Bayesian Tail Risk Forecasting using Realised Volatility DCC-Copula-GARCH Models

Christian Contino
*Discipline of Business Analytics, University of Sydney, Australia.

Richard H. Gerlach
Discipline of Business Analytics, University of Sydney, Australia.

*Correspondence address: Christian Contino, Merewether Building (H04), Discipline of Business Analytics, The University of Sydney, Sydney, NSW, 2006 Australia (christianco@econ.sydney.edu.au)
+61 (0)435 785 084
Abstract

A Skewed Student-t Realised DCC copula model using Realised Volatility GARCH marginal functions is developed within a Bayesian framework for the purpose of forecasting portfolio Value at Risk and Conditional Value at Risk. The use of copulas is implemented so that the marginal distributions can be separated from the dependence structure to produce tail forecasts. This is compared to using traditional GARCH-copula models, and GARCH on an aggregated portfolio, with weekly returns of five financial assets spanning from March 1971 to May 2014. An initial sample of 773 weeks is used to estimate the models, while 1000 one-week ahead forecasts are produced to compare out-of-sample forecast performance. Copula models implementing a Realised Volatility GARCH framework show an improvement over traditional GARCH models over a variety of formal and informal tests.

Keywords: Copula, Realised Volatility, GARCH, Value-at-Risk, Expected Shortfall, Skewed Student-t, High-Frequency Data, Risk Management.

1 Introduction

The past 30 years in the financial markets have seen a diverse mix of behaviour, including much prosperity, but also much financial uncertainty due to fast and deep crises in many international markets and assets. This has given rise to a large number of research initiatives in the subject, by academia, financial institutions and especially regulators who strive to attain better risk and forecast indicators (see Jorion, 2007, Crouhy, Mark, & Galai, 2001 Dowd, 1998). One of the most commonly used risk measure is the Value-at-Risk ("VaR"), which has gained much attention in the business community due to its relatively easy estimation and comprehension by a non-technical audience, especially following its introduction by official institutions such as the Basel Committee on Banking supervision (Basel Committe on Banking Supervision (1998)), and the Bank for International Settlements ("BIS") (Fisher-Report (1994)).

Embrechts (2000) and McNeil, Frey, and Embrechts (2005) provide an overview of the shortcomings of VaR, chiefly not being sub-additive, so that VaR is not a coherent risk measure, as the total risk of a portfolio can be larger than the sum of the individual risks (Artzner, Delbaen, Eber, and Heath (1997) and Artzner, Delbaen, Eber, and Heath (1999)). In this paper we explore the multivariate VaR approach with the use of copula models, which allow for a flexible multivariate distribution with different margins and different dependence structure, without the usual limitations of traditional joint-normal distributions.

The use of GARCH models (Engle (1982) and Bollerslev (1986)) in the marginals, so called copula-GARCH models is not novel, and has previously been used to model the joint distributions of multiple asset returns (see Patton (2004, 2006a, 2006b) and Jondeau and Rockinger (2006)) and has been shown to significantly improve forecast accuracy. Our
work implements a Bayesian approach, by drawing from the posterior distribution of the parameters by using a 'Metropolis-Hastings' algorithm, building on Xu (2004) who applies a Markov Chain Monte Carlo (MCMC) algorithm, but only assumes a constant mean and i.i.d. returns for the marginal distribution of each of the financial assets’ returns. This was further extended by Kang (2011), who within a Bayesian framework, implements a GARCH specification for the marginals. As the marginals each have their own set of parameters, together with the copula function, the model can be hard to estimate directly due to the large number of parameters to be estimated. To overcome this issue, we adopt an 'Inference For Margins' (IFM) approach (Joe and Xu (1996) and Joe (2005)), where the marginals are estimated first, and then the copula parameters are estimated conditional on this, where (Smith, Gan, & Kohn, 2012) and Kang (2011) implements this within a Bayesian framework. A comprehensive overview of Bayesian approaches to copula modelling is discussed in Smith (2013).

This paper combines three strands of recent popular research, the first being the use of Realised Volatility GARCH models (Andersen, Bollerslev, Diebold, and Labys (2003), Hansen and Lunde (2005), Watanabe (2012), Hansen, Huang, and Shek (2012)) which have become popular in recent times as they provide superior forecasting performance.

The second is the use of time varying dynamic copulas, with extensive literature on the matter (see Dias and Embrechts (2004), Patton (2004, 2006b), Chen and Fan (2006), Jondeau and Rockinger (2006), Giacomini, Härdle, and Spokoiny (2009), Jin (2009), Lai, Chen, and Gerlach (2009). Hafner and Manner (2012), Härdle, Okhrin, and Okhrin (2010)), with all these approaches implementing a copula structure that has time-varying parameters driven by past realization of the underlying data generating process, with this paper utilising a realised version of the Dynamic Conditional Correlation (DCC) framework of Engle and Sheppard (2001) and based on preliminary work by Bauwens, Storti, and Violante (2012). Recent work by Jin and Maheu (2012) and Jin and Maheu (2014) model Realised Covariance and returns within a Bayesian framework and compare their results to multivariate GARCH models and existing Realised Covariance models.


The bulk of the above research tries to model the data’s skewness and kurtosis in various ways, but their methods do not extend to all parts of their framework. While Jondeau and Rockinger (2006) use marginals employing a Skewed Student-t distribution, this does not extend to their copula, which is the opposite approach of Lai et al. (2009), in which their copula does account for asymmetry, but their marginal distribution does not. The use of Skewed Student-t copulas is the third strand, as it has only emerged in recent literature, with the arrival of multivariate Skewed Student-t distributions (Jones

By employing the univariate and multivariate Skewed Student-t framework of Fernández and Steel (1998) and Bauwens and Laurent (2005) respectively as one is the natural extension of the other, a new Realised Volatility Skewed Student-t copula-GARCH model is proposed, which is expected to model skewness and asymmetry in both the marginal and dependence structures.

We explore the use of high frequency data to improve the forecasting performance of tail estimates of a portfolio, by modelling the dependence structure of its constituents using a dynamic copula, and compare this to using traditional GARCH models also within a copula. This is also compared to a return series of the portfolio as a whole, once again using high-frequency data and weekly returns. This combination of high and low frequency data to improve the modelling of low frequency returns is similar in nature to research by Ghysels, Santa-Clara, and Valkanov (2004), Chen, Ghysels, and Wang (2011), Engle and Gallo (2006), Shephard and Sheppard (2010), Hansen et al. (2012) and Hansen, Lunde, and Voev (2014).

The main results are: the use of Realised Volatility GARCH as marginals results in a better forecasting performance than their low frequency counterpart, with the choice of copula distribution not being a significant factor. RV GARCH copulas also outperformed models using the return series of the single portfolio for certain tests, showing that modelling the dynamic of the individual assets is worthwhile.

The paper is structured in the following way; Section 2 specifies the dynamic copula model and Realised Volatility GARCH model. Section 3 gives a brief overview of the Bayesian approach and Markov-Chain Monte Carlo methods employed, Section 4 gives an overview of the forecasting methodologies employed and the relevant back testing, Section 5 presents the data and the empirical studies from five international markets for both Value-at-Risk and Conditional Value-at-Risk using formal and informal tests. Concluding remarks and possible extensions are given in Section 6. An appendix and tables with parameter estimation results follow.

2 Conditional Multivariate Modelling

2.1 Data Synchronisation

As discussed by Sheppard (2013), the issue of synchronisation is vital when using a cross-section of financial returns from international markets, as returns will be non-synchronous due to closing hour differences, local market closures due to public holidays, opening/closing delays and for certain assets, illiquidity and stale prices. To
counteract this problem, we extend the frequency of data from daily to weekly, which increases the period being synchronised from 30% to 86% when comparing New-York and London markets. As not all markets are open on the same day, the data needs to be cleaned by aligning all data sets, and removing days in which any of the asset’s markets are closed, and then summing the returns of five consecutive days to create a ’pseudo-week’ as detailed in Figure (1). The remaining daily returns are used to construct the Realised Volatility as detailed in Section 2.2.

2.2 Realised Volatility GARCH Model specifications

For a log return series $r_t$, we define the Realised Volatility (RV) over the time period $[t-j, t]$, for $0 < j \leq t \leq T$ as:

$$RV(t, j; m) \equiv \sum_{j=1}^{m} r(t - (j - 1)/m)^2,$$

where $m$ is the number of daily observations and $1/m$ is the sampling frequency. Our sample consists of 9,747 days, which translates into 1,949 weekly Realised Volatility data points. We define Realised Covariance (RCov) for assets $i$ to $k$ over the time period
\[ t - j, t \], for \( 0 < j \leq t \leq T \) as:

\[
RCov_{i,k}(t, j; m) \equiv \sum_{j=1}^{m} r_{i,k}(t - (j - 1)/m)r_{i,k}(t - (j - 1)/m)',
\]

where the diagonal elements in (2) are equal to (1).

The RV GARCH(1,1) model employing a simple log-linear specification is defined by the following volatility and measurement equation:

\[
\begin{align*}
\alpha_t &= r_t - \mu, \\
\alpha_t &= \sigma_t \epsilon_t, \\
\epsilon_t &\overset{\text{iid}}{\sim} t_\nu, \quad \text{or} \quad \epsilon_t &\overset{\text{iid}}{\sim} skt_{\nu, \lambda}, \\
\ln(\sigma_t^2) &= \alpha_0 + \alpha_1 \ln(RV_{t-1}) + \beta_1 \ln(\sigma_{t-1}^2) \\
\ln(RV_t) &= \xi + \phi \ln(\sigma_t^2) + \tau_1 \epsilon_t + \tau_2 (\epsilon_t^2 - 1) + u_t \\
u_t &\overset{\text{iid}}{\sim} t_\nu,
\end{align*}
\]

This is a similar specification as employed by Hansen et al. (2012) and Watanabe (2012) with the addition of a Student-t specification for the measurement equation in (4), which provides a coherent way to model the joint dependence between returns and the Realised Volatility measure.

