Stevens Institute of Technology

Financial Engineering

Systemic Risk Analysis and Visualization

Author:
Jordyn Auge
Yang Li
Chang Le

Supervisor:
Dr. Khaldoun Khashanah

May 10, 2018

Abstract

In this paper the Vine Copula-Based ARMA-EGARCH Model introduced by Chen and Khashanah (2017) is applied to ten market sectors and the S&P500 from 2006 to 2017 to create an early indicator of systemic risk with visual representations. In the financial system there are multiple market risk measures, however it is difficult to capture a systemic risk measure, much less an indicator of distress and potential collapse. The vine copula-based ARMA-EGARCH method provides a means to investigate multi-dimensional marginal distributions with predictive power based on in sample and out of sample data sets. The model predicts systemic risk up to 360 days ahead, shows the size of the shocks in the financial system, and identifies which sectors contributed most systemic risk in time of crisis. The model will be compared to basic forecasted value at risk (VaR) to confirm the superior nature and potential for improvement.

Keywords: vine copula, value at risk, ARMA-EGARCH, systemic risk, risk measures, financial crisis

1 Introduction and Related Work

Following recent financial crises (e.g. 2008 subprime mortgage crisis, and 2010 flash crash), systemic risk reemerged in the research spotlight in an effort to prevent or, at minimum, proactively prepare for such crises. While there is no globally accepted definition of systemic risk, it is important to address the concept before attempting to quantify it. Smaga (2014) dissects multiple sources regarding systemic risk and compares each definition to obtain the commonalities and important features of this unique risk type. It is determined that systemic risk involves a substantial number of institutions and disrupts the performance and functions of the financial system; this is assumed to be a result of shocks within individual institutions that are interconnected, and ultimately has a negative effect on the real economy (Smaga, 2014). Given the general concept of systemic risk, it is now plausible to look at the basic systemic risk measure that is elaborated upon in this study, value at risk (VaR).

VaR is one of the most common market risk measures, used by many financial institutions, and in Basel II and III as a tool for financial regulation. VaR focuses on the marginal distribution of potential losses within an institution, and provides the maximum loss of an institution at α confidence level (Brunnermeier and Adrian, 2011). While VaR provides valuable information about the risk exposure of an individual bank, this is an isolated value and it does not incorporate the effect on the financial system as a whole. Conditional or Comovement VaR (CoVaR) is an idea introduced by Brunnermeier and Adrian (2011), where the VaR is expanded upon to include information about the financial system as a whole. CoVaR captures the risk contribution of one institution on the system, providing a basic systemic risk measure.

Hakwa (2011) and Hakwa et al. (2012) introduce the idea of a modified CoVaR, providing a closed formula to calculate the CoVaR using copula theorem, which is further expanded up by Chen and Khashanah (2014) by using vine copula modeling in higher dimensions to calculate the CoVaR. Copula theory was introduced by Sklar (1959) and is advantageous in financial VaR calculations because it allows for the separation of univariate marginal distributions of a multivariate distribution from the dependence structure, allowing it to be more easily analyzed (Hakwa, 2011). Chen and Khashanah (2015) expanded upon the copula-based modelling of ΔCoVaR, to create the vine copula-based ARMA-EGARCH model which allows modelling in higher dimensions and a more sophisticated forecasting tool. ARMA-EGARCH (Autoregressive Moving Average Exponential General Autoregressive Conditional Heteroskedasticity) provides relatively accurate VaR predictions compared to other models because it allows for thick tail distributions and asymmetrical variance by forecasting values based on autocorrelations of the error term (Angelidis et al., 2004). In this paper, we explore the concept of copula modelling combined with ARMA-EGARCH from 2006 to 2017 and attempt to create a predictive systemic risk tool and identify the main contributors in the 2008 financial crisis, closely following Chen and Khashanah (2015).

In the rest of the report we show our methodology in Section 2, our numerical results in Section 3, and conclude in Section 4.

2 Methodology

2.1 VaR Ratio and ES Ratio

The definition of Value-at-Risk (VaR) is the worst loss expected of an asset or a portfolio to be suffered over a given period with a given probability of either 5%, 1% or 0.1%. To recognize which sector contributes more risk than others, we calculate the VaR ratio by

$$VaR_t^{i\to j} Ratio = \frac{VaR_t^i}{VaR_t^j} \tag{1}$$

The higher ratio shows that the sector provides more risk to the entire market. In addition, the expected shortfall ratio can be easily derived from the same methodology as above.

