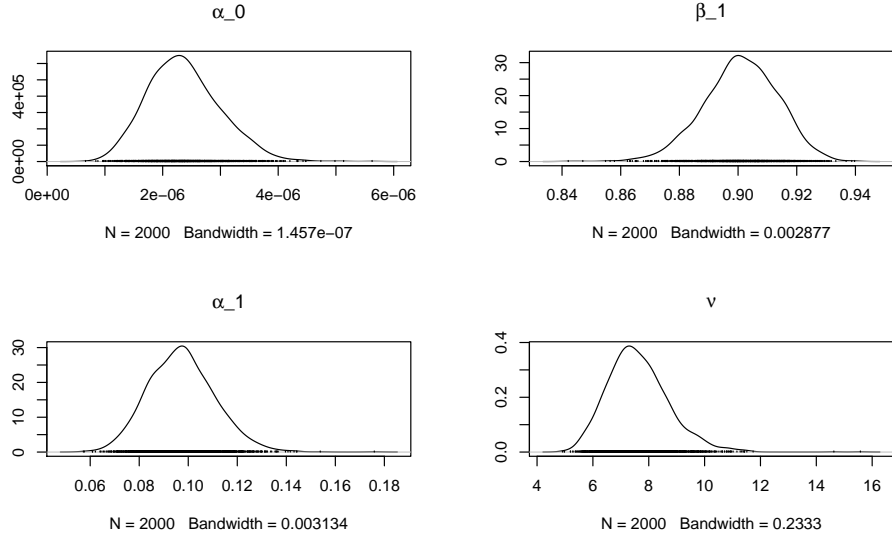
Figure 6. Density plots of the posterior distributions of the model parameters. nu = density of degrees of freedom. Density plots are used to test the covariance stationarity condition. For GARCH(1,1) model, $\alpha_1 + \beta_1 < 1$.

Table 2. Parameter estimates of Bayesian-GARCH(1,1) model with Student-t distribution.

| | Parameter | | | | Diagnostic check | | |
	α_0	α_1	β_1	v	a.rate (α)	a.rate (β_1)	psrf
UK.HSBA	2.078e-06(1.355e-08)	0.0934(3.105e-4)	0.9057(2.781e-4)	4.9740(1.097e-4)	92.72%	97.63%	1.01
UK.BARC	2.526e-06(1.733e-08)	0.0827(3.014e-04)	0.9171(2.663e-04)	5.2000(1.138e-02)	89.59%	97.63%	1.00
UK.LLOY	6.375e-06(3.827e-08)	0.1022(3.532e-04)	0.8951(3.259e-04)	5.3501(1.196e-02)	92.66%	97.71%	1.00
UK.RBS	3.446e-06(2.477e-08)	0.0876(4.046e-04)	0.9077(3.386e-04)	4.8620(1.328e-02)	90.98%	97.64%	1.00
UK.STAN	1.106e-05(7.034e-08)	0.1290(5.029e-04)	0.8558(5.383e-04)	5.3920(1.253e-02)	94.20%	97.70%	1.00
Portfolio	2.364e-06(1.4060e-09)	0.0973(3.0960e-05)	0.9009(2.8370e-05)	7.6750(2.480e-03)	90.17%	97.66%	1.00

Note: Standard error in parenthesis. a.rate = parameter acceptance rate, which is the proportion of the total number of single values in the MCMC chain to the total number of values in the chain. A high acceptance rate tells us that the chain does not get stuck in certain areas in the parameter space, thus producing good mixing as seen in the example of Figure 3. psrf $= \sqrt{\frac{\hat{Var}(x)}{W}}$; the potential scale reduction factor and should be < 1.2, where $\hat{Var}(x)$ is a weighted average of the average of the m within-sequence variance, s_j^2, each based on $n-1$ degrees of freedom, and the variance between the m sequence means, $\bar{x}_j$, each based on n values of x: $\hat{Var}(x) = (\frac{n-1}{n})W + \frac{1}{n}B$; $W = \frac{1}{m}\sum_{j=1}^{m}s_j^2$, $B = \frac{n}{m-1}\sum_{j=1}^{m}(\bar{x}_j - \bar{x}..)^2$, $s_j^2 = \frac{1}{n-1}\sum_{j=1}^{n}(x_{ij} - \bar{x}_j)^2$. If psrf > 1.2, then the length of the chain should be increased to improve convergence to a stationary distribution (see Gelman and Rubin (1992) for more details).

Applying Eqn.(3), we then obtain a matrix Σ, which consists of marginal standardised residuals $\{\epsilon_{i,t}\}_{t=1}^{T}$. That is

$$\Sigma_{i,t} = \epsilon_{i,t} = (r_{i,t} - \mu_i)\left(\frac{v-2}{v}\omega_{i,t}h_{i,t}(\alpha_i, \beta_i)\right)^{-1/2}, \quad i = 1,\ldots,5; t = 1,\ldots,T, \quad (40)$$

where the vectors $(r_{i,t} - \mu_i)'$ are the residuals of the mean equation.

To check for conditional heteroscedasticity in the standarsised residuals, we employ the Ljung-Box test statistics, $Q_k(m)$ and its modification known as robust, $Q_k^r(m)$. The modification 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 the $Q_k^r(m)$ test is because $Q_k(m)$ may fare poorly in finite samples when the residuals of the time series, a_t, have heavy tails (Tsay 2013). The Ljung-Box test statistics is given by

$$Q_k(m) = T^2 \sum_{i=1}^{m} \frac{1}{T-i} \mathbf{b}_i'(\hat{\boldsymbol{\rho}}_0^{-1} \otimes \hat{\boldsymbol{\rho}}_0^{-1})\mathbf{b}_i \approx \chi_{k^2}^2(m), \tag{41}$$

where k is the dimension of a_t, T is the sample size, $\mathbf{b}_i = vec(\hat{\rho}_i')$ with $\hat{\boldsymbol{\rho}}_j$ being the lag-j cross-correlation matrix of $a_{i,t}^2$. The test indicates the presence of conditional heteroscedasticity in the standardised residuals with $Q_k(10) = 92.6690$, p-value$= 1.5e^{-15}$, and $Q_k^r(10) = 318.6257$, p-value$= 0.0022$.

3.3. *Modelling dependence*

We model the dependence structure among the stock returns using vine copula functions. The C- and D-vine copula parameters are estimated by maximising the likelihood function:

$$l_{C-vine}(\theta|\mathbf{x}) = \sum_{j=1}^{d-1}\sum_{i=1}^{d-j}\sum_{t=1}^{T} \log[c_{j,j+i|1,\ldots,j-1}(F(x_{j,t}|x_{1,t},\ldots,x_{j-1,t}),$$

$$F(x_{j+i,t}|x_{1,t},\ldots,x_{j-1,t})|\theta_{j,j+i|1,\ldots,j-1})]; \tag{42a}$$

$$l_{D-vine}(\theta|\mathbf{x}) = \sum_{j=1}^{d-1}\sum_{i=1}^{d-j}\sum_{t=1}^{T} \log[c_{i,i+j|i+1,\ldots,i+j}(F(x_{i,t}|x_{i+1,t},\ldots,x_{i+j-1,t}),$$

$$F(x_{i+j,t}|x_{i+1,t},\ldots,x_{i+j-1,t})|\theta_{i,i+j|i+1,\ldots,i+j-1})], \tag{42b}$$

where $x_i \in [0,1]$ are pseudo observations of the standardised residuals and, θ_i and θ_j are the pair-copula parameters of the joint distribution function (Schepsmeier and Brechmann 2015, Aas et al. 2009).

The conditional distribution functions in C-vine (i.e. Eqn.(42a)) and D-vine (i.e. Eqn. (42b)) copulas are obtained from the conditional distribution

$$h(\mathbf{x}, \mathbf{v}, \theta) = F(\mathbf{x}|\mathbf{v}) = \frac{\partial C_{xv_j|\mathbf{v}_{-j}}(F(x|\mathbf{v}_{-j}), F(v_j|\mathbf{v}_{-j}))}{\partial F(v_j|\mathbf{v}_{-j})}, \tag{43}$$

where $C_{xv_j|\mathbf{v}_{-j}}$ is a bivariate copula distribution function.