We adopt a standardised Skewed Student-t distribution introduced by Fernández and Steel (1998) with takes the following form:

\[
skt(\epsilon | \nu, \lambda) = \begin{cases} \\
\frac{2\lambda}{1 + \lambda^2} \left[ f \left( \frac{\epsilon}{\lambda} \right) \right] & \text{if } \epsilon \geq 0 \\
\frac{2\lambda}{1 + \lambda^2} \left[ f(\lambda \epsilon) \right] & \text{if } \epsilon < 0 
\end{cases}
\]

where \( f \) is the standardised Student-t pdf:

\[
f_\nu(r | \mu, \sigma^2) = \frac{\Gamma\left(\frac{\nu+2}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)\sqrt{(\nu-2)\pi\sigma^2}} \left[ 1 + \frac{(r - \mu)^2}{(\nu - 2)\sigma^2} \right]^{\frac{\nu+1}{2}}
\]
by combining (5) and (6) we obtain the pdf:

\[
skt(\epsilon | \nu, \lambda) = \begin{cases} 
\frac{2 \lambda s \Gamma\left(\frac{\nu + 1}{2}\right)}{(1 + \lambda^2) \Gamma\left(\frac{\nu}{2}\right) \sqrt{(\nu - 2)\pi \sigma^2}} \left[1 + \frac{\lambda^{-2} (m \sigma + s \epsilon)^2}{(\nu - 2)\sigma^2}\right]^{-\frac{\nu + 1}{2}} & \text{if } \epsilon \geq -\frac{m \sigma}{s}, \\
\frac{2 \lambda s \Gamma\left(\frac{\nu + 1}{2}\right)}{(1 + \lambda^2) \Gamma\left(\frac{\nu}{2}\right) \sqrt{(\nu - 2)\pi \sigma^2}} \left[1 + \frac{\lambda^2 (m \sigma + s \epsilon)^2}{(\nu - 2)\sigma^2}\right]^{-\frac{\nu + 1}{2}} & \text{if } \epsilon < -\frac{m \sigma}{s}. 
\end{cases}
\]

(7)

where

\[
m = \frac{\Gamma\left(\frac{\nu - 1}{2}\right) \sqrt{\nu - 2}}{\Gamma\left(\frac{\nu}{2}\right) \sqrt{\pi}} \left(\lambda - \frac{1}{\lambda}\right), \quad s = \sqrt{\left(\lambda^2 + \frac{1}{\lambda^2} - 1 - m^2\right)}.
\]

\(\nu\) is the kurtosis parameter with range \(2 < \nu < \infty\) and \(\lambda\) is the skewness parameter with range \(0 < \lambda < \infty\), \(m\) and \(s\) are the mean and variance of the non-standardised Skewed Student-t distribution, and we standardise the distribution to have mean and variance \(\mu\) and \(\sigma^2\) respectively. Lambert and Laurent (2000) show that the direction of the skewness is determined by the sign of \(ln\lambda\), where the third moment is positive when if \(ln\lambda > 0\) and hence the density is skewed to the right and vice versa if \(ln\lambda < 0\).

2.3 Copulas

2.3.1 Copula Properties

A link between a joint distribution and a copula can be formulated by employing Sklar’s theorem (Sklar (1959)), in which an N-dimensional distribution \(F\) and marginal distribution \(F_i, i = \ldots, N\) there exists a copula \(C\), such that:

\[
F(x_1, \ldots, x_N) = C(F_1(x_1), \ldots, F_N(x_N)).
\]

(8)

The copula in (8) is unique given that the marginals employed are continuous, and from it we can obtain a formula for the copula:

\[
C(u_1, \ldots, u_N) = F(F_1^{-1}(u_1), \ldots, F_N^{-1}(u_N)),
\]

(9)

where

\[
u_i = F_i(x_i), \quad i = 1, \ldots, N.
\]
We can obtain the joint density if \( F \) is \( N \) times differentiable:

\[
\begin{align*}
  f(\mathbf{x}) &= \frac{\partial^N}{\partial x_1 \partial x_2 \ldots \partial x_N} F(\mathbf{x}) = \\
  &= \prod_{i=1}^{N} f_i(x_i) \frac{\partial^N}{\partial u_1 \partial u_2 \ldots \partial u_N} C(F_1(x_1), \ldots, F_N(x_N)) \\
  &= \prod_{i=1}^{N} f_i(x_i) c(F_1(x_1), \ldots, F_N(x_N))
\end{align*}
\]

with the copula density taking the form:

\[
  c(u_1, \ldots, u_N) = \frac{f(F_1^{-1}(u_1), \ldots, F_N^{-1}(u_N))}{\prod_{i=1}^{N} f_i(F_i^{-1}(u_i))} \tag{10}
\]

where \( \mathbf{x} = (x_1, \ldots, x_N) \) and \( c \) is the copula density. Evidence shows that the assumption of a joint normal distribution is not sufficient in capturing the behaviour and structural dependence of financial assets, as it has been widely observed that correlation is higher during downturns (Patton (2004)), so features such as asymmetric dependence and non-zero tail dependence can be modelled via copulas. This paper will only deal with three classes of copulas, Gaussian, Student-t and Skewed Student-t, and not deal with Archimedean copulas or the use of vines for higher dimensions.

By applying a Bayesian approach, we can formulate the cumulative distribution using copulas as follows:

\[
\begin{align*}
  F(r_1,t, \ldots, r_N,t | r_1,t-1, \ldots, r_N,t-1; \theta_1, \ldots, \theta_N, \delta) &= C(u_1,t(r_1,t-1, \theta_1), \ldots, u_N,t(r_N,t-1, \theta_N)|\delta) \tag{11}
\end{align*}
\]

where \( r_1,t, \ldots, r_N,t \) are excess returns of \( N \) financial assets at time \( t \), \( r_1,t-1, \ldots, r_N,t-1 \) are vectors of past excess return up to time \( t - 1 \). \( \theta_1, \ldots, \theta_N \) are the set of parameters of the marginal distributions from section (2.2) and \( \delta \) is the parameters of the copula being used. \( u_1, \ldots, u_N \) are the cumulative distributions of the respective excess returns.

To derive the joint density function we differentiate (11) w.r.t each excess return:

\[
\begin{align*}
  f(r_1,t, \ldots, r_N,t | r_1,t-1, \ldots, r_N,t-1; \theta_1, \ldots, \theta_N, \delta) &= c(u_1,t(r_1,t-1, \theta_1), \ldots, u_N,t(r_N,t-1, \theta_N)|\delta) \\
  &\times f_1(r_1,t|\xi_{1,t-1}, \theta_1) \ldots \times f_N(r_N,t|\xi_{N,t-1}, \theta_N) \tag{12}
\end{align*}
\]

where \( c \) is the density form of the copula being used and \( f_1, \ldots, f_N \) are the marginal densities of the respective asset returns.
2.3.2 Inference for the Margins

Joe and Xu (1996) proposed an Inference for the Margins (IFM) method in which the marginal distribution and copula parameters are estimated separately in a two-step process, as the number of parameters can be very large even in moderate dimensions. Joe (1997) shows that the IFM method is consistent and has the property of asymptotic normality under regular conditions, while Patton (2006a) shows that this two-step method yields asymptotically efficient and normal parameter estimates. The procedure is fairly straightforward, in which the marginal parameters in section (2.2) are estimated at the first step by a Bayesian MCMC procedure (as per section (3)) independently of each other, following which the copula parameters are optimised, conditional of the results obtained in the first step. Kang (2011) shows that the joint density given (12) and all observations can be written as:

\[
f(r_{1,0}, \ldots, r_{N,0}, \ldots, r_{1,T}, \ldots, r_{N,T}|\theta_1, \ldots, \theta_N, \delta)
= f(\xi_{1,0}, \ldots, \xi_{N,0}|\theta_1, \ldots, \theta_N, \delta)
\times \prod_{t=1}^T f(r_{1,t}, \ldots, r_{N,t}|\xi_{1,t-1}, \ldots, \xi_{N,t-1}; \theta_1, \ldots, \theta_N, \delta). \quad (13)
\]

The log-likelihood can be derived by omitting the first term:

\[
\mathcal{L}(\theta_1, \ldots, \theta_N, \delta) \approx \sum_{t=1}^T \log f(r_{1,t}, \ldots, r_{N,t}|\xi_{1,t-1}, \ldots, \xi_{N,t-1}; \theta_1, \ldots, \theta_N, \delta).
\]

the log-likelihood can be written as follows by using (12) and substitution:

\[
\mathcal{L}(\theta_1, \ldots, \theta_N, \delta) = \sum_{t=1}^T \log c(u_{1,t}(r_{1,t}|\xi_{1,t-1}, \theta_1), \ldots, u_{N,t}(r_{N,t}|\xi_{N,t-1}, \theta_N)|\delta)
+ \sum_{t=1}^T \sum_{i=1}^N \log f_i(r_{i,t}|\xi_{i,t-1}; \theta_i).
\]

The IFM two-step procedure can therefore be broken down so that each \( \theta_i \) is estimated separately, and then plugged into the copula in step two to estimate the copula parameters \( \delta \), which Ausin and Lopes (2010) have demonstrated to be valid in a Bayesian framework as follows:

**Step one:**

\[
\mathcal{L}(\hat{\theta}_i|r) = \sum_{t=1}^T \log f_i(r_{i,t}|\xi_{i,t-1}; \theta_i)
\]

for \( i = 1 \ldots N \)
Standardised residuals are then obtained as $z_{it} = \frac{r_{it} - \mu_i}{\sqrt{h_{it}}}$, following which $u_{it} = \text{CDF}(z_{it})$ so that $u \in [0, 1]$ using the CDF for the respective marginal distribution.

**Step two:**

$$L(\hat{\delta} | u t) = \sum_{t=1}^{T} \log c(u_{1,t}(r_{1,t} | r_{1,t-1}, \theta_1), \ldots, u_{N,t}(r_{N,t} | r_{N,t-1}, \theta_N) | \delta)$$

This paper utilises the IFM method detailed above, which while being less efficient, it is computationally benign and simple to implement. Härdle, Hautsch, and Overbeck (2008) give a comprehensive discussion of alternative methods to estimate copula models and their parameters.

### 2.3.3 Copula Models

This paper will look at the two most commonly used specifications for copulas, the Normal and the Student-t, together with a more flexible approach via the Skewed Student-t which will be discussed below. By combining the multivariate Normal and Student-t distribution PDFs given in (14) and (15) with (9) and (10) we can obtain their respective copulas.

**Normal Copula**

$$N(x | \mu, \Sigma) = (2\pi)^{-N/2} |\Sigma|^{-1/2} \exp \left[ -\frac{1}{2}(x - \mu)^\prime \Sigma^{-1}(x - \mu) \right]$$ (14)

$$c(u_{1,t}, \ldots, u_{N,t}) = |R_t|^{-\frac{1}{2}} \exp \left[ -\frac{1}{2} x_t' (R_t^{-1} - I_N) x_t \right]$$ (17)

where $\Phi$ is the multivariate normal CDF with zero mean vector, $\phi^{-1}$ is the univariate normal inverse CDF, $x_t = (x_{1,t}, \ldots, x_{N,t})$ and $x_{i,t} = \phi^{-1}(u_{i,t})$, $R_t = [p_{ij,t}]$ is a correlation matrix and $I_N$ is a $N \times N$ identity matrix.

**Student-t Copula**
\[ C(u_{1,t}, \ldots, u_{N,t}) = T_\nu(T_\nu^{-1}(u_{1,t}), \ldots, T_\nu^{-1}(u_{N,t}))[0, R_t], \]
\[ c(u_{1,t}, \ldots, u_{N,t}) = \frac{|R_t|^{-\frac{3}{2}} \Gamma\left(\nu + \frac{N}{2}\right) \left[\Gamma\left(\frac{\nu}{2}\right)\right]^{N-1} \left(1 + \frac{1}{\nu} x_i' R_t^{-1} x_i\right)^{-\frac{\nu + N}{2}}}{\left[\Gamma\left(\nu + \frac{1}{2}\right)\right] \prod_{i=1}^{N} \left(1 + \frac{x_{ii}^2}{\nu}\right)^{-\frac{\nu + 1}{2}}} \]

where \( T_\nu \) is the multivariate Student-t CDF with zero mean vector, \( T_\nu^{-1} \) is the univariate Student-t inverse CDF, \( x_t = (x_{1,t}, \ldots, x_{N,t}) \) and \( x_{i,t} = T_\nu^{-1}(u_{i,t}) \).

