ABSTRACT

HUYNH, TAM CONG. Goodness of Fit Tests for GARCH-Copula Models. (Under the direction of Peter Bloomfield.)

For years the Gaussian copula function was the golden egg in financial technology that allowed complex risks to be modeled with more ease and accuracy than ever before. The method was adopted by various financial institutions, from investors, banks, rating agencies and regulators. However, before long it became clear that the gaussian copula was not respected by the data especially in the extreme values it was being used to model. The Gaussian copula became instrumental in the world financial crisis in 2008. In this paper I address the question of the selection of garch-copula models in terms of forecasting accuracy of Value-at-Risk. I consider six different copulas, each paired with five unique garch models. I selected fourteen assets selected by volume from The New York Stock Exchange. I propose an alternative goodness of fit test that is more sensitive to deviation from data in the left tail. This test is adapted from the chi square test. Finally, the combinations of garch-copula models are benchmarked against dynamic condition correlation and HAR models.

Goodness of Fit Tests for GARCH-Copula Models

by
Tam Cong Huynh

A thesis submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the Degree of
Master of Science

Statistics

Raleigh, North Carolina

2014

APPROVED BY:

Soumendra Lahiri

Kimberly Weems

Peter Bloomfield
Chair of Advisory Committee

DEDICATION

To my parents.

BIOGRAPHY

The author was born in a small town . . .

ACKNOWLEDGEMENTS

I would like to thank my advisor, Professor Peter Bloomfield, for his invaluable guidance.

I would also like to thank Professor Jeffrey Scroggs, my former advisor who still insists on telling me what to do.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER

1

INTRODUCTION

Copula models have been successfully implemented in many applications but in no area is it as influential as in finance and risk management. As financial instruments become more sophisticated and complex, it becomes more difficult to find an informative and practical risk measure. The most widely used measure is Value at Risk, an estimate of the quantile of the profit-loss distribution of a portfolio. Calculations for portfolio VaR is often modeled by a variance-covariance approach that makes the assumption that asset returns follow a multivariate normal distribution. However, empirical studies show that financial asset returns tend to be negatively skewed and leptokurtic. When multiple sources of risk are undertaken in a portfolio an accurate measurement of dependence is essential to VaR estimation. Copula theory has proven to be very powerful in this respect because it allows expression of a multivariate distribution as a function of its marginals. This allows modeling of the dependence in a joint distribution to be separated from the modeling of the marginals of that joint distribution. Longin and Solnick (2001) and Ang and Chen (2002) found evidence that asset returns become more correlated during volatile markets and during deep bear markets. These long tails and deviations from normality can lead to very poor VaR estimates.

GARCH models, originally proposed by Engle, has led to a fundamental change to volatility modeling in the prices of financial assets. These models address many of the issues in modeling financial times series, such as, nonstationarity of price series, absence of autocorrelation for the price variations, autocorrelation of the squared price returns, volatility clustering, fat-tailed distributions, and leverage effects. Furthermore, GARCH models can be coupled with a copula, where the copula provides a more robust estimation of the garch innovations and can more accurately model the dependence structure between two residual series. There have been numerous successful applications of copulaes in conjunction with GARCH models. McNeil and Frey and Rockinger and Jondeau (2010) use GARCH models to estimate the current volatility of the log returns series and use the copula to estimate the distribution of the garch innovations. Bob (2013) show that GARCH-EVT-Copula approach outperforms the commonly used variance covariance methods. In this thesis, I adopt the same approach and try to answer the question of copula selection.

There have been studies that have shown when GARCH-copula models benchmarked against multivariate garch and correlation based models, GARCH-copula models tend to produce the best results and reliable VaR estimates. There have been few studies that address the issue of copula selection which significantly impacts the joint distribution and the resulting VaR estimates. Several studies have tried to answer the question of copula selection. Ane and Kharoubi (2003) investigate the dependence structure of international stock index returns and conclude that misspecification of the dependence structure can account for up to 20 percentof the error in the estimation of VaR. Palaro and Hotta (2006) found the SJC copula paired with egarch models allow for different dependence in the tails and produced better results than tradition methods. Kole et al. (2007) use copula goodness of fit tests to choose a parametric copula for the computation of value-at-risk and expected shortfall with inconclusive results. Weib (2011) suggest that goodness of fit tests have little power when it comes to selecting the optimal parametric copula for VaR and ES calculations. Genest (2009) show that that best procedures overall are those based on the cramer von mises test but reliability depends on the choice of parametric copula. Weib (2011) suggest that copula models are not necessarily better suited for risk estimation than a benchmark correlation-based model, however they only considered a garch(1,1) model in their simulation.

This paper extends previous research in three ways. Previous studies on goodness of fit either consider just the copula or a garch-copula pair where model is a garch(1,1). Here I consider five different garch specifications. Secondly, I implement a suite of tests that are more robust risk measures than solely VaR and ES. Thirdly, I propose an alternative goodness of fit test that is more sensitive to deviations in the tail than the current goodness of fit tests that exist. This goodness of fit test is adapted from the chi square test.

1.1 Design of the Empirical Study

An empirical study was conducted using financial assets from the United States, we try to answer whether goodness of fit testing can be used to choose the optimal garch-copula model for estimating VaR. To compare the forecasting accuracy of different models, the VaR of portfolios are estimated and backtested.

1.1.1 Assets

Fourteen assets were chosen from American assets sorted by highest volume in hopes of capturing a cross section of the US economy, eleven stocks were selected across major industries along with three different commodities. Every possible pair of the fourteen assets were chosen to create ninety-one bivariate portfolios. Five garch specifications were used, each paired with six different copulas fitted to the standardized garch residuals of the two assets in each portfolio.