2.2 ARMA-EGARCH Model

Bollerslev (1986) extended the ARCH model proposed by Engle (1986) to the GARCH model. What is more, to overcome the leverage effect in financial time series and the nonnegativity constraints in the linear GARCH model, we use the Exponential GARCH model which introduced by Nelson (1991) in dealing with the asymmetric effects between positive and negative asset returns.

$$X_t = \mu + \sum_{i=1}^{p} AR_i X_{t-i} + \sum_{j=1}^{q} MA_j \varepsilon_{t-j} + \varepsilon_t \tag{2}$$

$$\varepsilon_t = z_t \cdot \sigma_t \tag{3}$$

$$ln\sigma_t^2 = \omega + \beta \cdot ln\sigma_{t-1}^2 + \gamma \cdot \frac{\varepsilon_{t-1}}{\sqrt{\sigma_{t-1}^2}} + \alpha \cdot [\frac{|\varepsilon_{t-1}|}{\sqrt{\sigma_{t-1}^2}} - \sqrt{\frac{2}{\pi}}] \tag{4}$$

Where X is the log return, μ is the drift term, ε is the error term, γ captures the size effect, and z is the standardized residual with skewed students t distribution.
The ARMA-EGARCH model can be easily implemented by using R package 'rugarch' proposed by Ghalanos (2018).

2.3 Sklar's Theorem

Let F be a bivariate distribution with marginal distributions F1, F2.
There exists a two-dimensional copula C(·,·), such that

$$\forall (x_1, x_2) \in R^2 : F(x_1, x_2) = C(F_1(x_1), F_2(x_2)) \tag{5}$$

If F_1 and F_2 are continuous, the copula C is unique.
Since C_Y is the cumulative distribution function of $\{F_{Y_1}(Y_1), ..., F_{Y_d}(Y_d)\}$,

$$\begin{aligned} C_Y(u_1, ...u_d) &= P\{F_{Y_1}(Y_1) \le u_1, ..., F_{Y_d}(Y_d) \le u_d\} \\ &= P\{Y_1 \le F_{Y_1}^{-1}(u_1), ..., Y_d \le F_{Y_d}^{-1}(u_d)\} \\ &= F_Y\{F_{Y_1}^{-1}(u_1), ..., F_{Y_d}^{-1}(u_d)\} \end{aligned} \tag{6}$$

Next, letting $u_j = F_{Y_j}(y_j)$, for $j = 1, ..., d$, in 5 and interchanging the right and left sides, we see that

$$F_Y(y_1, ..., y_d) = C_Y\{F_{Y_1}(y_1), ..., F_{Y_d}(y_d)\} \tag{7}$$

2.4 Copula Family

2.4.1 Gaussian copula

$$\begin{aligned} C^{Gaussian}(u_1, u_2) &= \Phi_\rho(\Phi^{-1}(u_1), \Phi^{-1}(u_2)) \\ &= \int_{-\infty}^{\Phi^{-1}(u_1)} \int_{-\infty}^{\Phi^{-1}(u_2)} \frac{1}{2\pi\sqrt{1-\rho^2}} \exp \frac{2\rho st - s^2 - t^2}{2(1-\rho^2} dsdt \end{aligned} \tag{8}$$

The Gaussian copula only has one parameter ρ, which is the linear correlation coefficient.

2.4.2 Student's t copula

$$C^{student'st} = \int_{-\infty}^{u_1} \int_{-\infty}^{u_2} \frac{1}{2\pi\sqrt{1-\rho^2}} (1 + \frac{s^2 + t^2 - 2\rho st}{\nu(1-\rho^2)})^{-\frac{\nu+2}{2}} dsdt \tag{9}$$

The Student's t copula has two parameter ρ and ν, where ν is usually known as degree of freedom.

2.4.3 Clayton copula

$$C^{Clayton}(u_1, u_2) = (u_1^{-\theta} + u_2^{-\theta} - 1)^{-\frac{1}{\theta}} \tag{10}$$

where the Clayton generator is $\varphi(u) = u^{-\theta} - 1$ with $\theta \in (0, \infty)$.

2.4.4 Gumbel copula

$$C^{Gumbel}(u_1, u_2) = \exp\left(-\{(-lnu_1)^{\theta} + (-lnu_2)^{\theta}\}^{\frac{1}{\theta}}\right) \tag{11}$$

where the Gumbel generator is $\varphi(u) = (-lnu)^{\theta}$ with $\theta \in [1, \infty]$.

2.4.5 Frank copula

$$C^{Frank}(u_1, u_2) = -\frac{1}{\theta} ln(1 + \frac{(e^{-\theta u_1} - 1)(e^{-\theta u_2} - 1)}{e^{-\theta} - 1}) \tag{12}$$

where the Frank generator is $\varphi(u) = ln[\frac{(e^{-\theta u}-1)}{(e^{-\theta}-1)}]$ with $\theta \in (-\infty, 0) \cup (0, \infty)$.