We select the best pair-copula for the decomposition of the n-variate copula densities based on the paired copula with the smallest AIC value from a range of copula families (Table 3). To select which of the vine copulas is the best to model the dependence among the risk factors, we follow the procedure of Vuong (1989). That is, we use a likelihood-ratio based test to compare non-nested models. We cannot select between the two vine copulas based on their likelihoods because the two

copulas are non-nested (Aas et al. 2009). The test statistics is given by

$$\nu = \frac{\frac{1}{N}\sum_{i=1}^{N} m_i}{\sqrt{\sum_{i=1}^{N}(m_i - \bar{m})^2}} \tag{44}$$

$$m_i = log\left[\frac{c_1(u_i|\hat{\theta}_1)}{c_2(u_i|\hat{\theta}_2)}\right], \quad u_i \in [0,1], \quad i = 1,\dots,N$$

where ν is the standardised sum of the log differences of the pointwise likelihoods, m_i, between two competing vine copulas c_1 and c_2 and the estimated parameters $\hat{\theta}_1$ and $\hat{\theta}_2$. Vine model 1 is selected in favor of vine model 2 at a certain level of confidence α if and only if $\nu > \Phi^{-1}\left(1 - \frac{\alpha}{2}\right)$, and vine model 2 is selected in favor of vine model 1 if and only if $\nu < -\Phi^{-1}\left(1 - \frac{\alpha}{2}\right)$. If $|\nu| \le \Phi^{-1}\left(1 - \frac{\alpha}{2}\right)$ then no decision among the models is possible (Schepsmeier and Brechmann 2015). Based on the likelihood ratio test, C-vine copula is selected in favor of D-vine with $\nu = 2.4989$ and $\Phi^{-1}\left(1 - \frac{\alpha}{2}\right) = 1.96$ at $\alpha = 5\%$. From now on, we proceed with the C-vine copula. The estimated copula parameters are used to simulate 10000 pairs (u_i, u_j) of observations of $[0,1]$ uniformly distributed random variables with the joint distribution function $C(u_i, u_j)$ (see Cherubini et al. (2004) for a detailed simulation technique). The simulated data is then transformed to the original scales of the noise variables using the inverse quantile function $F_i^{-1}(u_i)$, of the desired marginal distributions to obtain a new matrix:

$$\hat{\Sigma} = \{\zeta_{i,t}\}, \quad i = 1,\dots,5,\, t = 1,\dots,10000, \tag{45}$$

which is free from assumptions of normality or linear correlations. Here, we compare between Student's-t and normal marginal distributions by changing the marginals but keeping the copula fixed.

Multivariate ARCH test, Eqn.(41), on $\{\zeta_{i,t}\}$ at 95% significance level shows no evidence of conditional heteroscedasticity (Table 4). Therefore, the Bayesian-GARCH(1,1) C-vine copula model is a better model in terms of describing the conditional heteroscedasticity in the log return series and modelling dependence as compared to the Bayesian-GARCH(1,1) model without copula functions where there is evidence of ARCH effect in the standardised residuals.

Table 3. C-vine and D-vine copula parameter estimates.

	Unconditional and conditional pairs copulas	Selected copula family	Parameter	Log-likelihood	paired AIC
C-vine Copula					
$Tree_1$	$C_{1,2}$	Student's-t copula	$\rho = 0.5536, v = 5.98$	545.2690	-1086.54
	$C_{1,3}$	Student's-t copula	$\rho = 0.6222, v = 6.27$	716.3512	-1428.70
	$C_{1,4}$	Student's-t copula	$\rho = 0.5701, v = 5.80$	573.0684	-1142.14
	$C_{1,5}$	Student's-t copula	$\rho = 0.6384, v = 6.63$	758.9981	-1513.10
$Tree_2$	$C_{2,3\mid1}$	Frank copula	$\lambda = 4.0043$	497.7722	-993.54
	$C_{2,4\mid1}$	Student's-t copula	$\rho = 0.5870, v = 7.94$	602.6068	-1201.21
	$C_{2,5\mid1}$	Frank copula	$\lambda = 1.9893$	140.8872	-279.77
$Tree_3$	$C_{3,4\mid1,2}$	Student's-t copula	$\rho = 0.3687, v = 8.7657$	222.7078	-441.42
	$C_{3,5\mid1,2}$	Student's-t copula	$\rho = 0.2124, v = 10.8326$	74.4631	-144.93
$Tree_4$	$C_{4,5\mid1,2,3}$	Frank copula	$\lambda = 0.5513$	10.8833	-19.77
Log-likelihood				4143.0070	
D-vine Copula					
$Tree_1$	$C_{1,2}$	Student's-t copula	$\rho = 0.5536, v = 5.98$	545.2690	-1086.54
	$C_{2,3}$	Student's-t copula	$\rho = 0.6969, v = 6.19$	688.4580	-1893.67
	$C_{3,4}$	Student's-t copula	$\rho = 0.7197, v = 4.91$	464.5665	-2106.17
	$C_{4,5}$	Student's-t copula	$\rho = 0.5385, v = 6.11$	726.9948	-972.03
$Tree_2$	$C_{1,3\mid2}$	Student's-t copula	$\rho = 0.3915, v = 8.26$	440.2784	-493.94
	$C_{2,4\mid3}$	Student's-t copula	$\rho = 0.4342, v = 8.50$	548.5746	-620.12
	$C_{3,5\mid4}$	Student's-t copula	$\rho = 0.3534, v = 9.67$	115.0825	-396.34
$Tree_3$	$C_{1,4\mid2,3}$	Student's-t copula	$\rho = 0.1499, v = 12.11$	184.4686	-76.18
	$C_{2,5\mid3,4}$	Student's-t copula	$\rho = 0.1434, v = 12.83$	50.6013	-71.75
$Tree_4$	$C_{1,5\mid2,3,4}$	Student's-t copula	$\rho = 0.3787, v = 12.44$	-189.6890	-447.73
Log-likelihood				3574.6050	

Note: The copula types for the decomposition of n-variate bivariate copulas for unconditional and conditional pairs are selected based on AIC values. That is, the paired copula with the smallest AIC value.

Table 4. Multivariate ARCH test (Eqn.(41)) on the standardised residuals after modelling dependence with copulas.

		$m = 5$	$m = 10$	$m = 15$	$m = 20$	$m = 30$
t-distribution	$Q_k(m)$	1.8923	4.9683	7.1327	9.3807	19.0364
	p-value	0.8638	0.8933	0.9539	0.9781	0.9392
	$Q_k^r(m)$	117.9122	251.0490	371.3950	519.748	764.4626
	p-value	0.6608	0.4694	0.5429	0.2619	0.3489
Normal distribution	$Q_k(m)$	2.6596	5.1806	10.9645	15.7037	28.9405
	p-value	0.7523	0.8788	0.7551	0.7348	0.5207
	$Q_k^r(m)$	123.6943	245.9011	348.5009	491.0212	705.8559
	p-value	0.5162	0.5614	0.8332	0.6043	0.8740

Note: p-values $> 5\%$ suggest no evidence of ARCH effect or conditional heteroscedasticity.

4. Results

4.1. *Threshold selection and robust VaR estimates*

We follow the approach by McNeil and Frey (2000) and apply the POT method of EVT to each of the marginal distributions of $\{\zeta_{i,t}\}$ (i.e. Eqn.(45)) to obtain the qth 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. For example, we use the mean excess function plot for Bank 4 (Figure 7), following the Bayesian-GARCH(1,1) C-vine copula with t-margins, to demonstrate the threshold selection method. Based on the outcome of the plot, we select a subjective threshold value of approximately 1.2. That is, $\eta_1 = 1.2$, which is the lowest point on the graph; note that the graph appears to be approximately linear. However, if we select this point as the threshold value, we will have 1470 exceedances, which are too many compared to the size of the data (i.e. $T = 10000$), and the number of exceedances will lie towards the center of the data, which will inevitably result in a poor approximation of the GPD parameters and hence lead to inaccuracies in the VaR estimate.

Figure 7. Mean excess function plot of standardised residuals for the threshold selection following the Bayesian-GARCH(1,1) C-vine copula model.