1.1.2 Models

The log returns of each asset over 2000 trading days back from September 8, 2005 to October 22, 2013 are modelled with each of the following garch models: aparch(1,1), egarch(1,1), gjr(1,1), igarch(1,1), garch(1,1). Each asset is then paired with every other asset creating 91 bivariate portfolios. Six copulas including: normal, t, Clayton, Frank, Amh, and Joe, are used to capture the dependence structure of the standardized residuals from each garch model.

1.1.3 Model evaluation

For each of the portfolios, the fit of 30 models was evaluated with the use of 7 goodness of fit tests, out of which 5 are based on the Rosenblatt transformation and 2 are based on the empirical copula.

1.1.4 Forecasting Accuracy

The predicted model VaR is compared with the empirical VaR through backtesting. The Kupiec test or the unconditional coverage test, determines the adequacy of the estimated VaR, by comparing the actual number of exceedances assuming the model is correct, against a confidence interval on the expected number of exeedances. The Christoffersen tests are used to determine whether the "hit sequence" or the times of exceedances are independent.

1.2 Garch Models

GARCH Model

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i y_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2$$

Heteroscedasticity is a stylized fact of financial time series. ARCH model introduced by Engle (1982) and GARCH models introduced by Bollerslev (1986) model time varying condition variance. The conditional variance of a univariate econometric time series is expressed as a linear function of the squared past values of the series.

Asymmetric Power ARCH

$$\sigma_t^\delta = \omega + \sum_{i=1}^{q} \alpha_i (|\epsilon_{t-i}| - \gamma_i \epsilon_{t-i})^\delta + \sum_{i=1}^{p} \beta_i \sigma_{t-i}^\delta$$

where $\delta > 0$ and $-1 < \gamma_j < 1, j = 1, ...p$ Large negative turns seem to increase volatility more than large positive returns, often referred to as the "leverage effect." Aparch models replace the square function ϵ_i^2 with a more flexible function that can capture this asymmetric response. Also adds another dimension to modeling with $delta > 0$ paramater.

Glosten, Jagannath, and Runkle Garch

$$\sigma_t^2 = \omega + \alpha \epsilon_{t-1}^2 + \gamma \epsilon_{t-1}^2 \mathbb{1}(\epsilon_{t-1} < 0) + \beta \sigma_{t-1}^2$$

Similarly to the Aparch model, the GJR model captures leverage effects but using the square function as in the original garch model.

Integrated GARCH

$$(1 - \alpha(L) - \beta(L))\epsilon_t^2 = \omega + (1 - \beta(L))\nu_t$$

where $\nu_t = \epsilon_t^2 - \sigma_t^2$ and $\alpha(L)$ and $\beta(L)$ denote appropriately defined lag polynomials. This model imposes an exact unit root in the corresponding autoregressive polynomial.

Exponential GARCH

$$log(\sigma_t^2) = \omega + \alpha(|z_{t-1}| - E(|z_{t-1}|)) + \gamma z_{t-1} + \beta log(\sigma_{t-1}^2)$$

For $\gamma < 0$ negative will model will capture the "leverage effect." By modeling the log transform of the conditional variance, this model avoids having to ensure the process remains positive.

1.3 Copulas

A copula expresses a multivariate distribution as a function of its marginals. A copula models the dependence between variates in a multivariate distribution and can be combined with any set of univariate distributions for the marginals. As a result, copulas allow us to tap into the wide variety of univariate distributions available.

1.3.1 Sklar's Theorem

Statement of Sklar's theorem an application of the probability integral transform. Sklar's theorem where the vector $U_1, ..., U_n$ are the psuedo observations.

$$\begin{aligned} F(y_1, ..., y_m) &= F(F_1^{-1}(u_1), ..., F_m^{-1}(u_m)) \\ &= Pr[U_1 \le u_1, ..., U_m \le u_m] \\ &= C(u_1, ..., u_m) \end{aligned}$$

Normal Copula

$$C(u_1, u_2; \theta) = \int_{-\infty}^{\Phi^{-1}(u_1)} \int_{-\infty}^{\Phi^{-1}(u_2)} \frac{1}{2\pi(1-\theta^2)^{1/2}} \frac{-(s^2 - 2\theta st + t^2}{2(1-\theta^2)} ds dt$$

where Φ is the distribution of the standard normal distribution. The normal copula is flexible in that it allows for equal degrees of positive and negative dependence and includes both Frechet bounds.

t copula

$$C^t(u_1, u_2; \theta_1, \theta_2) = \int_{-\infty}^{t_{\theta_1}^{-1}(u_1)} \int_{-\infty}^{t_{\theta_1}^{-1}(u_2)} \frac{1}{2\pi(1-\theta_2^2)^{1/2}} \left(1 + \frac{-(s^2 - 2\theta st + t^2}{v(1-\theta_2^2)}\right)^{-(\theta_1+2)/2} ds dt$$

where $t_{\theta_1}^{-1}(u_1)$ is the inverse distribution function of a t-distribution with θ_1 degrees of freedom. θ_1 controls the heaviness of the tails. For $\theta_1 < 3$, the variance does not exist and for $\theta_1 < 5$, the fourth moment does not exist.

Clayton Copula

$$C(u_1, u_2; \theta) = \left(u_1^{-\theta} + u_2^{-\theta} - 1\right)^{-1/\theta}$$

$\theta \in (0, \infty)$ As theta approaches zero, the marginals become independent. As θ approaches infinity, the copula attains the Frechet upper bound. The Clayton copula cannot account for negative dependence. Strong left tail dependence and relatively weak right tail dependence.

Frank Copula

$$C(u_1, u_2, ; \theta) = -\theta^{-1} log\left(1 + \frac{(e^{-\theta u_1} - 1)(e^{-\theta u_2} - 1)}{e^{-\theta} - 1}\right)$$

$\theta \in (-\infty, \infty)$ Values of $(-\infty, 0, \infty)$correspond to Frechet lower bound, independence, and Frechet upper bound, respectively. The Frank copula allows for negative dependence, symmetric in both tails, includes both frechet bounds. However, dependence in the tails of the Frank copula are relatively weak, its strongest dependence is centered in the middle of the distribution.