**Skewed Student-t Copula**

We use the multivariate Skewed Student-t distribution proposed by Bauwens and Laurent (2005) which is a natural extension of the univariate distribution of Fernández and Steel (1998) described in section (2.2) and its PDF \( \text{skt}_{\nu,\lambda} \) takes the following form:

\[ \text{skt}_{\nu,\lambda}(x|\mu, \Sigma) = \left(\frac{2}{\sqrt{\pi}}\right)^N \left(\prod_{i=1}^{N} \lambda_i s_i\right) \frac{\Gamma\left(\frac{\nu + N}{2}\right) |\Sigma|^{-\frac{3}{2}}}{\Gamma\left(\frac{\nu}{2}\right)^N} \left(1 + \frac{a'a}{\nu - 2}\right)^{-\frac{\nu + N}{2}} \]

where \( \nu \) and \( \lambda \) are the degree of freedom and skewness parameters respectively, with the elements of vector \( a = (a_1, \ldots, a_N) \) defined by \( a_i = \lambda_i^{-1}(m_i + s_i x_i^*) \) and \( x^* = (x_1^*, \ldots, x_N^*) = \Sigma^{-\frac{1}{2}}(x - \mu) \) and

\[ I_{i,t} = \begin{cases} 
1 & \text{if } \epsilon_{i,t} \geq -\frac{m_i \sigma_{i,t}}{s_i}, \\
1 & \text{if } \epsilon_{i,t} < -\frac{m_i \sigma_{i,t}}{s_i},
\end{cases} \]

\[ m_i = \frac{\Gamma\left(\frac{\nu - 1}{2}\right) \sqrt{\nu - 2}}{\Gamma\left(\frac{\nu}{2}\right) \sqrt{\pi}} \left(\lambda_i - \frac{1}{\lambda_i}\right), \quad s_i = \sqrt{\left(\lambda_i^2 + \frac{1}{\lambda_i^2} - 1 - m_i^2\right)}. \]

This allows us to formulate the copula for a Skewed Student-t distribution where \( \text{Skt}_{\nu,\lambda} \) and \( \text{Skt}_{\nu,\lambda}^{-1} \) are the CDF of the multivariate and inverse CDF of the univariate Skewed Student-t distribution with zero mean vector respectively:

\[ C(u_{1,t}, \ldots, u_{N,t}) = \text{Skt}_{\nu,\lambda}(\text{Skt}_{\nu,\lambda}^{-1}(u_{1,t}), \ldots, \text{Skt}_{\nu,\lambda}^{-1}(u_{N,t}))[0, R_t], \]
\[ c(u_{1,t}, \ldots, u_{N,t}) = \frac{|R_t|^{-\frac{3}{2}} \Gamma\left(\nu + \frac{N}{2}\right) \left[\Gamma\left(\frac{\nu}{2}\right)\right]^{N-1} \left(1 + \frac{a_i'a_i}{\nu - 2}\right)^{-\frac{\nu + N}{2}}}{\left[\Gamma\left(\nu + \frac{1}{2}\right)\right] \prod_{i=1}^{N} \left(1 + \frac{\lambda_i^{-2} s_i^2 (m_i C + s_i^2 x_{i,t})^2}{\nu - 2}\right)^{-\frac{\nu + 1}{2}}} \]
where

$$x_t = (x_{1N} \ldots x_{tN})$$, has elements \(x_{i,t} = Skt^{-1}_\nu(u_{i,t})\),

$$a_t = (a_{1N} \ldots a_{tN})$$, has elements \(a_{i,t} = \lambda_i^{-IC^*}{t}(m_i^C + s_i^C x^*_i)\)

given \(x^*_t, = R_t^{-1/2}x_t\), where \(R_t = [p_{ij,t}]\)

where

\[IC^*_{i,t} = \begin{cases} 1 & \text{if } x^*_{i,t} \geq -m_i^C \sqrt{s_i^C}, \\ -1 & \text{if } x^*_{i,t} < -m_i^C \sqrt{s_i^C} \end{cases} \]

\[I_C^*_{i,t} = \begin{cases} 1 & \text{if } x_{i,t} \geq -m_i^C \sqrt{s_i^C}, \\ -1 & \text{if } x_{i,t} < -m_i^C \sqrt{s_i^C} \end{cases} \]

given

\[m_i^C = \frac{\Gamma(\nu - 1)\sqrt{\nu - 2}}{\Gamma(\nu/2)\sqrt{\pi}} \left( \lambda_i - \frac{1}{\lambda_i} \right), \quad s_i^C = \sqrt{\left( \lambda_i^2 + \frac{1}{\lambda_i^2} - 1 - (m_i^C)^2 \right)}.\]

The correlation structure \(R_t = [p_{ij,t}]\) in all copulas above follows a Dynamic Conditional Correlation (DCC) structure similar to that of Engle (2002) or a Realised Dynamic Conditional Correlation (rDCC) based on preliminary work by Bauwens et al. (2012), with further details given in sections (2.3.4) and (2.3.5) respectively. A Skewed Student-t distribution is chosen for both the marginal and copula models as it provides a very flexible framework, which can capture the asymmetry of returns in the variance equation (Alberg, Shalit, and Yosef (2008) and Tu, Wong, and Chang (2008)) and a non linear dependence in the copula structure.

### 2.3.4 The Dynamic Conditional Correlation

The Dynamic Conditional Correlation (DCC) of Engle (2002) is used to model the evolution of the correlation structure \(R_t\) in the copulas by allowing it to evolve over time. A DCC(1,1) structure is used, taking the following form:

\[Q_t = (1 - \alpha - \beta) \cdot \hat{Q} + \alpha(x_{t-1}x_{t-1}') + \beta Q_{t-1}\]

\[R_t = \hat{Q}_t^{-1}Q_t\hat{Q}_t^{-1},\]

where \(\hat{Q}\) is the sample covariance of \(x\), \(\hat{Q}_t\) is a diagonal matrix with the diagonal elements taking the square root of \(Q_t\):

\[
Q_t = \begin{bmatrix}
\sqrt{q_{11}} & 0 & 0 & \ldots & 0 \\
0 & \sqrt{q_{22}} & 0 & \ldots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & \ldots & \sqrt{q_{NN}}
\end{bmatrix}
\]
therefore the elements of $R_t$ are in the form $p_{ij,t} = \frac{q_{ij,t}}{\sqrt{q_{ii,t}q_{jj,t}}}$, and the parameter constraints are identical to a univariate GARCH(1,1) model:

$$\alpha + \beta < 1, \quad \alpha, \beta \in (0, 1).$$

### 2.3.5 The Realised Covariance Dynamic Conditional Correlation

In order to take advantage of intra period dynamics, the Realised Covariance from (2) is used within a DCC framework, based on preliminary work by Bauwens et al. (2012), in order to model the evolution of the correlation structure $R_t$ in the copulas by allowing it to evolve over time. The Realised DCC(1,1) (rDCC) taking the following form:

$$Q_t = (1 - \alpha - \beta) \cdot \bar{Q} + \alpha P_{t-1} + \beta Q_{t-1}$$
$$R_t = \hat{Q}_t^{-1}Q_t\hat{Q}_t^{-1},$$

where

$$P_t = \{\text{diag}(C_t)\}^{-1/2}C_t\{\text{diag}(C_t)\}^{-1/2}$$

and

$$C_t = \begin{bmatrix} RV_{1,t} & RCov_{1,2,t} \\ RCov_{1,2,t} & RV_{2,t} \end{bmatrix}$$

where $\bar{Q}$ is the sample covariance of $x$, $\hat{Q}_t$ is a diagonal matrix with the diagonal elements taking the square root of $Q_t$ as per section 2.3.4.

### 3 Markov-Chain Monte Carlo Estimation

A likelihood function is specified according to the marginal and copula models, with Student-t and Skewed Student-t distribution for the former and Gaussian, Student-t and Skewed Student-t for the latter. A likelihood function for all distributions is provided for reference in Appendix A.

#### 3.1 Prior

A combination of mostly uninformative and Jeffrey’s priors are chosen over the possible region for the parameters in $\theta$, where:

$$\theta = [\alpha_0, \alpha_1, \beta_1, \nu, \lambda, \xi, \phi, \tau_1, \tau_2, \sigma^2_u, \nu_{mes}, \mu, \sigma^2_0, \alpha_c, \beta_c, \nu_c, \lambda_c]^t$$
The prior, \( \pi(\boldsymbol{\theta}) \propto \frac{1}{\sigma_u^2 \nu \nu_{mes} \nu_c} I_A \), (where A is the region described by the restrictions in Table 1) is used for the degree of freedom parameters in (3) and (4) and a standard Jeffery’s prior for the volatility of \( u_t \) in (4) as we are assuming they are a priori independent. A flat prior on the degree of freedom parameters would lead to an improper posterior distribution as shown in Bauwens and Lubrano (1998).

During the MCMC process, the following parameter restrictions are placed to achieve stationarity.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>( \mu, \alpha_0, \xi, \tau_1, \tau_2 )</th>
<th>( \alpha_1, \beta_1, \alpha_c, \beta_c )</th>
<th>( \nu, \nu_{mes}, \nu_c )</th>
<th>( \lambda, \lambda_c )</th>
<th>( \phi, \sigma_u^2, \sigma_0^2 )</th>
<th>( \phi \alpha_1 + \beta_1, \alpha_c + \beta_c )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bounds</td>
<td>( \pm \infty )</td>
<td>( \geq 0, &lt; 1 )</td>
<td>( &gt; 2, &lt; 200 )</td>
<td>( &gt; 0 )</td>
<td>( &gt; 0 )</td>
<td>( &lt; 1 )</td>
</tr>
</tbody>
</table>

The degree of freedom parameters \( \nu, \nu_{mes} \) and \( \nu_c \) are restricted to be above 2 and under 200 as without an upper limit, and with a normal distribution of the data, the likelihood is not integrable in terms of degrees of freedom, so a limit of 200 was put in place, as the Student-t distribution is practically normal at this level. The use of an inverse prior also places a stronger emphasis on lower values. The mean return and initial volatility are also optimised during MCMC via the \( \mu \) and \( \sigma_0^2 \) parameters respectively, while the \( \sigma_u^2 \) is only restricted positive in non RV GARCH models. Copula parameters are denoted with the subscript \( c \) in Table (1).

### 3.2 Sampling

In order to make inference about the parameters \( \boldsymbol{\theta} \), the random walk Metropolis (RWM) algorithm is used to draw samples from the posterior distribution. The adaptive proposal (AP) algorithm of Haario, Saksman, and Tamminen (1999) is used during burn-in iterations to produce an efficient proposal distribution for the RWM algorithm. The high number of iterations, combined with many unknown parameters, number of models, all for 1000 day forecast results in a very large computational burden. Using several computers simultaneously, the entire empirical forecasting exercise was completed in 5 days. Each single day’s forecast takes approximately 180 seconds.