Following the *hybrid* method, as demonstrated in Figure 8, we obtain a threshold value of 2.1146 and 465 exceedances. The threshold range plot; Figure 9, which is a plot of reparameterised scale of the shape parameter suggest that $\eta^* = 2.1146$ is a reasonable choice to use as the threshold value because 2.1146 seems to yield POT parameter estimates that will not change significantly within uncertainty bounds. The number of exceedances produced is also sufficient to allow for reasonable statistical inferences with EVT.
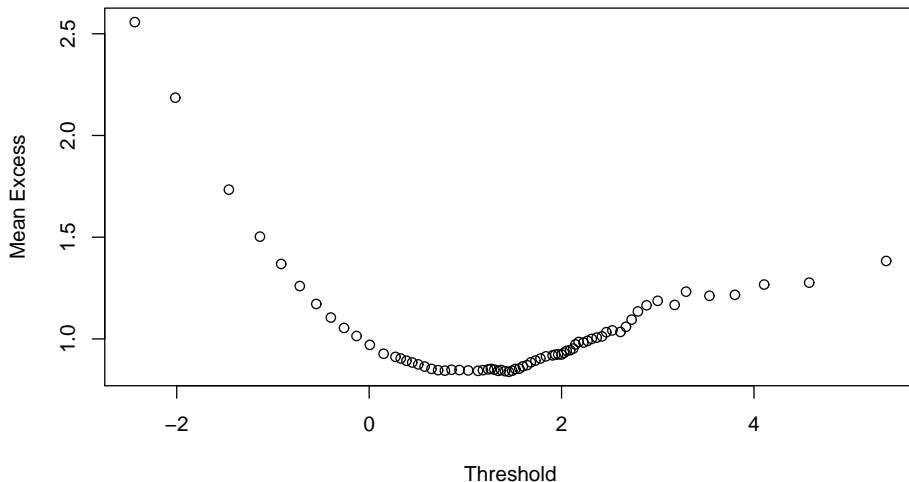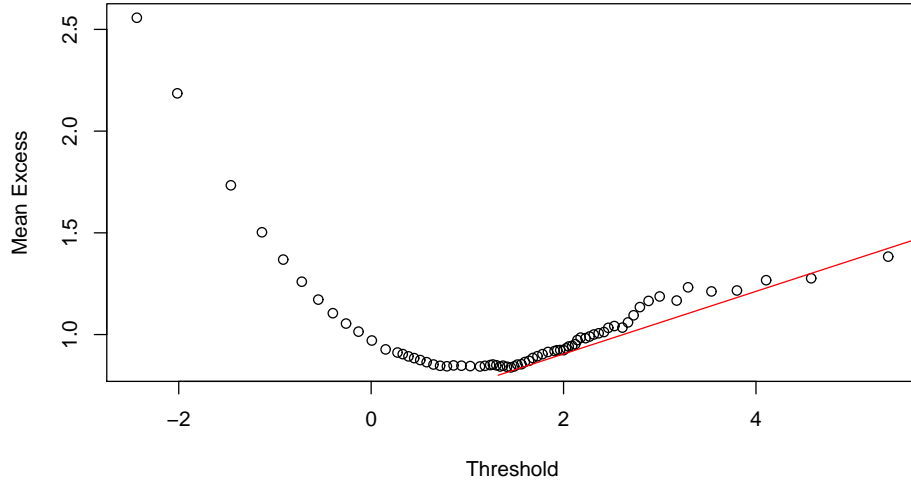
Figure 8. Mean excess function plot of standardised residuals following the Bayesian-GARCH(1,1) C-vine copula model demonstrate the *hybrid* method of threshold selection. The threshold is the average of all the points that lie on the tangent line.
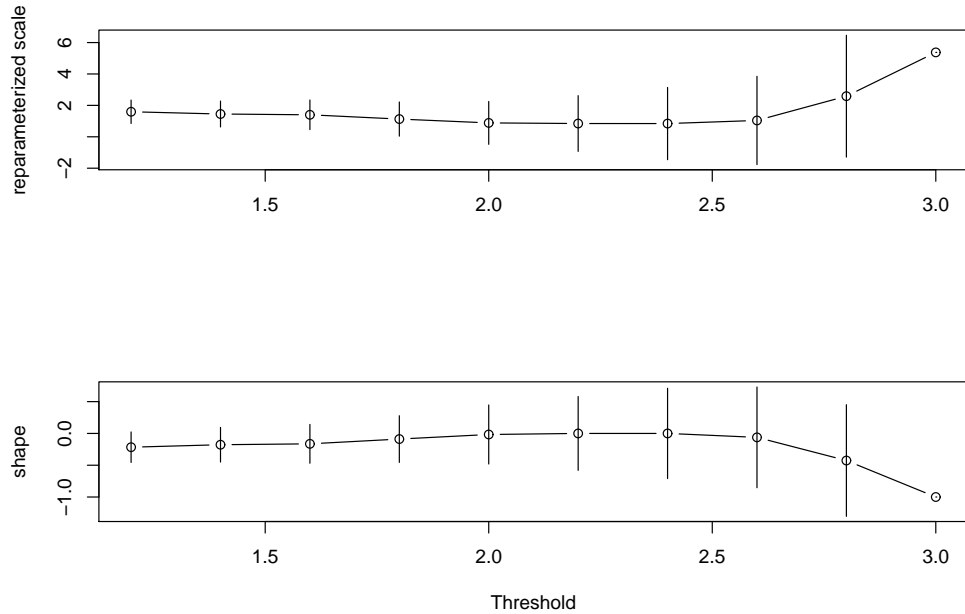
Figure 9. The threshold range plot shows that $\eta = 2.1146$ is appropriate to use as the threshold value as it seems to yield POT parameter estimates that will not change significantly within uncertainty bounds.

Table 5 shows the POT parameter estimates and the qth quantile VaR, $VaR_q(Z)$, estimates of the noise variables at $q = 99\%$ and $q = 95\%$. From Eqn. (29), where $\psi(\eta) > 0, y \in [0, x - \eta]$ when $\xi \geq 0$, and $y \in [0, -\frac{\psi(\eta)}{\xi}]$ when $\xi < 0$, thus the appropriate values for the shape parameter ξ for a financial time series data must be greater than zero since the upper bound of financial losses

22

cannot be fixed (Bhattacharyya and Ritolia 2008, Soltane et al. 2012, Gilli et al. 2006). Therefore, the GARCH(1,1) C-vine copula-EVT model with normal margins is not a suitable model in this case as the shape parameters are all less than zero. The shape parameter indicates the thickness of the tail of the distribution. $\xi < 0$ implies a lighter tail.

Table 5. POT parameter estimates and q^{th} quantile
VaRs, $VaR_q(Z)$.

| | | | | Parameters | | | | $VaR_q(Z)$ | |
		ξ	$\psi(\eta^*)$	η^*	N_{η^*}	μ	σ	99%	95%
C-vine copula: t margins	UK.HSBA	0.0400	0.9593	2.5481	270	-0.6778	0.8301	3.5200	1.9642
	UK.BARC	0.0889	0.9077	2.7751	200	-0.2242	0.6411	3.4240	1.9764
	UK.LLOY	0.1765	0.7017	2.6973	501	0.4316	0.4137	3.3712	2.0646
	UK.RBS	0.1654	0.7941	2.1146	465	0.2040	0.4780	3.5043	2.0573
	UK.STAN	0.0736	0.7696	1.9858	495	-0.0890	0.6166	3.2922	1.9781
C-vine copula: normal margins	UK.HSBA	-0.1523	0.4265	1.4231	560	-0.1204	0.6616	2.0694	1.4710
	UK.BARC	-0.1784	0.4202	1.5786	379	-0.2893	0.7536	2.0770	1.4593
	UK.LLOY	-0.0771	0.3749	1.5762	380	0.1820	0.4824	2.0517	1.4722
	UK.RBS	-0.1326	0.3880	1.6775	260	-0.1437	0.6294	2.0257	1.4124
	UK.STAN	-0.0665	0.3274	1.6628	280	0.3407	0.4154	1.9886	1.4692

Note: With normal margins, $\xi < 0$. For financial data, the upper bound losses cannot be fixed; thus, the only applicable situation is when $\xi \geq 0$. Hence normal margins with $\xi < 0$ is inappropriate.

Employing the risk formula, we can now compute the portfolio quantile VaRs, $VaR_q^p(Z)$ for holding multiple investments in all five banks as

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

where $\rho_{i,j}$ is the Pearson's cross-correlation coefficient between the returns of the ith and jth stocks. For simplicity, and since the stocks are all from banks of almost the same strength and ratings, i.e., the top 10%, we assume equal weights w. This gives: $VaR_{99\%}^p(Z) = 2.7891$ and $VaR_{95\%}^p(Z) = 1.6363$.