Ali, Mikhail, and Haq copula

$$C(u_1, u_2, ; \theta) = \frac{u_1 u_2}{1 - \theta(1 - u_1)(1 - u_2)}$$

where $\theta \in [-1, 1]$. This is the only copula whose parameter lies on a closed interval between -1 and 1 and measures both, positive and negative, dependence. Similar to the Gumbel copula.

Joe copula

$$C(u_1, u_2, ; \theta) = \exp(-[(1 - p_1)u_1 + (1 - p_2)u_2 + ((p_1 u_1)^{\delta_{12}} + (p_2 u_2)^{\delta_{12}})^{\frac{1}{\delta_{12}}}]^{\frac{1}{\theta}})$$

Overall strength parameter $\theta \in [-1, \infty)$. Parameter for each pair of variables $\delta_{i,j} \in [1, \infty]$, and an additional parameter p_j, where $\frac{1}{p_j} \geq$ m -1, for each of the m variables, allowing for more control over the tails. Not every possible correlation matrix can be matched by these copulas. The θ parameter plays an important role in determining what is possible. This parameterization of the Joe copula does not allow for lower tail dependence.

1.4 Testing Procedures

Testing procedures are based on the difference between a parametric copula and an empirical copula. I consider two alternative approaches to construction of a goodness of fit test, one type based on Dehuevel's empirical copula and another based on the Rosenblatt transformation.

1.4.1 Tests based on empirical copula

Below is a variant of Dehuevel's (1979) empirical copula that converges asymptotically to the one original proposed by Dehuevel,

$$C_n(u) = \frac{1}{n}\sum_{i=1}^{n} 1(U_{i1} \leq u_1, ..., U_{id} \leq u_d)$$

$\mathbf{u} = (u_1, ..., u_d) \in [0,1]^d$

Goodness of fit tests are based on the difference between this empirical copula and the assumed parametric copula.

$$\mathbb{C}_n = \sqrt{n}(C_n - C_{\theta_n})$$

Ranked based versions of the cramer-von Mises

$$S_n = \int_{[0,1]^2} \mathbb{C}_n(\mathbf{u})^2 d\, C_n(\mathbf{u})$$

Large values of this statistic leads to rejection of $H_0 : C \in C_0$ The limiting distribution of process $\mathbb{C}_n$ depends on the parametric copula C_{θ_n}. P-values have to be approximated using a specially adapted monte carlo method Genest and Remillard (2009).

1.4.2 Tests based on Rosenblatt's transform

We consider five test statistics based on the Rosenblatt transformation. By successively conditioning each random variable on the ones we have already conditioned upon, we should get a set of uniform and independently distributed random variables on [0,1]. Then we can test $H_0 : Z = TX \sim U(0,1)$ Consider random vector $X = (X_1, ..., X_k)$ with distribution function $F(x_1, ..., x_k)$. Let $z = (z_1, ..., z_n) = Tx = T(x_1, ..., x_n)$, where T is given by the transformation:

$z_1 = P(X_1 \leq x_1) = F_1(x_1)$

$z_2 = P(X_2 \leq x_2 | X_1 = x_1) = F_2(x_2|x_1)$
$\vdots$
$z_k = P(X_k \leq x_k | X_{k-1} = x_{k-1}, ..., X_1 = x_1) = F_k(x_k|x_{k-1}, ..., x_1)$
$Z = TX = (Z_1, ..., Z_k)$ is uniform and independently distributed on [0,1]

Under H_0 the empirical distribution of the pseudo observations should be the independence copula $C_\perp(u)$, and let the empirical distribution of the psuedo-observations be defined as,

$$D_n(u) = \frac{1}{n}\sum_{i=1}^{n} 1(z_i \leq u), \quad u \in [0,1]^d$$

Cramer von mises tests can be constructed from the difference between the empirical distribution of psuedo-obesrvations and the independence copula. Two tests that differ in their integration measure are presented here,

$$\begin{aligned} S_n^{(C)} &= n\int_{[0,1]^d} [D_n(u) - C_\perp(u)]^2, \quad u \in [0,1]^d \\ &= \sum_{i=1}^{n} [D_n(z_i)) - C_\perp(z_i)]^2 \end{aligned}$$

and

$$\begin{aligned} S_n^{(B)} &= n\int_{[0,1]^d} [D_n(u) - C_\perp(u)]^2, \quad u \in [0,1]^d \\ &= \frac{n}{3^d} - \frac{1}{2^{d-1}}\sum_{i=1}^{n}\prod_{k=1}^{d}(1 - E_{ik}^2) + \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}\prod_{k=1}^{d}(1 - E_{ik} \vee E_{jk}) \end{aligned}$$

the asymptotic null behavior of

$$\sqrt{n}(D_n - C_\perp)$$

was found by Ghoudi and Remillard (2004) but goodness of fit testing is carried out using the same parametric bootstrap procedure used in previous section .

Anderson Darling:

$$\chi_i = \left(\sum_{j=1}^{d} (\Phi^{-1}(u_{ij})^2 \right)$$

Anderson-Darling statistic proposed by Breyman et al.(2003). Φ is the normal cumulative distribution function of a standard normal random variable and χ_d^2 denotes the distribution function of the chi-square distribution with d degrees of freedom. If the psuedo-observations $E_i, ...E_n$ are uniformly and indepently distributed on unit hypercube then $\chi_i' s$ should follow a chi-square distribution and a test is constructed from the difference of these $\chi_i' s$ and the theoretical χ^2 distribution.

AnGamma:

$$\gamma_i = \sum_{j=1}^{d} -\log(u_{ij})$$

Similar to the anderson darling tests but transformation to the gamma distribution is used instead of chi-square distribution.