### 3.3 Burn-in

Assuming that the parameter vectors \( \boldsymbol{\theta}^{(1)}, \ldots, \boldsymbol{\theta}^{(i-1)} \), for \( i > H \), have been generated, to generate \( \boldsymbol{\theta}^{(i)} \), the AP algorithm first simulates \( \boldsymbol{\theta}^p \) from the multivariate Normal proposal distribution:

\[
\boldsymbol{\theta}^p \sim \mathcal{N}\left( \boldsymbol{\theta}^{(i-1)}, c_d \Sigma_{t-1} + c_d \varepsilon \mathbf{I}_d \right),
\]

where \( \Sigma_{t-1} \) is the sample covariance matrix of \( \boldsymbol{\theta}_{t-H}, \ldots, \boldsymbol{\theta}_{t-1} \), \( H \) is the “memory parameter”, which is the number of previously sampled parameter vectors used for
evaluating the covariance matrix, $c_d$ is a scale factor depending on the dimension, $d$, of the parameter vector, and $\varepsilon$ is a small positive constant, which is used to prevent zero variances. Following Haario et al. (1999) and Gelman, Roberts, and Gilks (1996), $c_d = 2.4^2/d$. The algorithm then accepts $\theta^p$ as $\theta^{(i)}$, i.e., $\theta^{(i)} = \theta^p$ with probability:

$$\alpha = \min \left\{ \frac{p(\theta^p | r)}{p(\theta^{(i-1)} | r)}, 1 \right\},$$

and rejects $\theta^p$, i.e., $\theta^{(i)} = \theta^{(i-1)}$ with probability $1 - \alpha$.

The AP algorithm is set to run for 20,000 iterations per parameter, in order to guarantee convergence of the parameters. So a five parameter model would have a chain length of 100,000. Convergence usually occurs quickly, by 25,000 iterations or so.

### 3.4 Post burn-in

As the stationary distribution of the AP algorithm is not strictly ergodic, and it samples from a slightly different distribution from the target posterior distribution, it is used only during the burn-in period. After burn-in, the standard random walk Metropolis algorithm is used, whose proposal distribution is a multivariate Normal with a fixed covariance matrix determined using the last 10,000 burn-in samples.

Assuming that the parameter vector $\theta^{(i-1)}$, for $i > 1$, has been generated, to generate $\theta^{(i)}$, the RWM algorithm first simulates $\theta^p$ from the multivariate Normal proposal distribution:

$$\theta^p \sim \mathcal{N}\left(\theta^{(i-1)}, c_d \Sigma\right).$$

The algorithm then accepts $\theta^p$ as $\theta^{(i)}$ with probability $\alpha$ given by equation (23), and rejects $\theta^p$ with probability $1 - \alpha$. For post burn-in iterations, $\theta^{(1)}$ and $\Sigma$ are chosen to be the sample mean and the sample covariance matrix, respectively, of the parameter vectors sampled in the last 1000 burn-in iterations. The RWM algorithm is set to run for 10,000 iterations per parameter, and a mean of the last 10,000 chains is used to generate the parameter values, standard errors and 95% credible intervals. Tables of parameter results are shown Tables 11 and 12.

### 4 Forecasting Methodology

#### 4.1 VaR Forecasts

Each of the univariate and multivariate models in the paper is used to produce a 1000 one-step ahead volatility forecasts, which in turn are used to form a Value-at-Risk (VaR) forecast at both the 95% and 99% confidence level $\alpha$ as per (25), (26) and section (4.1.1). The accuracy of each model can then be measured by their relative risk forecast.
performance. The conditional 1 period VaR forecast is formally defined as:

$$\alpha = Pr(a_{t+1} < VaR_\alpha | \Omega_t)$$ (24)

where $a_{t+1}$ is the one period return from time $t$ to time $t+1$, $\alpha$ is the quantile level and $\Omega_t$ is the information set at time $t$. For the univariate case using a Student-t distribution, VaR is calculated via the inverse CDF ($t^{-1}_\nu$) of the distribution as follows:

$$VaR_\alpha = \mu + \sigma t^{-1}_\nu(\alpha) \sqrt{\frac{\nu - 2}{\nu}},$$ (25)

similarly, for a Skewed Student-t using its inverse CDF ($skt^{-1}_\nu,\lambda$) as follows

$$VaR_\alpha = \mu + \sigma skt^{-1}_\nu,\lambda(\alpha),$$ (26)

where $skt^{-1}_\nu,\lambda$ is defined in Appendix B.2.

Our aim is to estimate the one-step ahead VaR for a portfolio consisting of various financial assets, by employing each of the respective time series and formulating the most appropriate joint distribution function. By choosing specific marginals and copulas to link the assets together we can form a joint distribution function which can be used to form a VaR figure. For a simple 2 assets portfolio composed of assets $x$ and $y$, and hence given their respective log-returns and allocation weight $\beta$, we can obtain the following conditional joint distribution function estimated at time $t-1$:

$$H_t(x,y|\Omega_{t-1}) = C_t(F_t(x|\Omega_{t-1}), G_t(y|\Omega_{t-1})|\Omega_{t-1})$$ (27)

and therefore the cumulative distribution function for the portfolio return $z_t = \beta x_t + (1 - \beta)y_t$:

$$\zeta(z) = Pr(Z \leq z_t) = Pr(\beta X + (1 - \beta)Y \leq z_t)$$

$$= \int_{-\infty}^{+\infty} \left[ \int_{-\infty}^{\frac{1}{\beta}z_t - \frac{1 - \beta}{\beta}y_t} c_t(F_t(x|\Omega_{t-1}), G_t(y|\Omega_{t-1})|\Omega_{t-1}) \cdot f_t(x|\Omega_{t-1}) dx \right] \cdot g_t(y|\Omega_{t-1}) dy$$

The solution $z^*$ for the equation $\zeta(z^*) = p$ where $p$ is the confidence level, will produce the VaR at time $t-1$. As there as no closed form solutions for the calculation of VaR for copulas using various marginals, we will need to employ the use of Monte Carlo simulations, as used by Bauwens and Laurent (2005) and Jorion (2007), where Giot and Laurent (2003) show that 100,000 simulations provide an adequate level of precision for quantile estimates.
4.1.1 Monte Carlo simulation VaR for Copulas

By using the methods detailed in Cherubini, Luciano, and Vecchiato (2004), in a multivariate setting, we simulate 100,000 scenarios for each of asset N in the portfolio, and together with the conditional distribution function (27), form a portfolio distribution of returns at time t, where $VaR(\alpha)$ is determined by taking the empirical quantile at $\alpha$.

1. Simulate $j=100,000$ Monte Carlo scenarios for each asset of the portfolio using the conditional joint distribution (27).

   (a) Simulate a $j$ random variate $(u_{j,1}, \ldots u_{j,N})'$ from the copula $C_t(\cdot)$.

   i. For an N-dimensional conditional Normal Copula (16) we use the following algorithm:

   • Calculate the Cholesky decomposition $A$ of the $N \times N$ correlation matrix $\Sigma$ i.e. $[1 \ \hat{p}_t, \hat{p}_t \ 1]$, where $\hat{p}_t$ is the forecast conditional correlation estimated using (2.3.4) or (2.3.5).
   • Simulate $N$ independent random variates from the standard normal distribution $z_j = (z_{j,1} \ldots z_{j,N})'$.
   • Set $b_j = Az_j$.
   • Calculate $(u_{j,1} \ldots u_{j,N})' = (\phi(b_{j,1}), \ldots \phi(b_{j,N}))'$ where $\phi$ is the standard normal CDF. This vector is a $j$ random variate from the N-dimensional Normal copula $C_N(\cdot; \hat{p}_t|\Omega_t^{-1})$

   ii. For an N-dimensional conditional Student-t Copula (18) we use the following algorithm:

   • Calculate the Cholesky decomposition $A$ of the $N \times N$ correlation matrix $\Sigma$ i.e. $[1 \ \hat{p}_t, \hat{p}_t \ 1]$, where $\hat{p}_t$ is the forecast conditional correlation estimated using (2.3.4) or (2.3.5).
   • Simulate $N$ independent random variates from the standard normal distribution $z_j = (z_{j,1} \ldots z_{j,N})'$.
   • Set $b_j = Az_j$.
   • Simulate a random variate $s_j$ from a $\chi^2_{\hat{\nu}_{c,t}}$ distribution where $\hat{\nu}_{c,t}$ is the forecast degree of freedom parameter of the copula.
   • Set $b_j = \frac{\sqrt{\hat{\nu}_{c,t}}}{\sqrt{s_j}}b_j$.
   • Calculate $(u_{j,1} \ldots u_{j,N})' = (t_{\hat{\nu}_{c,t}}(b_{j,1}), \ldots t_{\hat{\nu}_{c,t}}(b_{j,N}))'$ where $t_{\hat{\nu}_{c,t}}$ is the Student-t CDF where $\hat{\nu}_{c,t}$ is the forecast degree of freedom parameter of the copula. This vector is a $j$ random variate from the N-dimensional Student-t copula $C_t(\cdot; \hat{p}_t, \hat{\nu}_{c,t}|\Omega_t^{-1})$

iii. For an N-dimensional conditional Skewed Student-t Copula (21) we use the following algorithm:
• Calculate the Cholesky decomposition \( A \) of the \( NxN \) correlation matrix \( \Sigma \) i.e., \( \begin{bmatrix} 1 & \hat{p}_t, \hat{p}_t \end{bmatrix} \), where \( \hat{p}_t \) is the forecast conditional correlation estimated using (2.3.4) or (2.3.5).

• Simulate \( N \) independent random variates using the independent univariate standard Skewed Student-t distribution \( z_j = (z_{j,1}, \ldots, z_{j,N})' \) where \( z_{j,t} = Skt_{\nu_{c,t},\lambda_{c,t}}^{-1}(u_j) \) and \( u_j \) is drawn from a Unif[0,1] distribution.

• Set \( b_j = Az_j \).

• Calculate \( (u_{j,1}, \ldots, u_{j,N})' = (Skt_{\nu_{c,t},\lambda_{c,t}}^{-1}(b_{j,1}), \ldots, Skt_{\nu_{c,t},\lambda_{c,t}}^{-1}(b_{j,N}))' \) where \( Skt_{\nu_{c,t},\lambda_{c,t}} \) is the Skewed Student-t CDF where \( \nu_{c,t} \) and \( \lambda_{c,t} \) are the forecast degree of freedom and skewness parameter of the copula respectively. This vector is a \( j \) random variate from the \( N \)-dimensional Skewed Student-t copula \( C_{skt}(\cdot; \hat{p}_t, \nu_{c,t}, \lambda_{c,t} | \Omega_{t-1}) \).

(b) Calculate the standardized log-returns of each asset by using the inverse CDF of the distribution employed in the marginal model (See Appendix for details):

• Normal: \( Q_j = (q_{j,1}, \ldots, q_{j,N})' = \phi^{-1}(u_{j,1}) \ldots \phi^{-1}(u_{j,N}) \)

• Student-t: \( Q_j = (q_{j,1}, \ldots, q_{j,N})' = t_{\nu_{c,t}}^{-1}(u_{j,1}) \ldots t_{\nu_{c,t}}^{-1}(u_{j,N}) \)

• Skewed Student-t: \( Q_j = (q_{j,1}, \ldots, q_{j,N})' = Skt_{\nu_{c,t},\lambda_{c,t}}^{-1}(u_{j,1}) \ldots Skt_{\nu_{c,t},\lambda_{c,t}}^{-1}(u_{j,N}) \)

(c) Use the marginal models from section (2.2) to rescale the standardized log-returns using the forecast mean \( (\hat{\mu}_t) \) and variances \( (\sqrt{\hat{h}_t}) \):

\[
[z_{1,j,t}, \ldots, z_{N,j,t}] = (\hat{\mu}_{1,t} + q_{j,1} \cdot \sqrt{\hat{h}_{1,t}}, \ldots, \hat{\mu}_{N,t} + q_{j,N} \cdot \sqrt{\hat{h}_{N,t}}) \quad (28)
\]

(d) Repeat (a) to (c) for \( j = 100,000 \).