4.2. Robust VaR estimates

Using Eqn.(46), the portfolio VaR for a single period is defined as

$$VaR_{q,t}^p = \hat{\mu}_{t+1} + VaR_q^p(Z)\hat{h}_{t+1}^{1/2}, \quad (47)$$

where $\hat{h}_t$ is the estimate of the conditional variances and $\hat{\mu}_t$ is an estimate of the conditional mean obtained by fitting a univariate Bayesian GARCH(1,1) model with Student's-t distribution to the portfolio log returns (Eqn.(3a) and (3b)). The data used are from daily stock returns; thus, these are daily VaR estimates referred to as *robust* because they incorporate volatility clustering and are free from any normality assumption. The results of Eqn.(47) can be used in conjunction with the $\alpha - root$ of time rule to obtain VaR estimates of any desired time horizon $\Bbbk$ as

$$VaR_{q,t}^p(\Bbbk) = \Bbbk^{1/\alpha} VaR_{q,t}^p, \quad 1/\alpha = \xi, \quad (48)$$

where ξ is the shape parameter of the POT method of EVT and α is the tail index (Tsay 2014, Danielsson and De Vries 2000). Figure 10 shows a time plot of profit and loss (P&L) of the portfolio

return series and portfolio VaR estimates. A visual observation of the plot suggests that the VaR model performs quite well in capturing the dynamics in the portfolio return series. However, the model has to be validated through back-testing.

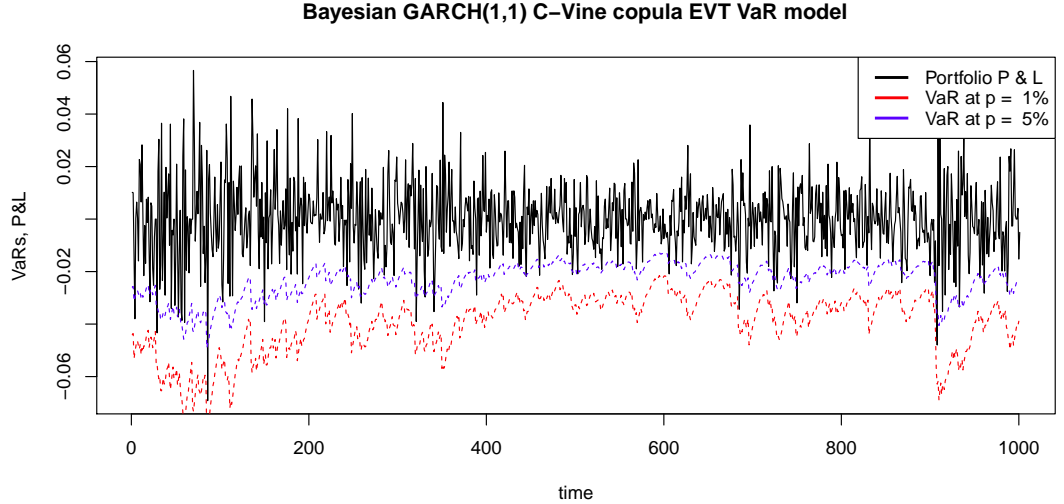**Bayesian GARCH(1,1) C–Vine copula EVT VaR model**

Figure 10. Estimated daily VaRs and profit and loss (P&L) plot for an investment in a portfolio consisting of all five banks following Bayesian-GARCH(1,1) with Student's-t distribution, C-vine copula functions, and EVT.

5.   Back-testing

To check the reliability of the VaR model in the portfolio, i.e. checking whether the model does not overestimate or underestimate the level of risk in the portfolio, we do back-testing on the model. This involves comparing the estimated VaRs over a given observation period with the subsequent returns. 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) property. Consider an indicator function on the exceptions

$$\mathbb{I}_t(1-q) = \mathbb{I}_{\left\{L_t > VaR_{q,t}^p\right\}} = \begin{cases} 1, & \text{if } L_t > VaR_{q,t}^p \\ 0, & \text{otherwise,} \end{cases} \tag{49}$$

explained as $\mathbb{I}$ registers a 1 on day t if the loss on the portfolio L_t on day $t > VaR_{q,t}^p$, and 0 if the loss on day $t \leq VaR_{q,t}^p$. Note that q is the choice of confidence level. For the UC property, $P[\mathbb{I}_t(1-q) = 1] \approx 1-q, \forall_t$; 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 being 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.

   Several back-testing methods have been proposed to test the UC and IND properties of reliable VaR models. The most common back-testing methods include Kupiec's "proportion of failures" (POF) test for the UC (Kupiec 1995), Christoffersen's test for the IND and conditional coverage

(CC) (Christoffersen 1998), and the Basel "traffic light" (Basel Committee on Banking Supervision 1996).

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

$$LR_{POF} = -2ln\frac{q^{T_0}(1-q)^{T_1}}{\left(1-\frac{T_1}{T_w}\right)^{T_0}\left(\frac{T_1}{T_w}\right)^{T_1}}, \tag{50}$$

asymptotically distributed χ_1^2 with one degree of freedom, 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 $H_a : E[I_t(1-q)] = \frac{T_1}{T_w} \neq (1-q)$. The VaR model is rejected if $LR_{POF} > \chi_1^2 = 3.841$. Alternatively, we can obtain a rejection region $[x_1, x_2]$ by equating Eqn.(50) to χ_1^2 and solving for T_1. The VaR model is rejected if $T_1 \notin [x_1, x_2]$ and accepted if $T_1 \in [x_1, x_2]$.

Christoffersen (1998) 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

$$LR_{IND} = -2ln\frac{(1-\pi)^{(T_{00}+T_{10})}\pi^{(T_{01}+T_{11})}}{(1-\pi_{01})^{T_{00}}\pi_{01}^{T_{01}}(1-\pi_{11})^{T_{10}}\pi_{11}^{T_{11}}}, \tag{51}$$

asymptotically distributed χ_1^2 with one degree of freedom, where T_{ij}, with $i, j = 0(noviolation), 1(violation)$, is the number of observed events with the jth event following i^{th}, and π_{01}, π_{01} and π are estimates of the probabilities of $T_{i,j}$ (Argyropoulos and Panopoulou 2017). The model is rejected for the independent property if $LR_{IND} > \chi_1^2 = 3.841$. Christoffersen's conditional coverage test is a joint test of Kupiec's POF test and the IND test that tests both properties of UC and IND instantaneously. The conditional coverage test has likelihood ratio

$$LR_{CC} = LR_{POF} + LR_{IND}, \tag{52}$$

asymptotically distributed χ_2^2 with two degree of freedom. The hypothesis is $P[\mathbb{I}_t(1-q) = 1|\Omega_{t-1}] = 1-q, \forall_t$ against $P[\mathbb{I}_t(1-q) = 1|\Omega_{t-1}] \neq 1-q, \forall_t$, where Ω_{t-1} is the information available on day $t-1$. The model is rejected for the conditional coverage property if $LR_{CC} > \chi_2^2 = 5.99$.

BCBS came up with a set of requirements that the VaR model must satisfy for it to be considered a reliable risk measure. 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 (Resti 2008). Table 6 summarises the acceptance region for the Basel "traffic light" approach to back-testing VaR models. In the red zone, the VaR model underestimates risk, which indicates that anything in this range should be rejected.

Table 6. Acceptance region for Basel "traffic light"
approach to back-testing VaR models.

Zone	Number of exceptions	Cumulative Probability
Green	≤ 4	89.22%
Yellow	5	95.88%
	6	98.63%
	7	99.60%
	8	99.89%
	9	99.97%
Red	≥ 10	99.99%

Note: $q = 99\%$, $T = 250$ (Jorion 2007).

We also consider the new independence test carried out by Santos and Alves (2012). The test statistics is defined as

$$T_{N,[N/2]} = \log 2 \frac{D_{N:N} - 1}{D_{[N/2]:N}} - \log N, \tag{53}$$

where $D_{1:N} \leq \ldots \leq D_{N:N}$ are the order statistics of durations $D_1, \ldots D_N$. $D_i = t_i - t_{t-1}$ is the duration between two consecutive exceptions and i is the time until the first exception. There are clustering of exceptions if the median of $D_{N:N}/D_{[N/2]:N}$ is higher than the median under the IDN hypothesis (see Santos and Alves (2012) for more details).

We divide the out-of-sample data into blocks of 250, 500, and 1000 trading days to observe how the model behaves for longer and shorter observation periods. The division of out-of-sample data is also employed to meet the BCBS requirements. Table 7 presents the expected and observed number of exceptions produced at $q = (1\%, 5\%)$ for a portfolio consisting of all five banks.

Table 7. Exceptions on the porfolio following
Bayesian-GARCH(1,1) C-vine copula-EVT VaR model.