1.5 Backtesting

A portfolio's value at risk is defined as the α quantile of the portfolio's profit and loss distribution.

$$\text{VaR}_t(\alpha) = -F^{-1}(\alpha) \tag{1.1}$$

VaR has grown increasingly popular, most financial institutions now use VaR as a measure of risk. It becomes increasingly important that statistical tests of VaR models be conducted rigorously.

1.5.1 Kupiec Test

The Kupiec test, or the proportion of failure (PoF) is a test of the unconditional coverage property. Let x be the number of realized VaR violations and n be the total number of observations during a time window. Assuming the VaR model is correct, then x $\sim$ binomial (n,p), where p is the failure rate and the probability of an observation exceeding a specified quantile.

$$f(x|n,p) = \binom{n}{x} p^x (1-p)^{n-x}$$

An unbiased estimator of p is the sample proportion of failure, $\frac{x}{n}$. The Kupiec test measures whether the observed failure rate $\hat{p}$ is significantly different from the true proportion of failure, $H_0 : p = \hat{p} = \frac{x}{n}$.

Kupiec suggests the likelihood ratio test (LRT) which takes the form:

$$\lambda(x) = -2\ln(\frac{(1-p)^{n-x}p^x}{(1-\frac{x}{n})^{n-x}(\frac{x}{n})^x})$$

Under H_0 the LRT is asymptotically χ^2 distributed with 1 degree of freedom. Unconditional coverage is both informative and compelling but it not reliable except for large samples. It has difficulty detecting VaR models that systematically under report risk. It also only considers the frequency of exceedances and not the time they occur, so it fails to detect clustering and dependence of the hit sequence.

1.5.2 Christoffersen Tests

Christoffersen tested the independence of exceedences, whether or not the probability of exceedence at any given time period depends on whether there was an exceendence the time period before. First let us define an indicator of a "hit" or exceedance:

$$f(x) = \begin{cases} 1 & \text{if there is an exceedance} \\ 0 & \text{if there is not an exceedance} \end{cases}$$

let $\pi_{ij} = \Pr(I_t(\alpha) = j | I_{t-1}(\alpha) = i)$. Then $I_t(\alpha)$ is a markov chain with transition matrix

$$\Pi_1 = \begin{pmatrix} 1-\pi_{01} & \pi_{01} \\ 1-\pi_{11} & \pi_{11} \end{pmatrix}$$

The likelihood function

$$L(\Pi_1, I_t(\alpha)) = (1-\pi_{01})^{N_{00}}\pi_{01}^{N_{01}}(1-\pi_{11})^{N_{10}}\pi_{11}^{N_{11}}$$

where N_{ij} is the number of observations where the process transitions from state i to state j. The transition matrix can be estimated by the ratio of counts of transitions from 0 to 1 over the total number of transitions out of state 0. Similarly other probabilities can be estimated by:

$$\hat{\Pi}_1 = \begin{pmatrix} \frac{N_{00}}{N_{00}+N_{01}} & \frac{N_{01}}{N_{00}+N_{01}} \\ \frac{N_{10}}{N_{10}+N_{11}} & \frac{N_{11}}{N_{10}+N_{11}} \end{pmatrix}$$

$H_0 : \pi_2 = \pi_{01} = \pi_{11}$, the probability of violation occurring after a day of violation is the same as the probability of an exceedance after a day of no violation. The likelihood under the this null hypothesis is

$$L(\Pi_1, I_t(\alpha)) = (1-\pi_{01})^{N_{00}+N_{10}}\pi_{01}^{N_{01}+N_{11}}$$

where

$$\Pi_2 = \begin{pmatrix} 1-\pi_2 & \pi_2 \\ 1-\pi_2 & \pi_2 \end{pmatrix}$$

and π_2 can be estimated as

$$\hat{\pi}_2 = \frac{N_{01}+N_{11}}{N_{00}+N_{01}+N_{10}+N_{11}}$$

and the likelihood ratio test statistic is

$$\lambda(\alpha) = L(\hat{\Pi}_1, I_t(\alpha)) - L(\hat{\Pi}_2, I_t(\alpha))$$

$\lambda(\alpha)$ converges asymptotically to a χ^2distribution with one degree of freedom. This test has limited power against clustering. Christoffersen and Pelletier argues that the time elapsed between violations should be independent of the time elapsed since the last violation.

CHAPTER

2

RESULTS

In this chapter we analyze the data, the fitted models, and examine the goodness of their fittings with diagnostics.

2.1 Data

The data are from 14 assets, 11 stocks, 3 commodities of 2000 daily adjusted closing prices taken from Yahoo finance. The log returns and absolute log returns are pictured below.

In figure 2.3 there is evidence of the stylized fact of volatility clustering. Tables 2.1 and 2.2 are summary statistics of the 14 stocks in our sample. The average return for all assets are about zero. Most assets show signs of skewness and all assets exhibit excess kurtosis. For all assets, the Jarque-Bera test rejects the null hypothesis that the skewness and kurtosis of log returns match the skewness and kurtos of a normal distribution. The autocorrelation plots are not presented but there is little evidence of auto-correlation. Very few lags are significant at large lag values and barely pierce the signficance line. The ACF and PACF of the squared returns show strong correlation.