2. Calculate the profit and loss distribution for the portfolio \( Z_t \) by using the portfolio weights \( (w_i) \) of each asset 1 to \( N \).

\[ Z_t^j = \sum_{i=1}^{N} (w_i) z_{i,j,t} \]

3. To calculate \( VaR_{\alpha,t} \) simply take the \( \alpha \) quantile of \( Z_t \) where \( Z_t^j \in Z_t \).

4.2 CVaR Forecasts

Conditional Value-at-Risk (CVaR) or Expected Shortfall (ES) is also used to evaluate models, as it has become preferred to VaR (McNeil et al., 2005) due to the latter’s shortcomings. If we let \( VaR_{\alpha,t} \) denote the VaR for an asset at confidence level \( \alpha \) at time
Then, we can define CVaR as:

$$CVaR_\alpha = \mu + \sigma E[\epsilon | \epsilon < VaR_{\alpha,t}]$$

(29)

For the univariate case using a Student-t distribution McNeil et al. (2005) define it as:

$$CVaR_\alpha = \mu + \sigma \frac{g_\nu(t_\nu^{-1}(\alpha))}{1 - \alpha} \left( \frac{\nu + (t_\nu^{-1}(\alpha))^2}{\nu - 1} \right) \sqrt{\frac{\nu - 2}{\nu}}$$

(30)

where $\nu$ is the estimated degrees of freedom restricted to $\nu > 2$ as per (3), $g_\nu$ and $t_\nu^{-1}$ are the Student-t probability density function and inverse cumulative distribution function respectively.

We can derive the CVaR for a Skewed Student-t distribution for a long position given the standardised VaR value $\Theta$:

$$CVaR_\alpha = \mu + \sigma E[\epsilon | \epsilon < \Theta]$$

(31)

where

$$E[\epsilon | \epsilon < \Theta] = \int_{-\infty}^{\Theta} bc \left( 1 + \frac{1}{\nu - 2} \left( \frac{b \epsilon + a}{1 - \lambda} \right)^2 \right)^{-\frac{\nu + 1}{2}} x dx = \frac{c(1 - \lambda)^2 \nu - 2}{b \cdot skt_{\nu,\lambda}(\Theta) 1 - \nu} \left( 1 + \frac{1}{\nu - 2} \left( \frac{b \Theta + a}{1 - \lambda} \right)^2 \right)^{\frac{1 - \nu}{2}} - \frac{a}{b}$$

where

$$a = 4\lambda c \left( \frac{\nu - 2}{\nu - 1} \right), \quad b = \sqrt{1 + 3\lambda^2 - a^2}, \quad c = \frac{\Gamma \left( \frac{\nu + 1}{2} \right)}{\sqrt{\pi(\nu - 2)} \Gamma \left( \frac{\nu}{2} \right)}$$

where $skt_{\nu,\lambda}$ is the CDF of the Skewed Student-t distribution as per Appendix B.1.

To calculate CVaR for a multivariate case we simply extend the simulations employed in (4.1.1) as follows:

1. Find the mean of the quantile given that $Z^j_t < VaR_{\alpha,t}$ at level $\alpha$.

4.3 Back-testing VaR forecast models

An informal way to assess the performance of VaR forecasts, is to compute their Violation Rate (VRate):

$$VRate = \sum_{t=s+1}^{s+f} I(a_t < VaR_t),$$

(32)
where \( s \) is the estimation period, \( f \) is the forecast period, \( I \) is the indicator taking a value of 1 when \( a_t < VaR_t \), with the scope of comparing individual violation ratios (VRatio) which is simply (32) divided by the respective level of \( \alpha \), with a target of 1. If competing models are equidistant from such target, the model that produces the lower VRatio is preferred as it indicates a more conservative model. Formal methods of evaluation are the Unconditional Coverage (UC) test introduced by Kupiec (1995), the Conditional Coverage (CC) of Christoffersen (1998) and the Dynamic Quantile (DQ) test using 3 lags of Engle and Manganelli (2004).

### 4.4 Back-testing CVaR forecast models

Expected Shortfall Rate (ESRate) is the equivalent violation rate for the CVaR models, and is defined as:

\[
ESRate = \sum_{t=s+1}^{s+f} I(a_t < ES_t),
\]

once again used to obtain an ESRatio, where 1 is preferred, by dividing (33) by the expected remaining quantile at the respective level \( \delta_\alpha \), where each \( \delta_\alpha \) is obtained by taking the average degree of freedom and skewness, calculating the theoretical standardised CVaR, and running this through the CDF of that distribution. \( \delta_\alpha \) levels used are shown in Table (2). \( \delta \) levels for the copula are an average of the \( \delta \) levels used in the portfolios. Similarly to VaR back-testing, the UC, CC and DQ test can also be applied to examine the violations from a CVaR forecast.

#### Table 2: \( \delta_\alpha \) level used to calculate ESRate for each distribution.

<table>
<thead>
<tr>
<th>Portfolio</th>
<th>( \delta_{\alpha=5%} )</th>
<th>Copula</th>
<th>( \delta_{\alpha=1%} )</th>
<th>Copula</th>
</tr>
</thead>
<tbody>
<tr>
<td>GARCH</td>
<td>0.01832 0.01840</td>
<td>Student-t Skew-t Normal</td>
<td>0.01840 0.01840</td>
<td>Student-t Skew-t Normal</td>
</tr>
<tr>
<td>RV GARCH</td>
<td>0.01832 0.01840</td>
<td>Student-t Skew-t Normal</td>
<td>0.01840 0.01840</td>
<td>Student-t Skew-t Normal</td>
</tr>
<tr>
<td>rDCC</td>
<td>0.01840 0.01840</td>
<td>Student-t Skew-t Normal</td>
<td>0.01840 0.01840</td>
<td>Student-t Skew-t Normal</td>
</tr>
</tbody>
</table>

### 5 Empirical Study

#### 5.1 Data

The models are tested by utilising both weekly and RV series from daily data from five international asset classes: Equity Index: Standard & Poor’s 500 (SP500) - USA; Equity - IBM - USA; Commodity: Gold; Currency: GBPUSD; Fixed Income: US Treasury 10 Year Bond (US10YRT). The data was obtained from Bloomberg, covering the period
Figure 2: Plots of Returns for all data sets. Red dotted line separates estimation and forecast period.
Table 3: Summary of descriptive statistics for each of the five asset return series. JB is the Jarque-Bera statistic to test the null hypothesis of normality.

<table>
<thead>
<tr>
<th>Period</th>
<th>Mean</th>
<th>Std</th>
<th>Skewness</th>
<th>Kurtosis</th>
<th>Min</th>
<th>Max</th>
<th>JB</th>
</tr>
</thead>
<tbody>
<tr>
<td>SP500</td>
<td>0.20</td>
<td>2.16</td>
<td>-0.57</td>
<td>7.68</td>
<td>-16.66</td>
<td>8.01</td>
<td>916.5</td>
</tr>
<tr>
<td></td>
<td>0.14</td>
<td>2.60</td>
<td>-0.87</td>
<td>12.28</td>
<td>-20.26</td>
<td>17.49</td>
<td>3715.1</td>
</tr>
<tr>
<td>IBM</td>
<td>0.02</td>
<td>3.23</td>
<td>-0.19</td>
<td>5.20</td>
<td>-16.83</td>
<td>11.87</td>
<td>196.9</td>
</tr>
<tr>
<td></td>
<td>0.26</td>
<td>4.12</td>
<td>0.01</td>
<td>7.37</td>
<td>-21.17</td>
<td>23.00</td>
<td>796.8</td>
</tr>
<tr>
<td>GOLD</td>
<td>0.08</td>
<td>3.32</td>
<td>0.19</td>
<td>21.69</td>
<td>-30.91</td>
<td>29.20</td>
<td>13817.1</td>
</tr>
<tr>
<td></td>
<td>0.12</td>
<td>2.40</td>
<td>-0.37</td>
<td>8.27</td>
<td>-14.79</td>
<td>13.27</td>
<td>1179.7</td>
</tr>
<tr>
<td>GBPUSD</td>
<td>-0.04</td>
<td>1.59</td>
<td>-0.28</td>
<td>5.83</td>
<td>-10.09</td>
<td>6.99</td>
<td>329.1</td>
</tr>
<tr>
<td></td>
<td>0.01</td>
<td>1.22</td>
<td>-0.31</td>
<td>4.93</td>
<td>-6.41</td>
<td>4.85</td>
<td>171.1</td>
</tr>
<tr>
<td>US10YRT</td>
<td>0.00</td>
<td>2.15</td>
<td>0.48</td>
<td>7.37</td>
<td>-8.70</td>
<td>13.60</td>
<td>790.7</td>
</tr>
<tr>
<td></td>
<td>0.05</td>
<td>1.34</td>
<td>0.04</td>
<td>5.15</td>
<td>-5.44</td>
<td>8.39</td>
<td>192.4</td>
</tr>
</tbody>
</table>

March 1971 to May 2014. Five day percentage log-returns are defined by

$$r_t = (ln(P_t) - ln(P_{t-5})) \times 100$$

where $P_t$ is the closing price on day $t$. RV is calculated as per Section 2 using daily data. This generates 9747 daily data points after adjusting for synchronisation issues and small sample size variations due to trading day holidays etc, which converts into 1949 weekly return observations. This sample is then divided into an initial learning period of 949 weeks from March 1971 to March 1993 which includes the Black Monday of 1987, and a forecasting period of 1000 weeks from March 1993 to May 2014 which includes the Dot.com crash and the Global Financial Crisis (GFC). Figure 2 shows a plot of returns for the all data sets. Summary statistics for each data series are shown in Table 3, which clearly show that the estimation period has higher maximums and bigger losses, all periods exhibit leptokurtosis and negative skewness. Formal tests (Jarque & Bera, 1987) for normality show that none of the series are normally distributed.

Table 4 shows the sample correlation coefficients for the underlying assets of the portfolio using three types of measures, with Pearson’s $r$ showing the strength in linear relationship, and both Spearman’s $\rho$ and Kendall’s $\tau$ being non-parametric and showing the strength of the non-linear relationship. The strongest relationship is between IBM and the SP500, showing 0.628 for Pearson, 0.631 for Spearman and 0.455 for Kendall. In the majority of sample pairs, the non-linear dependence is stronger than the linear dependence, indicating that a multivariate model would be ideal to capture such dependence. Figure 3 shows the pairwise correlation for the historical daily returns, while Figures 4 and 5 show the transformed data on the unit-hypercube for the GARCH and RV-GARCH marginals respectively, prior to fitting the copulas.
Table 4: Sample correlation coefficients for the learning period.