2.4.6 Joe copula

$$C^{Joe}(u_1, u_2) = 1 - (u_1^{-\theta} + u_2^{-\theta} - u_1^{-\theta} u_2^{-\theta})^{\frac{1}{\theta}} \tag{13}$$

where the Joe generator is $\varphi(u) = u^{-\theta} - 1$ with $\theta \in [1, \infty]$.

2.4.7 BB1 copula(Clayton-Gumbel)

$$C^{BB1}(u_1, u_2) = (1 + [(u_1^{-\theta} - 1)^{\delta} + (u_2^{-\theta} - 1)^{\delta}]^{\frac{1}{\delta}})^{-\frac{1}{\theta}} \tag{14}$$

where $\theta \in (0, \infty) \cup \delta \in [1, \infty)$.

2.4.8 BB6 copula(Joe-Gumbel)

$$C^{BB6}(u_1, u_2) = 1 - (1 - \exp\{-[(-ln(1 - u_1^{-\theta}))^{\delta} + (-ln(1 - u_2^{-\theta}))^{\delta}]^{\frac{1}{\delta}}\})^{\frac{1}{\theta}} \tag{15}$$

with $\theta \in [1, \infty) \cap \delta \in [0, \infty)$.

2.4.9 BB7 copula(Joe-Clayton)

$$C^{BB7}(u_1, u_2) = 1 - (1 - [(1 - u_1^{-\theta})^{-\delta} + (1 - u_2^{-\theta})^{-\delta} - 1]^{-1/\delta})^{1/\theta} \tag{16}$$

with $\theta \in [1, \infty) \cap \delta \in [0, \infty)$.

2.4.10 BB8 copula(Frank-Joe)

$$C^{BB8}(u_1, u_2) = \frac{1}{\delta}(1 - [1 - \frac{1}{1 - (1 - \delta)^{\theta}}(1 - (1 - \delta u_1)^{\theta})(1 - (1 - \delta u_2)^{\theta})]^{1/\theta} \tag{17}$$

with $\theta \in [1, \infty) \cap \delta \in [0, \infty)$.

2.5 High Dimensional Copula

2.5.1 Copula Density(2-Dimensional

$$c_{12}(u_1,u_2) = \frac{\partial^2 c_{12}(u_1,u_2)}{\partial u_1 \partial u_2} \tag{18}$$

This implies,
Joint density:

$$f(x_1,x_2) = c_{12}(F_1(x_1),F_2(x_2)) \cdot f_1(x_1) \cdot f_2(x_2) \tag{19}$$

Conditional density:

$$f(x_2|x_1) = c_{12}(F_1(x_1),F_2(x_2)) \cdot f_2(x_2) \tag{20}$$

2.5.2 Pair-Copula Constructions and Vine-Copula Construction

We can represent a density $f(x_1,\dots,x_d)$ as a product of pair copula densities and marginal densities. For example, d = 3 dimensions. One possible decomposition of $f(x_1,x_2,x_3)$ is:

$$f(x_1,x_2,x_3) = f_{3|12}(x_3|x_1,x_2) \cdot f_{2|1}(x_2|x_1) \cdot f_1(x_1) \tag{21}$$

$$f_{2|1}(x_2|x_1) = c_{12}(F_1(x_1),F_2(x_2)) \cdot f_2(x_2) \tag{22}$$

$$f_{3|12}(x_3|x_1,x_2) = c_{13|2}(F_{1|2}\left(F_{1|2}(x_1|x_2),F_{3|2}(x_3|x_2)\right) \cdot f_{3|2}(x_3|x_2) \tag{23}$$

$$f_{3|2}(x_3|x_2) = c_{23}(F_2(x_2),F_3(x_3)) \cdot f_3(x_3) \tag{24}$$

After rearranging the terms, the joint density can be written as

$$\begin{aligned} f(x_1,x_2,x_3) &= f_3(x_3) \cdot f_2(x_2) \cdot f_1(x_1) \quad (\mathit{marginals}) \\ &\times c_{12}(F_1(x_1),F_2(x_2)) \cdot c_{23}(F_2(x_2),F_3(x_3)) \quad (\mathit{unconditional\ pairs}) \\ &\times c_{13|2}\left(F_{1|2}(x_1|x_2),F_{3|2}(x_3|x_2)\right) \quad (\mathit{conditional\ pair}) \end{aligned} \tag{25}$$

Joe (1996), Bedford and Cooke (2002), Aas et al. (2009) summarized the formula of general d-dimensional regular vine copula, in which one variable plays a pivotal role, can be written as

$$f(x_1,\dots,x_d) = \prod_{j=1}^{d-1} \prod_{i=1}^{d-j} c_{i,(i+j)|(i+1),\dots,(i+j-1)} \cdot \prod_{k=1}^{d} f_k(x_k) \tag{26}$$

With

$$c_{i,j|i_1,\dots,i_k} = c_{i,j|i_1,\dots,i_k}\left(F(x_i|x_{i_1},\dots,x_{i_k}),F(x_j|x_{i_1},\dots,x_{i_k})\right) \tag{27}$$

For $i,j,i_1,...,i_k$ with $i<j$ and $i_1 < \dots < i_k$

Remark that the decomposition is <u>not</u> unique.