Table 2.1 Descriptive Statistics for Daily Log Returns of First 7 Stocks

Statistics	BOA	CSCO	FCX	F	FOX	GE	GOLD
Mean	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Standard Deviation	0.04	0.02	0.03	0.03	0.02	0.02	0.03
Minimum	-0.3	-0.15	-0.22	-0.26	-0.19	-0.18	-0.22
Median	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Maximum	0.34	0.18	0.21	0.29	0.17	0.14	0.20
Skewness	0.27	0.25	0.48	0.12	-0.19	0.01	-0.19
Kurtosis	16.64	10.36	5.08	14.06	8.91	9.79	5.38
Jarque-Bera	1	1	1	1	1	1	1

Table 2.2 Descriptive Statistics for Daily Log Returns of Last 7 Stocks

Statistics	INTC	MU	MSFT	NOK	PFE	SLV	USA
Mean	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Standard Deviation	0.02	0.04	0.02	0.03	0.02	0.04	0.05
Minimum	-0.11	-0.21	-0.17	-0.27	-0.10	-0.28	-0.46
Median	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Maximum	0.13	0.20	0.13	0.17	0.11	0.24	0.36
Skewness	0.25	0.14	-0.04	0.08	0.16	0.00	-0.18
Kurtosis	4.87	4.22	10.18	7.49	6.93	4.90	7.45
Jarque-Bera	1	1	1	1	1	1	1

2.2 Modeling Univariate Distributions

Due to the heaviness of the tails exhibited by garch models, t distributions were fit to the standardized residuals or innovations of the garch models using maximum likelihood estimates for the degrees of freedom. Below are diagnostic plots of distribution fittings for garch(1,1) standardized residuals of Pfizer and US Airways.

Diagnostics generally look good. There do not appear to be huge deviations from a theoretical t distribution with 19 degrees of freedom. Some observations imply a distribution with even fatter tails than is modelled by the t-distribution. A skewed-t distribution would be more appropriate in this case.

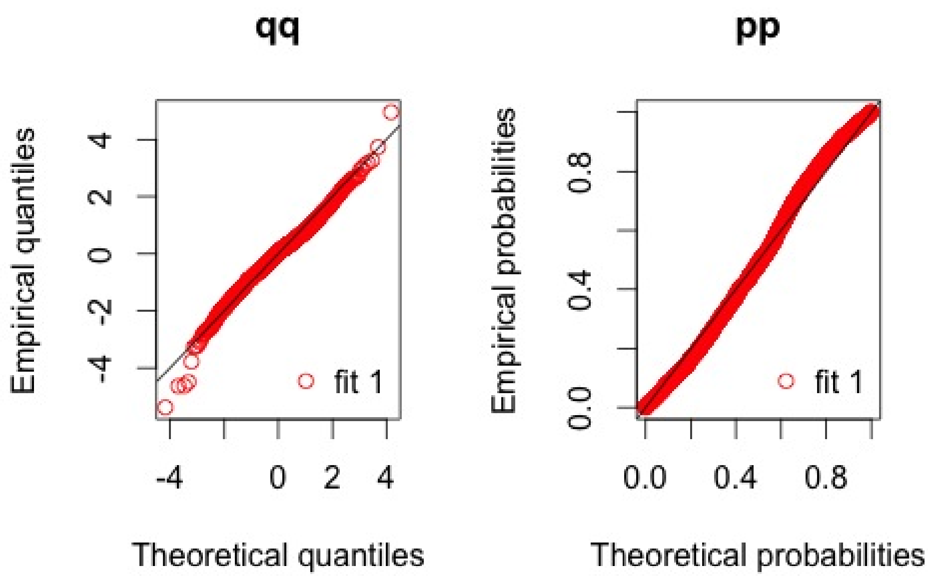

Figure 2.1 QQ and PP plots for Pfizer

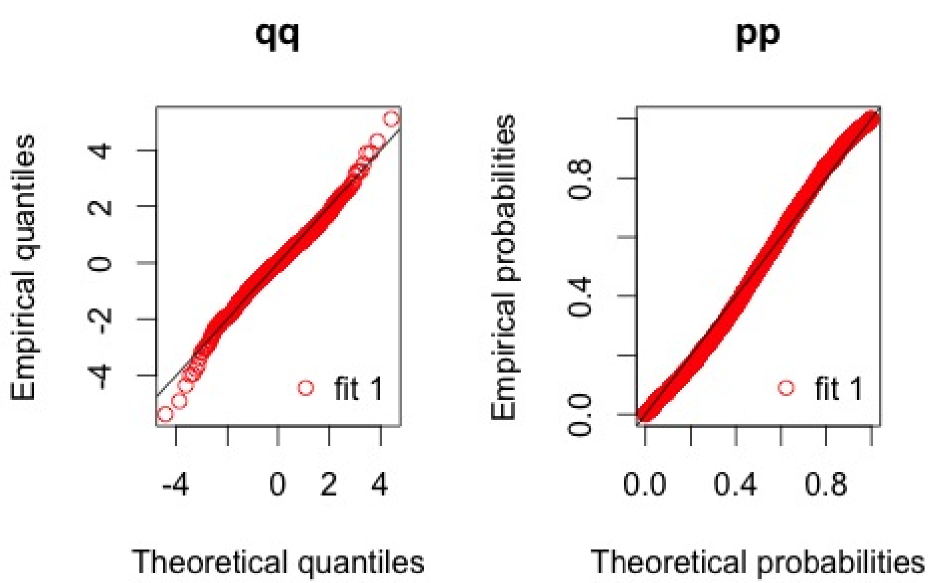

Figure 2.2 QQ and PP plots for US Airways

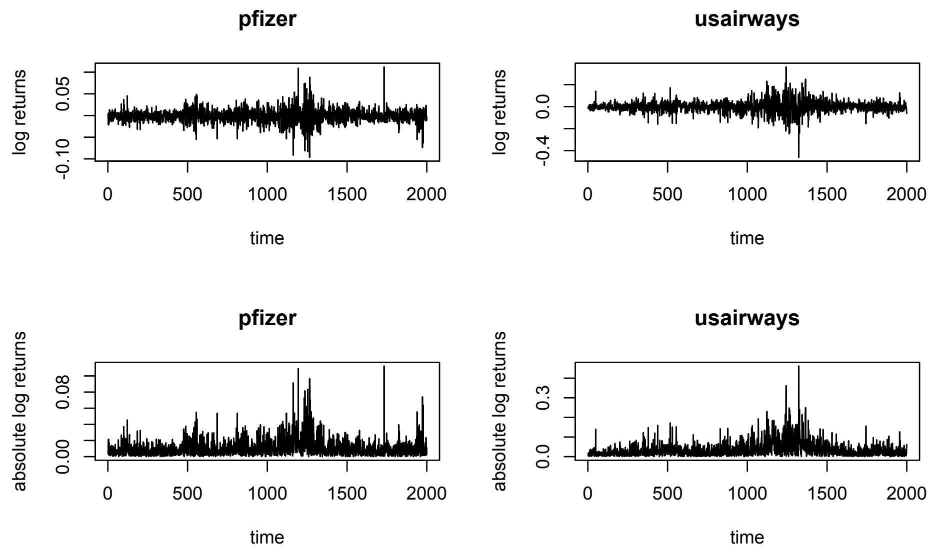

Figure 2.3 Log and absolute log returns of two sample stocks, Pfizer and US Airways