	Expected no. of exceptions			Observed no. of exceptions		
q	250	500	1000	250	500	1000
99%	3	5	10	0	0	0
95%	13	25	50	7	16	32

Note: The out-of-sample data after the 2011 financial crisis is divided into blocks of 250, 500 and 1000 trading days (observation periods); time horizon = 1 day.

At the 99% confidence level, we expect to observe 3, 5, and 10 exceptions for 250, 500, and 1000 trading days, respectively. The model produced zero exceptions. The model does not underestimate risk; rather, it is assumed to be too *conservative*. With zero exceptions, the model passed the "traffic light" test and is placed in the green zone (see Table 6). Most financial institutions will prefer this model at 99% confidence level. In this context, routinely produced profit and loss (P&L) plots of financial institutions show no violation of their 99% confidence VaR over long periods, proclaiming that this supports their risk model. This reveals the amount of the economic capital banks currently hold is in excess of their regulatory capital. Thus, banks may prefer to report higher VaR numbers to avoid a possibility of regulatory intrusion (Jorion 2007).

Tables 8 presents back-testing results at $q = 95\%$. We expect to observe 13, 25, and 50 exceptions for 250, 500, and 1000 trading days, respectively. The model produced 7, 16, and 32, respectively, and performs better at shorter observation periods compared to longer observation periods. During longer observation periods, the model will have longer term memory and is not be easily affected by sudden changes in the underlying volatility. Volatility in financial markets fluctuates as time

passes with most recent price changes providing more information with regards to current volatility as compared to older price changes. Thus, shorter observation periods will be more responsive to changes in volatility than longer observation periods. These results confirmed the findings of Best (2000) where he showed that VaR at 95% was more effective with shorter observation periods.

Table 8. Back-testing results after the 2011 financial
crisis following Bayesian GARCH(1,1) C-Vine Copula
EVT VaR model.

Trading window	Exceptions	$\frac{T_1}{T_w}$	LR_{POF}	LR_{IND}	LR_{CC}	$T_{N,[N/2]}$	Test results
				Back-test type			
250 trading days	7	0.028	3.009 (0.083)	0.952 (0.329)	3.961 (0.138)	1.916 (0.448)	(AAAA)
500 trading days	16	0.032	3.888 (0.049)	0.011 (0.916)	3.899 (0.142)	-0.693 (0.882)	(RAAA)
1000 trading days	32	0.032	7.777 (0.005)	0.813 (0.367)	8.590 (0.014)	-0.693 (0.861)	(RARA)

Note: The analysis is based on 5% significance levels with p-values in parenthesis. The VaR model performs better at shorter observation periods compared to longer observation periods. A = Accept, R = Reject. Coverage rate = $\frac{T_1}{T_w} \approx 1 - q$.

The out of sample data was taken immediately after the 2011 financial crisis. We, therefore, use VaR estimates during this period to calculate MCR for market risk in accordance with Basel II (i.e. Eqn.(1a)). To comply with Basel III rules, we consider back-testing for the periods of January to December 2008 and January to December 2011. This constitutes a continuous 12-months crisis period for the 2008 global financial crisis and the 2011 European financial crisis, which gives an observation period of 262 trading days for 2008 and 260 trading days for 2011. VaR estimates during these periods are referred to as stress VaR (sVaR) estimates because they are calculated during a period of significant financial distress. Moreover, sVaR estimates are used to calculate MCR for market risk in accordance with Basel III rules (Eqn.(2)). As shown in Tables 9, back-testing results during these periods suggest that the VaR measure can capture the dynamics of volatility in periods of severe crisis.

Table 9. Back-testing results conducted separately for
the 2008 and 2011 financial crises periods at $p = 5\%$.

Trading window	Exceptions	$\frac{T_1}{T_w}$	LR_{POF}	LR_{IND}	LR_{CC}	$T_{N,[N/2]}$	Test results
				Back-test type			
262 trading days (2008)	16	0.061	0.633 (0.426)	3.918 (0.048)	4.551 (0.103)	1.802 (0.275)	(ARAA)
260 trading days (2011)	10	0.038	0.789 (0.374)	1.628 (0.202)	2.417 (0.299)	-0.322 (0.808)	(AAAA)

Note: The VaR model shows reliability in periods of severe crisis. A = Accept, R = Reject, Coverage rate = $\frac{T_1}{T_w} \approx 1 - q$. We expect to have 13 exceptions for the 2008 and 2011 crisis period. However, the model produced 16 exceptions for 2008 and 10 exceptions for 2011.

Table 10 shows MCR for market risk in accordance with Basel II rules, calculated from current VaR estimates; from February 2012 to December 2015, i.e., covering the out of sample data. MCR for market risk in accordance with Basel III rules are calculated from current VaR estimates for a continuous 12-month period of significant financial stress. We consider the 2008 and 2011 financial crises periods. Regulatory multiplier is fixed to $k = 3$ in all cases because at 99% confidence level we observe ≤ 4 exceptions.

27

Table 10. Minimum Capital Requirements (MCR) for
market risk in accordance with Basel II and Basel III.

VaR estimation period	Current VaR estimates $VaR(99\%)$	$\frac{1}{60}\sum_{i=1}^{60}(VaR(99\%))$	$k=3$ $\frac{k}{60}\sum_{i=1}^{60}(VaR(99\%))$	MCR Basel II	Basel III
Feb.2012 to Dec.2015	3.75%	3.98%	11.94%	11.94%	-
sVaR: Jan.2008 to Dec.2008	9.05%	15.24%	45.73%	45.73%	57.67%
sVaR: Jan.2011 to Dec.2011	6.42%	5.99%	19.27%	19.27%	31.21%

Allen et al. (2017) employed R-Vine and C-Vine copula functions to model dependencies between 10 major European stock markets using their individual market indices and the composite blue chip STOXX50 European index. Their conclusion was that the R-Vine approach gave potentially better results. In this light, we also test the performance of the VaR model using R-Vine copula functions. For more details on R-Vine copula functions, see Schellhase and Spanhel (2018), Barthel et al. (2018), Allen et al. (2017) and the complete R-package by Schepsmeier et al. (2018).

Table 11 presents the unconditional and conditional pair copula type selection following R-Vine copula functions. Figure 11, a plot of the estimated daily VaRs and daily profit and loss of the portfolio returns series, show that the VaR model is capable of capturing the dynamics in the return series. Back-testing results are presented in Tables 12, 13, and 14. We also observe zero exceptions at $q = 99\%$. At $q = 95\%$ confidence level and 250 observation period, the R-Vine copula VaR model produce exactly the same number of exceptions as the C-Vine copula VaR model. For 500 and 1000 observation periods, the number of exceptions produced is increased by one. None of the back-testing methods rejects the R-Vine copula VaR model for the 2008 and 2011 crisis periods, whereas Christoffersen's conditional coverage test rejects the C-Vine copula VaR model for the 2008 crisis period (see Tables 9 and 14). The MCR for market Risk following the R-Vine copula VaR model, presented in Table 15, is slightly lower than the MCR for market risk following C-Vine copula VaR model.

We use as a benchmark, VaR models constructed using well known GJR-GARCH(1,1) and sGARCH(1,1) (standard GARCH(1,1)) volatility models with skewed Student's-t distribution to compare the performance of the vine copula EVT VaR models. In reference to Eqn.(4), the PDF of a standardized Student's-t distribution, skewness is introduced by an additional parameter $\gamma_k > 0$ as defined in Fernández and Steel (1998); that is

$$p(\epsilon_k|v,\gamma_k) = \frac{2}{\gamma_k + \frac{1}{\gamma_k}}\left\{f_s\left(\frac{\epsilon_k}{\gamma_k}\right)\mathbb{I}_{[0,\infty)}(\epsilon_k) + f_s(\gamma_k\epsilon_k)\mathbb{I}_{(-\infty,0)}(\epsilon_k)\right\}. \tag{54}$$

When $\gamma_k \neq 1$, the posterior distribution, $p(\epsilon_k|v,\gamma_k)$ loses symmetry (see Trottier and Ardia (2016), Ardia et al. (2016), Fernández and Steel (1998) for more details on skewed Student's-t probability distribution). Results are presented in Tables 16 and 17. Comparing these results with the Bayesian GARCH(1,1) vine copula EVT VaR models, the GJR-GARCH(1,1) and sGARCH(1,1) VaR models also record zero exceptions at 99% confidence level. At 95% confidence level, we record the best performance at 250 observation periods, i.e, no rejection and the worst performance at 1000 observation periods for the four VaR models. The GJR-GARCH(1,1) VaR model performs the least at both 500 and 1000 observation periods. With exception of the crisis periods, the the sGARCH(1,1) VaR model seems to produce the same results as the Bayesian GARCH(1,1) C-Vine copula VaR model as seen in Table 18. We should keep in mind that the stocks used to construct these VaR models are the top performing stocks in the banking sector in UK. Thus, it is no surprise that the results are quite close.