<table>
<thead>
<tr>
<th>Return</th>
<th>Correlation Type</th>
<th>SP500</th>
<th>IBM</th>
<th>GOLD</th>
<th>GBPUSD</th>
<th>US10YRT</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pearson (r)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SP500</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Spearman (ρ)</td>
<td>0.6284</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Kendall (τ)</td>
<td>0.4549</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IBM</td>
<td></td>
<td>0.0428</td>
<td>0.6683</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Pearson (r)</td>
<td>0.6305</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Spearman (ρ)</td>
<td>0.0178</td>
<td>0.0464</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Kendall (τ)</td>
<td>0.4549</td>
<td>0.0311</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GOLD</td>
<td></td>
<td>0.0703</td>
<td>0.6900</td>
<td>0.2351</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Pearson (r)</td>
<td>0.0906</td>
<td>0.0555</td>
<td>0.2842</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Spearman (ρ)</td>
<td>0.0178</td>
<td>0.0464</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Kendall (τ)</td>
<td>0.0120</td>
<td>0.0372</td>
<td>0.1940</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>GBPUSD</td>
<td></td>
<td>0.3337</td>
<td>0.2184</td>
<td>-0.0020</td>
<td>0.1528</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Pearson (r)</td>
<td>0.3888</td>
<td>0.2519</td>
<td>-0.0376</td>
<td>0.1146</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Spearman (ρ)</td>
<td>0.0028</td>
<td>0.0564</td>
<td>0.2842</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Kendall (τ)</td>
<td>0.2705</td>
<td>0.1731</td>
<td>-0.0266</td>
<td>0.0766</td>
<td></td>
</tr>
</tbody>
</table>

Figure 3: Plots of pairwise correlation for historical weekly returns
Figure 4: GARCH Skew-t transformed returns prior to fitting a copula

Figure 5: RV GARCH Skew-t transformed returns prior to fitting a copula
5.2 Estimation Results

In this section we present the results of the parameter estimates obtained from the MCMC chains during the learning period. Tables 11 and 12 summarise the estimation results for the GARCH and copula parameters respectively for each of the marginal models together with the GARCH on the aggregate portfolios for day 1 of the forecast period.

The GARCH parameter results for the aggregate portfolio are similar for both the GARCH (PGt and PGSkt) and the Realised Volatility GARCH (PRVtt and PRVSktt) using either the Student-t or Skewed Student-t distributions. There is a slight reduction in the \( \alpha_0 \) intercept parameter between GARCH and RV-GARCH, but all are statistically significant. The shift in \( \alpha_0 \) is far more pronounced when comparing the results for the individual assets between GARCH and RV-GARCH.

In terms of ARCH and GARCH parameters \( \alpha_1 \) and \( \beta_1 \), there is a mild shift up and down respectively when comparing GARCH and RV-GARCH, while the sum of the two parameters always remains below unity, even if this restriction is lifted while using RV-GARCH as per Table 1. This shift is again more pronounced when comparing the individual asset return series. The joint sum being close to unity is the usual characteristic of high volatility persistence in asset returns, which is present in all models and data sets.

The degree of freedom parameter \( \nu \) is similar between GARCH and RV-GARCH, and is fairly low in all cases (below 10), as is common with financial data as it exhibits leptokurtic tails.

Skewness parameter \( \lambda \) is close to 1 (implying no skewness) or marginally below (implying negative skewness), and is statistically significant for all data sets. Absence of large negative skewness is probably due to using weekly data rather than daily, as it is expected that the left tail be longer than the right when analysing financial returns. The parameter \( \xi \) should be close to zero if no bias is present, and in all of the data sets the value is greater than -1. Estimates of \( \phi \) are all very close to unity, which suggest that the realised measure RV is roughly proportional to the conditional variance.

The parameters in the leverage function \( \tau_1 \) and \( \tau_2 \) are all of the expected sign and magnitude, except RGSkt SP500 and RGSkt IBM, with \( \tau_1 \) expected to be negative as is consistent with the well known (Nelson (1991)) phenomenon in financial returns of a negative correlation between adjoining periods’ volatility, with \( \tau_2 \) expected to be positive. The size of the asymmetry between a negative and positive shock is dictated by \( \tau_1 \) while the level at which this occurs is governed by \( \tau_2 \).

The parameters \( \sigma_u^2 \) and \( \nu_{mes} \) have no practical interpretation, but are required as an input for the likelihood function of the measurement equation in (4).

Financial return mean \( \mu \) is small and not statistically significant in essentially all data sets, which is expected.

\( \sigma_0 \) is the starting value for the volatility in GARCH, and shows that a constant value
as is typically used might not be ideal.

The estimates for the copula parameters in Table 12 show that the values for $\alpha_c$ and $\beta_c$ are similar for the copula using GARCH marginals and RV marginals, while the copula using the realised DCC exhibits lower $\beta_c$ values, with the sum being further from unity.

Copulas using the Student-t distribution exhibit a similar value for the degree of freedom parameter $\nu_c$, while the copulas using GARCH marginals and the Skewed Student-t distribution exhibit a much lower degree of freedom value.

Skewness parameters $\lambda_c,1$ to $\lambda_c,5$ are all slightly above one, indicating a positive skewness, but this has no interpretation as it is related to the dependency. All values are statistically significant.

A plot of the one-step ahead correlation forecast using the RV Skewed Student-t marginal and Student-t copula (other combinations gave similar results) is provided in Figure 6 and shows the evolving nature of the correlation dynamics, and emphasises the need to accurately model such dependence structure using a copula model. Using a static correlation between return pairs is clearly over simplistic, and might result in an incorrect portfolio allocation or risk forecast.

Figure 6: RV GARCH Skew-t Marginal and Student-t Copula one-step ahead correlation forecast.

5.3 VaR forecast comparison

Table 5 shows a summary of the violation ratios for each model for both the aggregate portfolio and when copula models are employed. At the $\alpha = 5\%$ level, RV GARCH using univariate portfolio data provides the best estimate, getting very close to 1. For models
using copulas, there is a clear improvement from GARCH to RV-GARCH marginals, with a further improvement using the realised DCC copulas and RV GARCH as marginals. The worst estimate is the Student-t copula using the GARCH marginal. A similar situation arises at the $\alpha = 1\%$ level, but in this case the RV GARCH marginals provide superior estimates to the rDCC model. The worst performing model is the Skewed Student-t copula with GARCH marginals. Even though the best forecasting model is one employing a univariate return series, modelling the portfolio constituent dependencies might still be a worthwhile endeavour as the outcome could be used for optimal portfolio allocation rather than a Value-at-Risk application.

In all cases, both uni and multivariate, it is clear that using RV GARCH does provide superior estimates, while it is ambiguous if a Realised Covariance structure within the rDCC is providing an improvement using this test.

Table 6 shows the results of the UC, CC and DQ tests as described above, displaying if the model is rejected by the null at either the 5% and 1% significance level, where 1 indicates a rejection of the null. There is only one instance of a null rejection, the DQ test for the portfolio RV GARCH model with Student-t distribution at the $\alpha = 1\%$ level. A good VaR forecasting performance is translated into a favourable test result, with essentially all models doing well.

Table 5: Value at Risk Violation Ratios for each model.

<table>
<thead>
<tr>
<th>Portfolio Copula</th>
<th>$\alpha = 5%$</th>
<th>$\alpha = 1%$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Copula</td>
<td>GARCH</td>
<td>RV GARCH</td>
</tr>
<tr>
<td></td>
<td>Student-t Skew-t</td>
<td>Student-t Skew-t</td>
</tr>
<tr>
<td>Student-t</td>
<td>1.06 0.90</td>
<td>1.14 1.18</td>
</tr>
<tr>
<td>Skew-t</td>
<td>1.08 1.10</td>
<td>1.10 1.20</td>
</tr>
<tr>
<td>Normal</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 6: Value at Risk Unconditional Coverage, Conditional Coverage and Dynamic Quantile Tests Results. 1 indicates rejection by test at relevant $\alpha$ level.

<table>
<thead>
<tr>
<th>Portfolio Copula</th>
<th>$\alpha = 5%$</th>
<th>$\alpha = 1%$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Copula</td>
<td>GARCH</td>
<td>RV GARCH</td>
</tr>
<tr>
<td></td>
<td>Student-t Skew-t</td>
<td>Student-t Skew-t</td>
</tr>
<tr>
<td>Student-t</td>
<td>0000</td>
<td>0000</td>
</tr>
<tr>
<td>Skew-t</td>
<td>0000</td>
<td>0000</td>
</tr>
<tr>
<td>Normal</td>
<td>0000</td>
<td>0000</td>
</tr>
</tbody>
</table>

5.4 CVaR forecast comparison

Figure 7 shows the 1000 one-step ahead CVaR forecast at the 1% level for a selection of models under consideration; Portfolio GARCH with Student-t (black dashed dotted), Portfolio RV GARCH with Skewed Student-t (magenta dotted), GARCH and Skewed
Student-t copula (blue dashed) and RV GARCH and Student-t copula (green solid). It is interesting to note how the univariate models follow a different forecast compared to multivariate models using copulas. The GARCH and Skewed Student-t copula models consistently underestimates risk with an ESRatio of 2.50 over the whole period, while at the other end of the performance spectrum, the portfolio RV GARCH with Skewed Student-t model has an ESRatio of 1.11, an identical result to the portfolio RV GARCH with Student-t. As per the VaR results above, the univariate models using RV GARCH perform better than the models using any copula and any marginal. There is again a clear performance improvement when using RV GARCH models for both the univariate and multivariate setting, while the rDCC copula with RV GARCH marginals has identical performance to the DCC copulas with RV GARCH marginals. There is again minimal model rejection, with the portfolio RV GARCH with Skewed Student-t being rejected by all tests, and the DQ test rejecting both the GARCH and rDCC models with Skewed Student-t copula.

In summary, the use of Realised Volatility in either a univariate or multivariate setting does provide a consistent improvement compared to lower frequency data sets at both quantiles. This is an attractive quality if being used for regulatory purposes, as this would forecast an optimal amount of capital allocation for an institution. Table 7 shows a similar summary described in Section 5.3 but for the ESRatios. As above, the values which are closest to 1 are boxed while the values furthest from 1 are displayed in bold.