2.3 Modeling Dependence Structure

Looking at plots of the bivariate residuals, there appears to be a small positive dependence between the two series. Most of the points are concentrated in the center with large values for each stock appearing at different time periods.

Table 2.3 are Kendall tau estimates from different copulas fit across the different standardized GARCH residuals (rows). The Kendall tau of the log returns from the two stocks is 0.19 and the correlation is 0.34. It appears that the GARCH model has little to do with the correlation estimates.

Table 2.3 Kendall's Tau of 30 models fit to Pfizer and US Airways

Garch Model	Normal	t	Clayton	Frank	Amh	Joe
1	0.2042	0.1922	0.1512	0.1952	0.1755	0.1650
2	0.2043	0.1923	0.1514	0.1952	0.1757	0.1650
3	0.2039	0.1919	0.1512	0.1945	0.1752	0.1642
4	0.2039	0.1919	0.1507	0.1950	0.1749	0.1649
5	0.2044	0.1921	0.1520	0.1954	0.1760	0.1646

The GARCH models seem to translate the dependence structure in the same way. However, there is a noticeable difference between some of the copula fittings. *fit copulas to stock returns and compare kendall's tau to ones here*

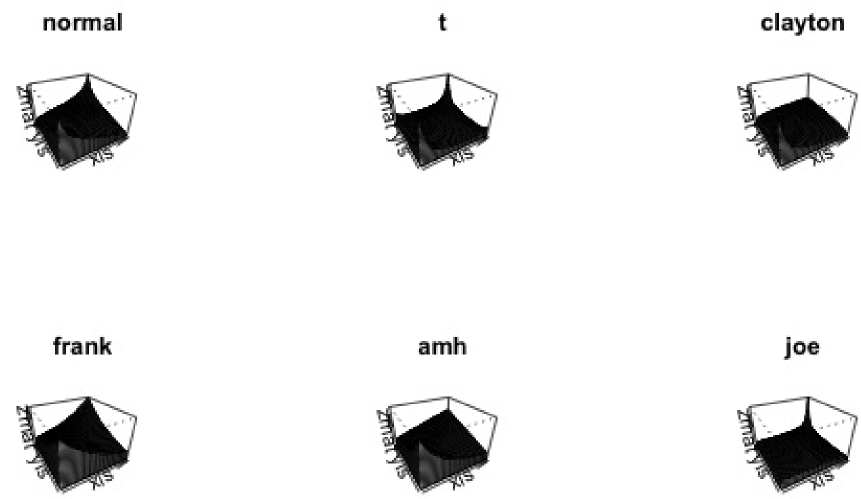

Figure 2.4 Perspective plots of density of copulas fit to standardized garch(1,1) residuals of Pfizer and US Airways

Figure 2.4 show perspectives plots of the densities of the six different copulas. The Normal, t, and Frank copulas have similar association estimates and allow for negative dependence. All three "lift" up in the bottom left and upper right corners of the perspective plots, which imply that these copulas model high density when returns are in either extreme, very low or very high. The Clayton copula does not allow for negative dependence and has the lowest association measure, this seems to contradict previous literature which says that asset returns become more correlated during deep bear markets. The source of this inconsistency may be because the data was fit over a very large period of time, 2000 tradings days which include periods with a diverse range of returns and economic environments. This form of the Joe copula does models only correlations in the right upper corner, extreme positive returns and estimates an average correlation. High positive returns of Pfizer and US airways do not seem to be strongly correlated or modeling just the upper right corner does not capture enough

Table 2.4 Goodness of Fit statistics across 30 models for Pfizer and US Airways

Garch Model	Copulas					
	Normal	t	Clayton	Frank	Amh	Joe
aparch(1,1)	0.08611	0.05094	0.2732	0.1152	0.1570	0.07804
egarch(1,1)	0.08517	0.05073	0.2737	0.1141	0.1566	0.07904
gjr(1,1)	0.08416	0.04981	0.2670	0.1115	0.1524	0.07956
igarch(1,1)	0.08747	0.05223	0.2771	0.1162	0.1590	0.07519
garch(1,1)	0.08561	0.05132	0.2814	0.1146	0.1606	0.08072

Table 2.5 Counts of Chosen model using a Cramer von Mises type goodness of fit test based on the empirical process

GoF Test	Garch Model	Copulas					
		Normal	t	Clayton	Frank	Amh	Joe
Sn	aparch(1,1)	0	36	0	0	0	0
	egarch(1,1)	0	12	0	0	1	0
	gjr(1,1)	0	8	0	0	0	0
	igarch(1,1)	0	5	0	0	0	0
	garch(1,1)	0	4	0	0	25	0

information. It seems that copulas which can capture negative dependence and are sensitive to both extreme ends, estimate higher correlations.

2.4 Model Evaluation

Table 2.4 show Cramer von Mises type GoF statistics based on the empirical copula. This is a sample of GoF test statistics of a portfolio across all 30 models. The minimum of all of these statistics are taken and the corersponding GARCH copula model combination is considered to be the preferred model among all 30. A count of the preferred model combination is taken across all portfolios and presented in Table 2.5. Here the t copula is preferred, followed by the joe and normal copulas.