Table 11. R-vine parameter estimates.

	Unconditional and conditional pairs copulas	Selected copula family	Parameter	Log-likelihood	paired AIC
$Tree_1$	$C_{3,1}$	Student's-t copula	$\rho = 0.6221\ v = 6.23$	16.4518	-30.90
	$C_{4,2}$	Student's-t copula	$\rho = 0.7150,\ v = 5.12$	25.7210	-49.44
	$C_{4,3}$	Student's-t copula	$\rho = 0.7197,\ v = 4.91$	174.1620	-344.32
	$C_{1,5}$	Student's-t copula	$\rho = 0.6384,\ v = 6.63$	758.9981	-1514.00
$Tree_2$	$C_{3,2\mid4}$	Student's-t copula	$\rho = 0.3736,\ v = 7.83$	38.6780	-73.36
	$C_{3,5\mid1}$	survival $BB8$ copula	$\Lambda = 6.0000,\ v = 0.34$	81.7662	-159.53
	$C_{4,1\mid3}$	Student's-t copula	$\rho = 0.2255,\ v = 10.74$	716.3524	-1428.70
$Tree_3$	$C_{2,1\mid4,3}$	Student's-t copula	$\rho = 0.1424,\ v = 11.03$	222.8991	-441.80
	$C_{4,5\mid3,1}$	Frank copula	$\lambda = 0.84$	1034.6563	-2065.31
$Tree_4$	$C_{2,5\mid4,3,1}$	Frank copula	$\lambda = 0.6694$	1055.084	-2106.168
Log-likelihood				3574.6050	

Note: The copula types for the decomposition of n-variate bivariate copulas for unconditional and conditional pairs are selected based on AIC values. That is, the paired copula with the smallest AIC value.
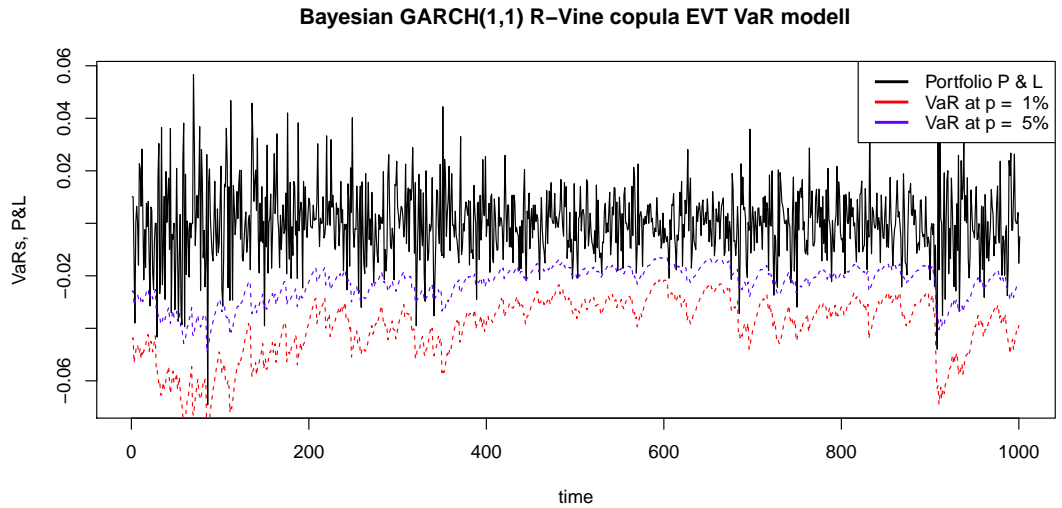
Figure 11. Estimated daily VaRs and profit and loss (P&L) plot for an investment in a portfolio consisting of all five banks following Bayesian-GARCH(1,1) with Student's-t distribution, R-vine copula functions, and EVT.

Table 12. Exceptions on the porfolio following
Bayesian-GARCH(1,1) R-vine copula-EVT VaR model.

	Expected no. of exceptions			Observed no. of exceptions		
q	250	500	1000	250	500	1000
99%	3	5	10	0	0	0
95%	13	25	50	7	17	33

Note: Time horizon = 1 day.

Table 13. Back-testing results after the 2011 financial
crisis following Bayesian GARCH(1,1) R-Vine Copula
EVT VaR model.

				Back-test type				
Trading window	Exceptions	$\frac{T_1}{T_w}$	LR_{POF}	LR_{IND}	LR_{CC}	$T_{N,[N/2]}$	Test results	
250 trading days	7	0.028	3.009	(0.083)	0.952 (0.329)	3.961 (0.138)	1.916 (0.449)	(AAAA)
500 trading days	17	0.034	3.021	(0.082)	0.057 (0.811)	3.078 (0.215)	1.029 (0.487)	(AAAA)
1000 trading days	33	0.033	6.878	(0.009)	0.976 (0.323)	7.854 (0.020)	1.009 (0.433)	(RARA)

Note: The analysis is based on 5% significance levels with p-values in parenthesis. The VaR model performs better at shorter observation periods compared to longer observation periods. A = Accept, R = Reject, Coverage rate = $\frac{T_1}{T_w} \approx 1 - q$.

Table 14. Back-testing results conducted separately for
the 2008 and 2011 financial crises periods at $p = 5\%$
following analysis with R-Vine copulas.

			Back-test type				
Trading window	Exceptions	$\frac{T_1}{T_w}$	LR_{POF}	LR_{IND}	LR_{CC}	$T_{N,[N/2]}$	Test results
262 trading days (2008)	17	0.065	1.122 (0.289)	1.118 (0.290)	2.240 (0.326)	1.530 (0.364)	(AAAA)
260 trading days (2011)	10	0.038	0.789 (0.374)	1.628 (0.202)	2.417 (0.299)	-0.322 (0.806)	(AAAA)

Note: The VaR model shows reliability in periods of severe crisis. A = Accept, R = Reject, Coverage rate = $\frac{T_1}{T_w} \approx 1 - q$. The model produced 17 exceptions for 2008 and 10 exceptions for 2011.

Table 15. Minimum Capital Requirements (MCR) for
market risk in accordance with Basel II and Basel III
following analysis with R-Vine.

	Current VaR estimates		$k = 3$	MCR	
VaR estimation period	$VaR(99\%)$	$\frac{1}{60}\sum_{i=1}^{60}(VaR(99\%))$	$\frac{k}{60}\sum_{i=1}^{60}(VaR(99\%))$	Basel II	Basel III
Feb.2012 to Dec.2015	3.60%	3.82%	11.47%	11.47%	-
sVaR: Jan.2008 to Dec.2008	8.70%	14.64%	43.92%	43.92%	55.39%
sVaR: Jan.2011 to Dec.2011	6.17%	5.76%	18.50%	18.50%	29.98%

Table 16. Back-testing results following sGARCH(1,1)
and GJR-GARCH(1,1) models with skewed student's-t
errors at 5% significance level

sGARCH				Back-test type				
Window	Exceptions	$\frac{T_1}{T_w}$	LR_{UC}	LR_{IND}	LR_{CC}	$T_{N,[N/2]}$	Test results
250	7	0.028	3.009 (0.083)	0.962 (0.327)	3.971 (0.137)	1.916 (0.448)	(A A A A)
500	16	0.032	3.888 (0.049)	0.011 (0.916)	3.899 (0.142)	-0.069 (0.708)	(R A A A)
1000	32	0.032	7.777 (0.005)	0.007 (0.933)	7.784 (0.020)	0.416 (0.541)	(R A R A)
GJR-GARCH(1,1)				Back-test type				
Window	Exceptions	$\frac{T_1}{T_w}$	LR_{UC}	LR_{IND}	LR_{CC}	$T_{N,[N/2]}$	Test results
250	7	0.028	3.009 (0.083)	0.962 (0.327)	3.971 (0.137)	4.119 (0.228)	(A A A A)
500	14	0.028	6.018 (0.014)	0.032 (0.858)	6.050 (0.049)	-0.773 (0.908)	(R A R A)
1000	26	0.026	14.597 (0.000)	0.263 (0.608)	14.860 (0.001)	-0.975 (0.928)	(R A R A)

Note: p-values in parenthesis.