Table 7: Conditional Value at Risk Violation Ratios for each model.

| Portfolio | | $\delta = 5\%$ | | $\delta = 1\%$ | |
|-----------|-----------------|----------------|-----------------|-----------------|
|           | Student-t Skew-t | Normal Student-t Skew-t | Student-t Skew-t | Normal Student-t Skew-t |
| GARCH     | 1.47 1.31 1.47 1.47 1.52 | | 1.39 1.12 1.67 1.67 2.50 |
| RV GARCH  | 1.30 1.14 1.30 1.30 1.30 | | 1.11 1.11 1.39 1.39 2.22 |
| rDCC      | 1.30 1.25 1.30 | | 1.39 1.39 2.22 | |

Table 8: Conditional Value at Risk Unconditional Coverage, Conditional Coverage and Dynamic Quantile Tests Results. 1 indicates rejection by test at relevant $\alpha$ level.

| Portfolio | | $\delta_{\alpha=0\%}$ | | $\delta_{\alpha=1\%}$ | |
|-----------|-----------------|-----------------|-----------------|-----------------|
|           | Student-t Skew-t | Normal Student-t Skew-t | Student-t Skew-t | Normal Student-t Skew-t |
| GARCH     | 0 0 0 0 0 | | 0 0 0 0 0 |
| RV GARCH  | 0 0 0 0 0 | | 0 1 1 0 0 |
| rDCC      | 0 0 0 0 0 | | 0 0 0 0 0 |

5.5 Loss function

A loss function is employed in order to further assess both VaR and CVaR tail forecasts generated. As per Koenker and Bassett Jr (1978) this is done through a criterion function,
Figure 7: 1% CVaR Forecast for Portfolio GARCH with Student-t (black dashed dotted), Portfolio RV GARCH with Skewed Student-t (magenta dotted), GARCH and Skewed Student-t Copula (blue dashed) and RV GARCH and Student-t Copula (green solid) over the whole 1000 day forecast period. Portfolio Returns (Solid Red).
which should be minimized in quantile regression estimation, defined as:

\[ LF = \sum_{t=s+1}^{s+f} (a_t - Q_t)(\alpha - I_t). \]

Where I is an indicator taking the value of 1 if there is a violation for either VaR or CVaR \((a_t < Q_t)\), \(Q_t\) is the quantile forecast, and \(\alpha\) is the evaluation quantile for VaR, taking \(\delta_\alpha\) values described in Table 2 for CVaR. The model with the lowest loss function figure is preferred. Results for Value-at-Risk and Conditional Value-at-Risk loss function are provided in Tables 9 and 10 respectively, where lowest values are boxed and highest values are displayed in bold.

For VaR, at the 5% quantile level, the lowest loss is exhibited by the RV GARCH marginal with Normal copula and the highest by the portfolio GARCH with Skewed Student-t distribution. As can be seen in Figure 8, there is always a clear improvement when going from GARCH to RV GARCH at the 5% quantile level.

At the 1% quantile level, the best performing is again the RV GARCH marginal with Normal copula, with the worst performing being the GARCH marginal with Skewed Student-t copula. As per Figure 8, there is a worsening when using RV GARCH compared to GARCH for the portfolio, a trend that is reversed when any copula is employed.

A similar situation is present for CVaR, with again the RV GARCH marginal with Normal copula being the best performer at the 5% quantile level, and the worst being the GARCH marginal with Skewed Student-t copula. Figure 9, shows a mild improvement when using RV GARCH for the portfolio compared to GARCH, but again a drastic improvement when copulas are used.

At the 1% quantile level, the best performing model is the RV GARCH marginal with rDCC Student-t copula with the worst being the portfolio RV GARCH. A similar finding as per VaR can be observed in Figure 9, at the 1% quantile level for the portfolio models, while the improvement is milder when comparing GARCH and RV GARCH models. Figures 8, and 9, reinforce the findings of RV GARCH models being superior to GARCH models, but also do show that copula models can provide more accurate forecasts if the Loss Function metric is used as a basis.

### Table 9: Value at Risk Loss Function.

<table>
<thead>
<tr>
<th></th>
<th>( \alpha = 5% )</th>
<th></th>
<th>( \alpha = 1% )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Portfolio</td>
<td>Copula</td>
<td>Portfolio</td>
</tr>
<tr>
<td></td>
<td>GARCH</td>
<td>Normal</td>
<td>Skew-t</td>
</tr>
<tr>
<td>RV GARCH</td>
<td>150.71</td>
<td>150.63</td>
<td>146.17</td>
</tr>
<tr>
<td>GARCH</td>
<td>152.86</td>
<td>153.11</td>
<td>148.20</td>
</tr>
<tr>
<td>rDCC</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

30
Table 10: Conditional Value at Risk Loss Function.

<table>
<thead>
<tr>
<th>Portfolio</th>
<th>$\alpha=5%$</th>
<th>$\alpha=1%$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Student-t</td>
<td>Skew-t</td>
</tr>
<tr>
<td>GARCH</td>
<td>70.72</td>
<td>70.25</td>
</tr>
<tr>
<td>RV GARCH</td>
<td>70.28</td>
<td>69.88</td>
</tr>
<tr>
<td>rDCC</td>
<td>67.50</td>
<td>67.61</td>
</tr>
</tbody>
</table>

Figure 8: Loss Function plot for each model and Copula Distribution for the Value at Risk forecast at the 5% and 1% level. Legend: Blue - GARCH; Green - RV GARCH; Red - rDCC RV GARCH.

Figure 9: Loss Function plot for each model and Copula Distribution for the Conditional Value at Risk forecast at the $\delta_{\alpha}=5\%$ and $\delta_{\alpha}=1\%$ level. Legend: Blue - GARCH; Green - RV GARCH; Red - rDCC RV GARCH.
6 Conclusion

This paper utilises the Realised Volatility GARCH model as a marginal in a multivariate setting by employing various copula models, which are compared to univariate GARCH models and copula models with GARCH marginals. Our findings show that using Realised Volatility produces a more accurate forecast for both VaR and CVaR at the 5% and 1% quantiles when using a variety of formal and informal tests. The choice of distribution does not appear to be a major factor in producing accurate forecasts, at either the univariate or copula level. The use of high-frequency data in the covariance structure does seem to offer an identical result or marginal improvement, but is not as evident as high frequency data in the volatility equation. Several extension could compare the performance of Archimedean copulas when GARCH and RV GARCH marginals are used, as well as Vine-copulas. The use of alternative realised measures such as the Realised Range could also be considered, while also looking at domestic market assets and reverting to intra-day data for the calculation of Realised Volatility.

Appendix A LogLikelihood Functions

c: Student-t distribution

\[ \mathcal{L}(\theta | r) = T \left[ \ln \Gamma \left( \frac{\nu + 1}{2} \right) - \ln \Gamma \left( \frac{\nu}{2} \right) - 0.5 \ln(\nu - 2\pi) \right] \\
- \sum_{t=1}^{T} \left[ \frac{\nu + 1}{2} \ln \left( 1 + \frac{a_t^2}{(\nu - 2)\sigma_t^2} \right) + 0.5 \ln(\sigma_t^2) \right] \]

\[ \text{I} = \begin{cases} 
1 & \text{if } a_t \geq -\frac{m}{s} \\
-1 & \text{if } a_t < -\frac{m}{s} 
\end{cases} \]

\[ u_t : \text{Normal distribution} \]

\[ \mathcal{L}(\theta | RV) = -0.5 \sum_{t=1}^{T} \left( \ln \sigma_{a,t}^2 + \frac{u_{a,t}^2}{\sigma_{a,t}^2} \right) \]
$u_t$ : Student-t distribution

$$
\mathcal{L}(\theta | RV) = T \left[ \ln \Gamma \left( \frac{\nu_{mes} + 1}{2} \right) - \ln \Gamma \left( \frac{\nu_{mes}}{2} \right) - 0.5 [\ln (\nu_{mes} - 2) \pi] \right]
- \sum_{t=1}^{T} \left[ \frac{\nu_{mes} + 1}{2} \ln \left( 1 + \frac{u_{t}^2}{(\nu_{mes} - 2)\sigma_{u,t}^2} \right) + 0.5 \ln (\sigma_{u,t}^2) \right]
$$

Normal Copula

$$
\mathcal{L}(R|u_t) = \frac{1}{2} \sum_{t=1}^{T} (\ln |R| + x_t (R^{-1} - I)x_t)
$$

where

$$
x_t = (x_{1N}...x_{tN}), \quad x_{it} = \Phi^{-1}(u_{i,t})
$$

Student-t Copula

$$
\mathcal{L}(R, \nu | u_t) = -T \left[ \ln \Gamma \left( \frac{\nu + N}{2} \right) + (N - 1) \ln \Gamma \left( \frac{\nu}{2} \right) - N \cdot \ln \Gamma \left( \frac{\nu + 1}{2} \right) \right]
- \frac{\nu + N}{2} \sum_{t=1}^{T} \ln \left( 1 + \frac{x_t R_i^{-1} x_t}{\nu} \right) - \frac{1}{2} \sum_{t=1}^{T} \ln |R_t| + \frac{\nu + 1}{2} \sum_{t=1}^{T} \sum_{i=1}^{N} \ln \left( 1 + \frac{x_{i,t}^2}{\nu} \right)
$$

where

$$
x_t = (x_{1N}...x_{tN}), \quad x_{it} = \Phi^{-1}(u_{i,t})
$$

Skewed Student-t Copula

$$
\mathcal{L}(R, \nu, \lambda | u_t) = -T \left[ \ln \Gamma \left( \frac{\nu + N}{2} \right) + (N - 1) \ln \Gamma \left( \frac{\nu}{2} \right) - N \cdot \ln \Gamma \left( \frac{\nu + 1}{2} \right) \right]
- \frac{\nu + N}{2} \sum_{t=1}^{T} \ln \left( 1 + \frac{a_t^2}{\nu - 2} \right) - \frac{1}{2} \sum_{t=1}^{T} \ln |R_t|
+ \frac{\nu + 1}{2} \sum_{t=1}^{T} \sum_{i=1}^{N} \ln \left( 1 + \frac{\lambda_{i,t}^{-2} (m_i^C + s_i^C x_{i,t})^2}{\nu - 2} \right)
$$

where

$$
\text{vector} \quad x_t = (x_{1N}...x_{tN}), \quad \text{has elements} \quad x_{it} = \text{skt}^{-1}_{\nu, \lambda}(u_{i,t}),
$$

$$
\text{vector} \quad a_t = (a_{1N}...a_{tN}), \quad \text{has elements} \quad a_{it} = \lambda_{i,t}^{-1} (m_i^C + s_i^C x_{i,t})
$$

given $x_t^* = R_t^{-1} x_t$. 

33
where

\[
I_{i,t}^C = \begin{cases} 
1 & \text{if } x_{i,t}^s \geq -\frac{m_i^C}{s_i^C}, \\
-1 & \text{if } x_{i,t}^s < -\frac{m_i^C}{s_i^C}, 
\end{cases}
\]

\[
I_{i,t}^C = \begin{cases} 
1 & \text{if } x_{i,t} \geq -\frac{m_i^C}{s_i^C}, \\
-1 & \text{if } x_{i,t} < -\frac{m_i^C}{s_i^C}, 
\end{cases}
\]

given

\[
m_i^C = \frac{\Gamma\left(\frac{\nu-1}{2}\right)\sqrt{\nu-2}}{\Gamma\left(\frac{\nu}{2}\right)} \left(\lambda_i - \frac{1}{\lambda_i}\right),
\]

\[
s_i^C = \sqrt{\left(\frac{\lambda_i^2}{\lambda_i^2} + 1 - (m_i^C)^2 \right)}.
\]

### Appendix B Properties of the Skewed Student-t Distribution

#### B.1 The Cumulative Distribution Function

The CDF of the Skewed Student-t Distribution is denoted by $Skt_{\nu,\lambda}(z)$. As per the PDF at (7), this is defined over 2 regions either side of $-\frac{m}{s}$.

If $z < -\frac{m}{s}$:

\[
Skt_{\nu,\lambda}(z) = \frac{2\lambda s \Gamma\left(\frac{\nu+1}{2}\right)}{(1 + \lambda^2) \Gamma\left(\frac{\nu}{2}\right) \sqrt{\nu-2} \pi} \int_{-\infty}^{z} \left[1 + \frac{\lambda^2 (m + st)^2}{(\nu - 2)}\right]^{-\frac{\nu+1}{2}} dt.
\]

By using the following substitutions:

- $w = \frac{\nu}{\nu-2}[\lambda(m + st)]$, therefore $dw = \frac{\nu}{\nu-2}[\lambda s]dt$.