Table 2.5 are counts of the chosen model using the Cramer-Von-Mises type GoF test, "Sn", the summation of squared deviations from the empirical copula to the parametric copula. The "Sn" test prefers the t and Amh copulas. The goodness of fit statistics are often more sensitive to the copula selection than the GARCH selection. The GARCH models seem to have little impact on the test statistic but the Amh copula here is being selected significantly more when paired with the GARCH(1,1) model than other GARCH specifications.

Table 2.6 Counts of Chosen Model Based on Goodness of Fit Tests Using Rosenblatt Transformation

GoF Test	Garch Model	Copulas					
		Normal	t	Clayton	Frank	Amh	Joe
Sn	aparch(1,1)	0	21	0	0	0	16
	egarch(1,1)	0	0	11	0	0	6
	gjr(1,1)	0	0	3	0	0	6
	igarch(1,1)	0	1	2	0	0	5
	garch(1,1)	0	0	8	0	0	12
SnB	aparch(1,1)	1	26	0	0	0	1
	egarch(1,1)	21	0	0	0	0	0
	gjr(1,1)	0	7	0	0	0	0
	igarch(1,1)	1	8	0	0	0	0
	garch(1,1)	1	31	0	1	0	1
SnC	aparch(1,1)	0	21	0	0	0	0
	egarch(1,1)	1	14	0	1	0	0
	gjr(1,1)	0	8	0	0	0	0
	igarch(1,1)	0	6	0	0	0	0
	garch(1,1)	0	30	0	0	0	0
AnChisq	aparch(1,1)	0	36	0	0	0	0
	egarch(1,1)	0	11	0	0	0	0
	gjr(1,1)	0	9	0	0	0	0
	igarch(1,1)	0	13	0	0	0	0
	garch(1,1)	0	22	0	0	0	0
AnGamma	aparch(1,1)	1	30	0	0	0	0
	egarch(1,1)	1	15	0	0	0	1
	gjr(1,1)	0	7	0	0	0	0
	igarch(1,1)	0	14	0	1	0	1
	garch(1,1)	1	15	0	0	0	4
Total	aparch(1,1)	2	144	0	0	0	17
	egarch(1,1)	3	52	11	1	0	7
	gjr(1,1)	0	31	3	0	0	6
	igarch(1,1)	1	42	2	1	0	6
	garch(1,1)	2	98	8	1	0	17

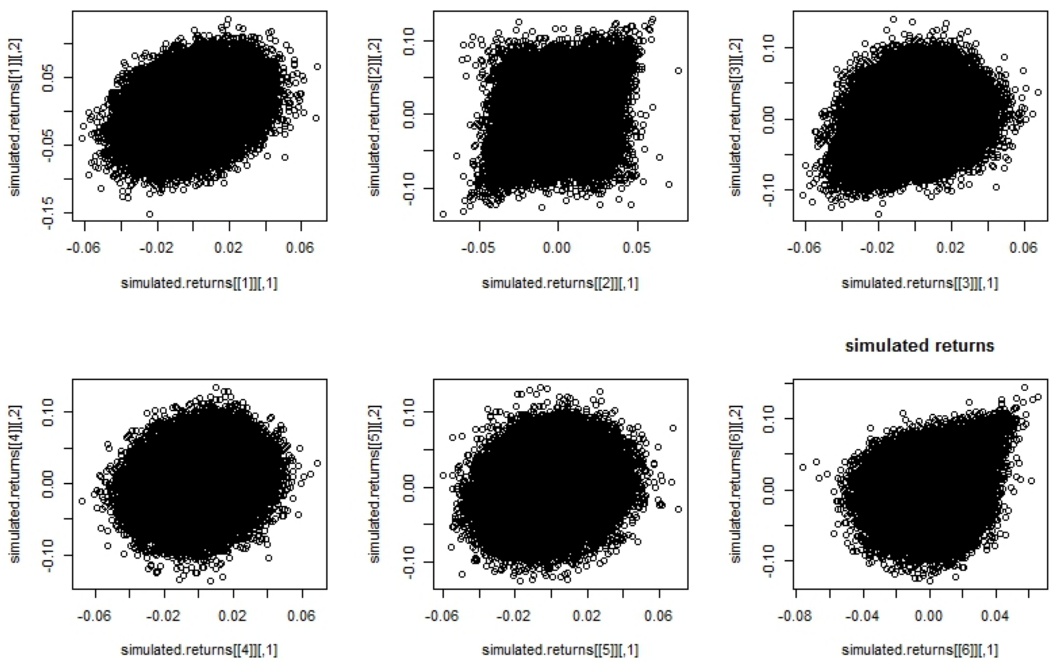

Figure 2.5 Sample of simulated returns of Pfizer and US airways, using different copulas fit to GARCH(1,1) residuals. From top left to bottom right, Normal, t , Clayton, Frank, AMH, and Joe copulas

2.5 Value at Risk

We consider bivariate portfolios of every possible pair out of the 14 assets, each asset in the portfolios have equal weights. We consider 5% and 1% one day VaR estimates and backtested the method at 95% and 99% confidence levels. Firstly, garch models are fit to the log returns of each asset, then pairing the standardized residuals in each portfolio, six different copulas are used to estimate the joint distribution of the pair of vectors of innovations. One day ahead garch forecasts are multiplied by one million random numbers pulled from the copulas to simulate one million possible portfolio returns. The 0.01 and 0.05 quantiles are taken of the simulated portfolio returns and those are the model VaR estimates.

Table 2.7 Number of portfolios where the 5% VaR model for n=1000 trading days was not rejected by the Kupiec test for all models.