Table 17. Back-testing results following sGARCH(1,1)
and GJR-GARCH(1,1) models with skewed student's-t
errors during the 2008 and 2011 financial crisis.

	sGARCH				Back-test type				
p	Window	Exceptions	$\frac{T_1}{T_w}$	LR_{UC}	LR_{IND}	LR_{CC}	$T_{N,[N/2]}$	Test results
1%	262 trading days (2008)	0	0.000	-	-	-	-	(- - - -)
	260 trading days (2011)	0	0.000	-	-	-	-	(- - - -)
5%	262 trading days (2008)	16	0.061	0.633 (0.426)	0.454 (0.500)	1.087 (0.581)	2.946 (0.148)	(A A A A)
	260 trading days (2011)	10	0.038	0.789 (0.374)	1.615 (0.204)	2.404 (0.301)	-0.322 (0.808)	(A A A A)
	GJR-GARCH(1,1)				Back-test type				
p	Window	Exceptions	$\frac{T_1}{T_w}$	LR_{UC}	LR_{IND}	LR_{CC}	$T_{N,[N/2]}$	Test results
1%	262 trading days (2008)	0	0.000	-	-	-	-	(- - - -)
	260 trading days (2011)	0	0.000	-	-	-	-	(- - - -)
5%	262 trading days (2008)	14	0.053	0.064 (0.800)	0.110 (0.740)	0.174 (0.917)	1.173 (0.392)	(A A A A)
	260 trading days (2011)	8	0.030	2.332 (0.127)	1.017 (0.313)	3.349 (0.187)	-0.539 (0.990)	(A A A A)

Note: p-values in parenthesis.

Table 18. Exceptions produced for a time horizon of
one day.

		Observed number of exceptions: VaR Model			
p	Trading window	B-GARCH(1,1) EVT C-Vine	B-GARCH(1,1) EVT R-Vine	GJR-GARCH(1,1)	sGARCH(1,1)
1%	250	0	0	0	0
	500	0	0	0	0
	1000	0	0	0	0
	262	0	0	0	0
	260	0	0	0	0
5%	250	7	7	7	7
	500	16	17	14	16
	1000	32	33	26	32
	262	16	17	16	14
	260	10	10	10	8

Note: We expect to observe 3, 5, and 10 exceptions for 250, 500, and 1000 trading days, respectively, and 13, 25, and 50 exceptions for 250, 500, and 1000 trading days, respectively.

6. Conclusion

VaR is the most common risk measure used by financial institutions to assess market risk of financial assets. Because VaR models often focus on the behaviour of asset returns in the left tail, the models must be calibrated in such a way that they do not underestimate or overestimate the proportion of outliers, as this will have significant effects on the allocation of the economic capital of investments. To implement a reliable VaR model, the chosen time horizon and type of volatility model are very important. We construct our VaR model by combining a Bayesian-GARCH(1,1) model with Student's-t distribution as the underlying volatility model, vine copula functions to model dependence, and EVT to model the left tail, thereby creating the Bayesian-GARCH(1,1) vine copula-EVT VaR models. Back-testing results show that the VaR models are reliable for forecasting risk of financial assets in periods when the market is relatively calm and periods of severe financial crisis.

Looking at Tables 10 and 15, the MCRs for market risk in relation to Basel II are almost three times the maximum loss per day and much higher in relation to Basel III. This results suggest that with the correct VaR model, Basel III is not needed. Moreover, these results also confirmed the previous findings by McAleer et al. (2011) that indicated that the global financial crisis could not be associated with the failure of Basel II since it was implemented in Europe prior to 2008 and never in the United States. Banks that displayed higher solvency ratios and higher credit-to-GDP ratios prior to their collapse or bailouts, probably manipulated their internal risk models for market risk to show positive results, or because of poor VaR models that were unable to capture the fat-tail risk. However, this assertion is not proven as the model needs to be tested in countries where banks were severely affected by the crisis such as Greece, which was in recession for 63 months from the third quarter of 2008 to the second quarter of 2014 and 27 months from the first quarter of 2015 to the first quarter of 2017 (DAmuri and Peri 2014, Jenkins et al. 2012, Lin et al. 2012).

It is also important to draw attention to the fact that Eqn.(32) depends on the threshold and the number of points above the threshold (i.e. exceedances) because the parameters are estimated based on the exceedances. Thus, it is logical to say that the reliability of Eqn.(32) rests solely on the choice of the subjective thresholds. Adopting the proposed *hybrid* method for the threshold selection reduces the possibility of selecting varying thresholds by different analysts on the same data, which can lead to different VaR estimates.

Statistical tools

We use R-software to estimate Bayesian GARCH(1,1) parameters, vine copula functions, and EVT analysis. GARCH(1,1) parameters estimated using the bayesGarch package, vine copula functions are estimated using the VineCopula and CDVine packages, and finally the POT method of EVT is estimated using the evir package.

References

Aas, K., Czado, C., Frigessi, A., Bakken, H., 2009. Pair-copula constructions of multiple dependence. Insurance: Mathematics and economics 44 (2), 182–198.

Allen, D. E., McAleer, M., Singh, A. K., 2017. Risk measurement and risk modelling using applications of vine copulas. Sustainability 9 (10), 1762.

Ardia, D., 2008. Financial risk management with Bayesian estimation of GARCH models: Theory and applications. Vol. 612 of Lecture notes in economics and mathematical systems. Springer-Verlag, Berlin, Germany.

Ardia, D., 2015. Bayesian Estimation of the GARCH(1,1) Model with Student-t Innovations in R. R package version R-3.3.1.

Ardia, D., Bluteau, K., Boudt, K., Trottier, D.-A., 2016. Markov-switching GARCH models in R: The MSGARCH package. R package version R-3.3.1.

Ardia, D., Hoogerheide, L. F., 2010. Bayesian estimation of the garch (1, 1) model with student-t innovations. The R Journal 2 (2), 41–47.

Argyropoulos, C., Panopoulou, E., 2017. Measuring the market risk of freight rates: A forecast combination approach. Journal of Forecasting.

Barthel, N., Geerdens, C., Killiches, M., Janssen, P., Czado, C., 2018. Vine copula based likelihood estimation of dependence patterns in multivariate event time data. Computational Statistics & Data Analysis 117, 109–127.

Basel Committee on Banking Supervision, 1996. Supervisory framework for the use of" backtesting" in conjunction with the internal models approach to market risk capital requirements. Bank for International Settlements.

Bedford, T., Cooke, R. M., 2001. Probability density decomposition for conditionally dependent random variables modeled by vines. Annals of Mathematics and Artificial intelligence 32 (1), 245–268.

Berkowitz, J., Christoffersen, P., Pelletier, D., 2011. Evaluating value-at-risk models with desk-level data. Management Science 57 (12), 2213–2227.

Best, P., 2000. Implementing value at risk. John Wiley & Sons, Chichester, United Kingdom.

Bhattacharyya, M., Ritolia, G., 2008. Conditional var using evt–towards a planned margin scheme. International Review of Financial Analysis 17 (2), 382–395.

Bob, N. K., 2013. Value at risk estimation. a garch-evt-copula approach. Mathematiska institutionen, 1–41.

Bollerslev, T., 1986. Generalized autoregressive conditional heteroskedasticity. Journal of econometrics 31 (3), 307–327.

Carmassi, J., Micossi, S., 2012. Time to set banking regulation right. Centre for European Policy Studies (CEPS), Brussels, Belgium.

Cherubini, U., Luciano, E., 2001. Value-at-risk trade-off and capital allocation with copulas. Economic notes 30 (2), 235–256.

Cherubini, U., Luciano, E., Vecchiato, W., 2004. Copula methods in finance. John Wiley & Sons, Chichester.

Cherubini, U., Mulinacci, S., Gobbi, F., Romagnoli, S., 2011. Dynamic Copula methods in finance. Vol. 625. John Wiley & Sons, Chichester.

Christoffersen, P. F., 1998. Evaluating interval forecasts. International economic review, 841–862.

Clayton, D. G., 1978. A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence. Biometrika 65 (1), 141–151.

DAmuri, F., Peri, G., 2014. Immigration, jobs, and employment protection: evidence from europe before and during the great recession. Journal of the European Economic Association 12 (2), 432–464.

Danielsson, J., de Haan, L., Peng, L., de Vries, C. G., 2001. Using a bootstrap method to choose the sample fraction in tail index estimation. Journal of Multivariate analysis 76 (2), 226–248.

Danielsson, J., de Vries, C., 1998. Beyond the sample: Extreme quantile and probability estimation. Tinbergen Institute Discussion Paper No. 98-016/2.