- When $t = -\infty$, $w = -\infty$ and when $t = z$, $w = \sqrt{\frac{\nu}{\nu-2}[\lambda(m + st)]}$

Therefore:

\[
Skt_{\nu,\lambda}(z) = \frac{2\lambda s \Gamma\left(\frac{\nu+1}{2}\right)}{(1 + \lambda^2) \Gamma\left(\frac{\nu}{2}\right) \sqrt{\nu-2} \pi} \sqrt{\frac{\nu-2}{\nu\lambda^2 s^2}} \int_{-\infty}^{z} \left[1 + \frac{\lambda^2 (m + st)^2}{(\nu - 2)}\right]^{-\frac{\nu+1}{2}} \sqrt{\frac{\nu\lambda^2 s^2}{\nu-2}} dt
\]

\[
= \frac{2\Gamma\left(\frac{\nu+1}{2}\right)}{(1 + \lambda^2) \Gamma\left(\frac{\nu}{2}\right) \sqrt{\nu\pi}} \int_{-\infty}^{\sqrt{\frac{\nu}{\nu-2}[\lambda(m + s)]}} \left(1 + \frac{w^2}{\nu}\right)^{-\frac{\nu+1}{2}} dw 
\]

\[
= \frac{2}{1 + \lambda^2} T_{\nu} \left(\sqrt{\frac{\nu}{\nu-2}[\lambda(m + s)]}\right).
\]

(34)
If \( z \geq -\frac{m}{s} \), then the CDF is composed of the following two integrals \( I_1 \) and \( I_2 \):

\[
Skt_{\nu, \lambda}(z) = \frac{2\lambda s \Gamma\left(\frac{\nu+1}{2}\right)}{(1 + \lambda^2) \Gamma\left(\frac{\nu}{2}\right) \sqrt{\nu - 2}} \int_{-\infty}^{-\frac{m}{s}} \left[ 1 + \frac{\lambda^2 (m + st)^2}{(\nu - 2)} \right]^{-\frac{\nu}{2}} dt + \frac{2\lambda s \Gamma\left(\frac{\nu+1}{2}\right)}{(1 + \lambda^2) \Gamma\left(\frac{\nu}{2}\right) \sqrt{\nu - 2}} \frac{1}{\lambda} \int_{-\frac{m}{s}}^{z} \left[ 1 + (m + st)^2 \right]^{-\frac{\nu}{2}} dt = I_1 + I_2.
\]

To obtain \( I_1 \) we substitute \( z = -\frac{m}{s} \) into (34):

\[
I_1 = \frac{2}{1 + \lambda^2} T_\nu(0) = \frac{1}{1 + \lambda^2}.
\]

Similarly, to calculate \( I_2 \) we can do the following substitution:

- \( w = \sqrt{\frac{\nu}{\nu - 2}} \left[ \frac{1}{\lambda} (m + st) \right] \), therefore \( dw = \sqrt{\frac{\nu}{\nu - 2}} \left[ \frac{s}{\lambda} \right] dt \).
- When \( t = -\frac{m}{s} \), \( w = 0 \) and when \( t = z \), \( w = \sqrt{\frac{\nu}{\nu - 2}} \left[ \frac{1}{\lambda} (m + sz) \right] \)

giving:

\[
I_2 = \frac{2\lambda s \Gamma\left(\frac{\nu+1}{2}\right)}{(1 + \lambda^2) \Gamma\left(\frac{\nu}{2}\right) \sqrt{\nu - 2}} \int_{0}^{z} \frac{1}{\lambda} \int_{-\frac{m}{s}}^{w} \left[ 1 + (m + sz)^2 \lambda^2 (m + sz) \right]^{-\frac{\nu}{2}} \sqrt{\frac{\nu s^2}{\lambda^2 (\nu - 2)}} dt \frac{\nu}{\lambda^2 (\nu - 2)} dw
\]

\[
= \frac{2\lambda^2}{1 + \lambda^2} \left[ T_\nu\left(\sqrt{\frac{\nu}{\nu - 2}} \left[ \frac{1}{\lambda} (m + sz) \right] \right) - T_\nu(0) \right] - \lambda^2 \frac{1}{1 + \lambda^2}
\]

So given \( z = -\frac{m}{s} \):

\[
Skt_{\nu, \lambda}(z) = I_1 + I_2 = \frac{1 - \lambda^2}{1 + \lambda^2} + \frac{2\lambda^2}{1 + \lambda^2} \left[ T_\nu\left(\sqrt{\frac{\nu}{\nu - 2}} \left[ \frac{1}{\lambda} (m + sz) \right] \right) - \lambda^2 \frac{1}{1 + \lambda^2} \right]
\]

(35)

By combining (34) and (35) we obtain the CDF of the Skewed Student-t Distribution:

\[
Skt_{\nu, \lambda}(z) = \begin{cases} 
\frac{2\lambda^2}{1 + \lambda^2} \left[ T_\nu\left(\sqrt{\frac{\nu}{\nu - 2}} \left[ \frac{1}{\lambda} (m + sz) \right] \right) - \lambda^2 \frac{1}{1 + \lambda^2} \right] & \text{if } z < -\frac{m}{s}, \\
\frac{2\lambda^2}{1 + \lambda^2} \left[ T_\nu\left(\sqrt{\frac{\nu}{\nu - 2}} \left[ \frac{1}{\lambda} (m + sz) \right] \right) - \lambda^2 \frac{1}{1 + \lambda^2} \right] + \frac{1 - \lambda^2}{1 + \lambda^2} & \text{if } z \geq -\frac{m}{s}.
\end{cases}
\]

(36)
B.2 The Inverse Cumulative Distribution Function

The Inverse CDF of the Skewed Student-t Distribution is denoted by $Skt_{\nu,\lambda}^{-1}(u)$ where $u \in [0,1]$. To determine the change point $-\frac{m}{s}$ on the inverse scale we use the CDF (36) so that:

$$Skt_{\nu,\lambda}^{-1}\left(-\frac{m}{s}\right) = \frac{1}{1 + \lambda^2}.$$ 

Therefore the region $z < -\frac{m}{s}$ is equivalent to $u < \frac{1}{1 + \lambda^2}$ and $z \geq -\frac{m}{s}$ is equivalent to $u \geq \frac{1}{1 + \lambda^2}$ on the inverse scale. To obtain the inverse CDF we interchange the variables into the CDF and solve the resulting equation. Let $u = Skt_{\nu,\lambda}(z)$ and $q = z$ in (36). Over the region $u < \frac{1}{1 + \lambda^2}$ we obtain:

$$u = \frac{2}{1 + \lambda^2} T_{\nu} \left(\sqrt{\frac{\nu}{\nu - 2}}[\lambda(m + sq)]\right)$$

solving for $q$ yields:

$$q = \frac{1}{s} \left[\frac{1}{\lambda} \sqrt{\frac{\nu - 2}{\nu}} T_{\nu}^{-1}\left(\frac{1 + \lambda^2}{2} - m\right) - m\right]. \quad (37)$$

Similarly, over the region $u \geq \frac{1}{1 + \lambda^2}$ :

$$u = \frac{1 - \lambda^2}{1 + \lambda^2} + \frac{2\lambda^2}{1 + \lambda^2} T_{\nu} \left(\sqrt{\frac{\nu}{\nu - 2}}\left[\frac{1 - \lambda^2}{2\lambda^2 u} - \frac{1 - \lambda^2}{2\lambda^2}\right]\right)$$

solving for $q$ yields:

$$q = \frac{1}{s} \left[\lambda \sqrt{\frac{\nu - 2}{\nu}} T_{\nu}^{-1}\left(\frac{1 + \lambda^2}{2\lambda^2 u} - \frac{1 - \lambda^2}{2\lambda^2}\right) - m\right]. \quad (38)$$

By combining (37) and (38), the Inverse CDF for the Skewed Student-t distribution is defined as:

$$Skt_{\nu,\lambda}^{-1}(u) = \begin{cases} \frac{1}{s} \left[\frac{1}{\lambda} \sqrt{\frac{\nu - 2}{\nu}} T_{\nu}^{-1}\left(\frac{1 + \lambda^2}{2} - m\right) - m\right] & \text{if } u < \frac{1}{1 + \lambda^2}, \\ \frac{1}{s} \left[\lambda \sqrt{\frac{\nu - 2}{\nu}} T_{\nu}^{-1}\left(\frac{1 + \lambda^2}{2\lambda^2 u} - \frac{1 - \lambda^2}{2\lambda^2}\right) - m\right] & \text{if } u \geq \frac{1}{1 + \lambda^2}. \end{cases} \quad (39)$$
References


39, 841–862.


Fisher-Report. (1994). Basel euro-currency standing committee of the central banks of the group of ten countries. public disclosure of market and credit risks by financial intermediaries..


### Table 11: GARCH MCMC Parameter Estimate with 95% Credible Interval.

<table>
<thead>
<tr>
<th>TABLE 11: GARCH MCMC PARAMETER ESTIMATE WITH 95% CREDIBLE INTERVAL.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Parameter</strong></td>
</tr>
<tr>
<td>----------------</td>
</tr>
<tr>
<td>α0</td>
</tr>
<tr>
<td>β0</td>
</tr>
<tr>
<td>λ</td>
</tr>
<tr>
<td>ξ</td>
</tr>
<tr>
<td>φ</td>
</tr>
<tr>
<td>σ2</td>
</tr>
<tr>
<td>PGt</td>
</tr>
<tr>
<td>PRVtt</td>
</tr>
<tr>
<td>PRVskt</td>
</tr>
<tr>
<td>GSktp</td>
</tr>
<tr>
<td>GSktp IBM</td>
</tr>
<tr>
<td>GSktp GOLD</td>
</tr>
<tr>
<td>GSktp GBPUSD</td>
</tr>
<tr>
<td>GSktp US10YRT</td>
</tr>
<tr>
<td>RGSktp SP500</td>
</tr>
<tr>
<td>RGSktp IBM</td>
</tr>
<tr>
<td>RGSktp GOLD</td>
</tr>
<tr>
<td>RGSktp GBPUSD</td>
</tr>
<tr>
<td>RGSktp US10YRT</td>
</tr>
</tbody>
</table>

### Table 12: Copula MCMC Parameter Estimate with 95% Credible Interval.

<table>
<thead>
<tr>
<th>TABLE 12: COPULA MCMC PARAMETER ESTIMATE WITH 95% CREDIBLE INTERVAL.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Parameter</strong></td>
</tr>
<tr>
<td>----------------</td>
</tr>
<tr>
<td>N GSkt</td>
</tr>
<tr>
<td>t GSkt</td>
</tr>
<tr>
<td>Skt GSkt</td>
</tr>
<tr>
<td>N RGSkt</td>
</tr>
<tr>
<td>t RGt</td>
</tr>
<tr>
<td>Skt RGSk</td>
</tr>
<tr>
<td>rDCC N RGSk</td>
</tr>
<tr>
<td>rDCC t RGSk</td>
</tr>
<tr>
<td>rDCC Skt RGSk</td>
</tr>
</tbody>
</table>

42