Garch Model	Copulas					
	Normal	t	Clayton	Frank	Amh	Joe
aparch(1,1)	26	25	26	25	22	18
egarch(1,1)	28	27	28	29	26	18
gjr(1,1)	27	25	27	26	23	17
igarch(1,1)	16	16	16	16	14	13
garch(1,1)	30	28	32	31	24	25

Table 2.8 Number of portfolios where the 1% VaR model for n=1000 trading days was not rejected by the Kupiec test for all models.

Garch Model	Copulas					
	Normal	t	Clayton	Frank	Amh	Joe
aparch(1,1)	25	32	32	21	18	23
egarch(1,1)	33	34	33	23	27	22
gjr(1,1)	30	34	38	24	22	22
igarch(1,1)	0	0	0	0	0	0
garch(1,1)	22	23	22	20	13	26

Table 2.9 Number of portfolios where the 5% VaR model for n=250 trading days was not rejected by the Kupiec test for all models.

Garch Model	Copulas					
	Normal	t	Clayton	Frank	Amh	Joe
aparch(1,1)	51	52	51	49	37	47
egarch(1,1)	57	61	57	58	42	60
gjr(1,1)	45	46	46	45	36	42
igarch(1,1)	0	0	0	0	0	0
garch(1,1)	29	31	31	31	22	38

Table 2.10 Number of portfolios where the null hypothesis of indepedence was not rejected by the Christoffersen test for n=250 trading days.

Garch Model	Copulas					
	Normal	t	Clayton	Frank	Amh	Joe
aparch(1,1)	73	73	75	73	62	62
egarch(1,1)	83	81	83	80	65	73
gjr(1,1)	77	76	77	75	60	70
igarch(1,1)	51	49	51	50	46	45
garch(1,1)	78	78	81	78	64	76

Table 2.11 Number of portfolios where the null hypothesis of indepedence was not rejected by the Christoffersen test for n=1000 trading days.

Garch Model	Copulas					
	Normal	t	Clayton	Frank	Amh	Joe
aparch(1,1)	7	7	8	7	6	4
egarch(1,1)	6	4	7	5	5	1
gjr(1,1)	14	13	15	13	10	8
igarch(1,1)	19	17	19	18	17	13
garch(1,1)	38	36	38	38	33	25

2.6 Summary

In this paper we determine which goodness of fit statistic wouild be most appropriate for in sample selection of GARCH copula models when applied to value at risk backtesting. We took log returns of 14 US assets over 2000 trading days and paired each asset with every other asset creating 91 bivariate portfolios. 5 GARCH models are paired with 6 copulas making a total of 30 models fit to those 91 portfolios. Goodness of fit testing was conducted using the empirical copula and the Rosenblatt transformation. The Cramer von Mises type test based on the empirical process says that the t copula is usually the best fit of the data. Similarly, the 7 goodness of fit tests are conducted under the Rosenblatt transformation also conclude that the t copula is usually the best fit of the data. In general, it is recommended that t copulas be used to capture the dependence structure between assets. The VaR backtesting was inconclusive. There does not appear to be a clear "winner" among the 30 models. The percentage of non-rejected models across the 91 portfolios are very low. Previous papers usually get percentages of non-rejected models to be between 60-80 percent. Some possibilities for poor VaR backtesting results is lack of fit of marginal distributions. Maximum likelihood estimation of the degrees of freedom of the t distributions fit to the marginals of the GARCH residuals sometimes resulted in extremely low degrees of freedom. Empirical marginal distribution may result in better estimates. Also, the data came from Yahoo finance which has been known to have questionable accuracy at times. More work needs to be done before the reliability of these backtesting results can be established. Extensions of this work would add a sliding window to the VaR backtesting. The time window can span 250 trading days over the past 2000 days over non-overlapping time periods. The GARCH-copula models can be benchmarked against alternative volatility forecasting models, such as, DCC and HAR models.

BIBLIOGRAPHY

[1] Gregor N.F. Weib (2011) "Are Copula-GoF-tests of any practical use? Empirical evidence for stocks, commodities and FX futures" *The Quarterly Review of Economics and Finance*, 51, 173-188.

[2] Gregor N.F. Weib (2013) "Copula-GARCH versus dynamic conditional correlation: an empirical study on VaR and ES forecasting accuracy" *Review of Quantitative Finance and Accounting*, 41 (2), 173-188.

[3] Daniel Berg (2009) "Copula goodness-of-fit testing: an overview and power comparison." *The European Journal of Finance*, 15 (7-8), 675-701.

[4] Sebastien Laurent, Jeroen V.K. Rombouts, and Rrancesco Violante (2012). "On the Forecasting Accuracy of Multivariate Garch Models." *Journal of Applied Econometrics,* 27, 934-955.

[5] Peter Reinhard Hansen, Asger Lunde, and James M. Nason (2011). "Model Confidence Sets for Forecasting Models." *Econometrica*, 79 (2), 453-497.

[6] Robert Engle (2002). Dynamic Conditional Correlation Simple Class of Multivariate Garch Models." *Journal of Business and Economic Statistics*, 20 (3), 339-350.

[7] Thierry Ane and Cecile Kharoubi (2003). "Dependence Structure and Risk Measure." *Chicago Journals*, 76 (3), 411-438.

[8] P. Deheuvels (2009). "A Multivariate Bahadur-Kiefer Representation for the Emperical Copula Process." *Journal of Mathematical Sciences*, 163 (4).

[9] Ngoga Bob(2013). "Value at Risk Estimation. A Garch-EVT-Copula Approach" *Mathematiska institutionen*, 2013 (6).

[10] Gary Venter, Jack Barnett, Rodney Kreps and John Major (2007). "Multivariate Copulas for Financial Modeling." *Variance Advanding the Science of Risk*, 1 (1), 103-119.

[11] Pranesh Kumer (2010). "Probability Distributions and Estimation of Ali-Mikhail-Haq Copula" *Applied Mathematical Sciences*, 4 (14), 657-666.

[12] Alex Huang. "Backtesting Value at Risk Models"