Danielsson, J., De Vries, C. G., 2000. Value-at-risk and extreme returns. Annales d'Economie et de Statistique, 239–270.

Davison, A. C., Smith, R. L., 1990. Models for exceedances over high thresholds. Journal of the Royal Statistical Society. Series B (Methodological), 393–442.

Deschamps, P. J., 2006. A flexible prior distribution for markov switching autoregressions with student-t errors. Journal of Econometrics 133 (1), 153–190.

Embrechts, P., Lindskog, F., McNeil, A., 2001. Modelling dependence with copulas. Rapport technique, Département de mathématiques, Institut Fédéral de Technologie de Zurich, Zurich.

Embrechts, P., McNeil, A., 1999. D. straumann [1999], correlation and dependency in risk management: properties and pitfalls, departement of mathematik, ethz. Tech. rep., Zürich, Working Paper.

Embrechts, P., McNeil, A., Straumann, D., 2002. Correlation and dependence in risk management: properties and pitfalls. Risk management: value at risk and beyond, 176–223.

Fantazzini, D., 2008. Dynamic copula modelling for value at risk. Frontiers in Finance and Economics 5 (2), 72–108.

Fernández, C., Steel, M. F., 1998. On bayesian modeling of fat tails and skewness. Journal of the American Statistical Association 93 (441), 359–371.

Frank, M. J., 1979. On the simultaneous associativity of f(x,y) and x+y-f(x,y). Aequationes mathematicae 19 (1), 194–226.

Frey, R., McNeil, A. J., McNeil, A. J., McNeil, A. J., 2001. Modelling dependent defaults. ETH, Eidgenössische Technische Hochschule Zürich, Department of Mathematics.

Gelman, A., Rubin, D. B., 1992. Inference from iterative simulation using multiple sequences. Statistical science, 457–472.

Geweke, J., 1993. Bayesian treatment of the independent student-t linear model. Journal of Applied Econometrics 8 (S1), S19–S40.

Ghalanos, A., 2015. The rmgarch models: Background and properties.(version 1.2-8). URL http://cran. r-project. org/web/packages/rmgarch/index. html.

Gilli, M., et al., 2006. An application of extreme value theory for measuring financial risk. Computational Economics 27 (2-3), 207–228.

Goodhart, C., 2011. The Basel Committee on Banking Supervision: a history of the early years 1974–1997. Cambridge University Press.

Greenberg, E., 2012. Introduction to Bayesian econometrics. Cambridge University Press.

Gumbel, E. J., 1960. Bivariate exponential distributions. Journal of the American Statistical Association 55 (292), 698–707.

Hall, P., Yao, Q., 2003. Inference in arch and garch models with heavy–tailed errors. Econometrica 71 (1), 285–317.

Hill, B. M., et al., 1975. A simple general approach to inference about the tail of a distribution. The annals of statistics 3 (5), 1163–1174.

Holton, G. A., 2014. Value-at-risk: theory and practice. e-book at http://value-at-risk.net.

Hsu, C.-P., Huang, C.-W., Chiou, W.-J. P., 2012. Effectiveness of copula-extreme value theory in estimating value-at-risk: empirical evidence from asian emerging markets. Review of Quantitative Finance and Accounting 39 (4), 447–468.

Huang, J.-J., Lee, K.-J., Liang, H., Lin, W.-F., 2009. Estimating value at risk of portfolio by conditional copula-garch method. Insurance: Mathematics and economics 45 (3), 315–324.

International Monetary Fund, 2009. Global financial stability report. World economic and financial surveys. International Monetary Fund.

Jenkins, S. P., Brandolini, A., Micklewright, J., Nolan, B., 2012. The great recession and the distribution of household income. Oxford University Press, Oxford OX2 6DP, United Kingdom.

Jorion, P., 2007. Value at risk: the new benchmark for managing financial risk. Vol. 3. McGraw-Hill New York.

Krämer, N., Schepsmeier, U., 2011. Introduction to vine copulas. In: NIPS Workshop, Granada.

Kupiec, P. H., 1995. Techniques for verifying the accuracy of risk measurement models. The journal of Derivatives 3 (2), 73–84.

Li, D. X., 2000. On default correlation: A copula function approach. The Journal of Fixed Income 9 (4), 43–54.

Lin, C. Y.-Y., Edvinsson, L., Chen, J., Beding, T., 2012. National intellectual capital and the financial crisis in Greece, Italy, Portugal, and Spain. Vol. 7. Springer Science & Business Media, London, United Kingdom.

Malz, A. M., 2011. Financial risk management: Models, history, and institutions. Vol. 538. John Wiley & Sons.

McAleer, M., Jimenez-Martin, J.-A., Perez Amaral, T., 2011. Has the basel ii accord encouraged risk management during the 2008-09 financial crisis? KIER Discussion Paper No. 767.

McNeil, A. J., Frey, R., 2000. Estimation of tail-related risk measures for heteroscedastic financial time series: an extreme value approach. Journal of empirical finance 7 (3), 271–300.

Patton, A. J., 2004. On the out-of-sample importance of skewness and asymmetric dependence for asset allocation. Journal of Financial Econometrics 2 (1), 130–168.

Resti, A., 2008. Pillar II in the new Basel Accord : the challenge of economic capital. Risk Books, London.

Rossignolo, A. F., Fethi, M. D., Shaban, M., 2012. Value-at-risk models and basel capital charges: Evidence from emerging and frontier stock markets. Journal of Financial Stability 8 (4), 303–319.

Sampid, M. G., Hasim, H. M., 2018. Estimating value-at-risk using a multivariate copula-based volatility model: Evidence from european banks. International Economics 165.

Santos, P. A., Alves, M. F., 2012. A new class of independence tests for interval forecasts evaluation. Computational Statistics & Data Analysis 56 (11), 3366–3380.

Schellhase, C., Spanhel, F., 2018. Estimating non-simplified vine copulas using penalized splines. Statistics and Computing 28 (2), 387–409.

Schepsmeier, U., Brechmann, E., 2015. Package cdvine: Statistical inference of c-and d-vine copulas. R-Project CRAN Repository.

Schepsmeier, U., Stoeber, J., Brechmann, E. C., Graeler, B., Nagler, T., Erhardt, T., Almeida, C., Min, A., Czado, C., Hofmann, M., et al., 2018. Package vinecopula.

Sheikh, A. Z., Qiao, H., 2010. Non-normality of market returns: A framework for asset allocation decision making (digest summary). Journal of Alternative Investments 12 (3), 8–35.

Silva Filho, O. C., Ziegelmann, F. A., Dueker, M. J., 2014. Assessing dependence between financial market indexes using conditional time-varying copulas: applications to value at risk (var). Quantitative Finance 14 (12), 2155–2170.

Sklar, M., 1959. Fonctions de répartition à n dimensions et leurs marges. Université Paris 8.

Soltane, H. B., Karaa, A., Bellalah, M., 2012. Conditional var using garch-evt approach: Forecasting volatility in tunisian financial market. Journal of Computations & Modelling 2 (2), 95–115.

Trottier, D.-A., Ardia, D., 2016. Moments of standardized fernandez–steel skewed distributions: Applications to the estimation of garch-type models. Finance Research Letters 18, 311–316.

Tsay, R. S., 2005. Analysis of financial time series. Vol. 543. John Wiley & Sons.

Tsay, R. S., 2013. Multivariate Time Series Analysis: With R and Financial Applications. John Wiley & Sons.

Tsay, R. S., 2014. An introduction to analysis of financial data with R. John Wiley & Sons.

Turner, A., et al., 2009. The Turner Review: A regulatory response to the global banking crisis. Vol. 7. Financial Services Authority London.

Uylangco, K., Li, S., 2016. An evaluation of the effectiveness of value-at-risk (var) models for australian banks under basel iii. Australian Journal of Management 41 (4), 699–718.

Virbickaite, A., Ausín, M. C., Galeano, P., 2015. Bayesian inference methods for univariate and multivariate garch models: A survey. Journal of Economic Surveys 29 (1), 76–96.

Vuong, Q. H., 1989. Likelihood ratio tests for model selection and non-nested hypotheses. Econometrica: Journal of the Econometric Society, 307–333.

Wong, M. C., 2013. Bubble Value at Risk: A Countercyclical Risk Management Approach. John Wiley & Sons.

Yan, J., et al., 2007. Enjoy the joy of copulas: with a package copula. Journal of Statistical Software 21 (4), 1–21.