Bedford and Cooke (2002) introduced a graphical structure called regular vine structure to help organize them. Dissmann et al. (2013) introduced an automated algorithm including finding out an optimal R-vine tree structure, the pair-copula families, and the parameter values of each best pair-copula families based on AIC value. The practical procedure is as follows:

i. Calculate the Kendalls tau for all possible variable pairs

ii. Select the tree that maximized the sum of absolute values of taus

iii. Select the pair-copula and fit the corresponding parameters based on AIC

iv. Transform the observations using the copula and parameters from step iii to obtain the transformed values

v. Use transformed observations to calculate Kendalls taus for all possible pairs

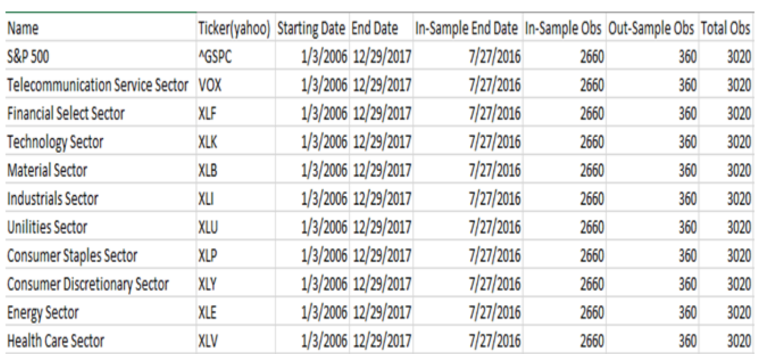
vi. Repeat step iii to v until the R-vine is fully specified

The above algorithm can be easily implemented by using VineCopula library proposed by Schepsmeier et al. (2018) in R.

3 Data and Empirical Findings

3.1 Data Representation

We decided to use indices prices instead of other financial accounting numbers because an index price could reflect a timely financial environment in contrast to financial accounting numbers that are published quarterly, and indices can easily be constructed and tell us which sector contributes more risk to the entire market. We use S&P 500 index as market proxy and 10 different sectors indices at daily frequency from January 1, 2006 to the last trading day of 2017. We divide them into in-sample and out-sample which serves as our forecast validation dataset.



Name	Ticker(yahoo)	Starting Date	End Date	In-Sample End Date	In-Sample Obs	Out-Sample Obs	Total Obs
S&P 500	^GSPC	1/3/2006	12/29/2017	7/27/2016	2660	360	3020
Telecommunication Service Sector	VOX	1/3/2006	12/29/2017	7/27/2016	2660	360	3020
Financial Select Sector	XLF	1/3/2006	12/29/2017	7/27/2016	2660	360	3020
Technology Sector	XLK	1/3/2006	12/29/2017	7/27/2016	2660	360	3020
Material Sector	XLB	1/3/2006	12/29/2017	7/27/2016	2660	360	3020
Industrials Sector	XLI	1/3/2006	12/29/2017	7/27/2016	2660	360	3020
Unilities Sector	XLU	1/3/2006	12/29/2017	7/27/2016	2660	360	3020
Consumer Staples Sector	XLP	1/3/2006	12/29/2017	7/27/2016	2660	360	3020
Consumer Discretionary Sector	XLY	1/3/2006	12/29/2017	7/27/2016	2660	360	3020
Energy Sector	XLE	1/3/2006	12/29/2017	7/27/2016	2660	360	3020
Health Care Sector	XLV	1/3/2006	12/29/2017	7/27/2016	2660	360	3020

Figure 1: Summary of Data

Index	Mean(%)	Sigma(%)	Skew	Kurt	ADF test	J-B test	ARCH test	Min(%)	Max(%)
S&P 500	0.0285	1.29	-0.0921	10.5	1	1	1	-9.03	11.6
Telecommunic…	0.0431	1.32	0.191	10.7	1	1	1	-8.1	14
Materials	0.0399	1.63	-0.0866	6.93	1	1	1	-12.4	14.1
Energy	0.0343	1.91	-0.157	9.89	1	1	1	-14.4	16.5
Financials	0.021	2.21	0.421	13.1	1	1	1	-16.7	16.5
Industrials	0.0409	1.4	-0.0868	6.15	1	1	1	-9.4	10.7
Technology	0.0438	1.33	0.261	10.2	1	1	1	-8.65	13.9
Consumer Staples	0.0456	0.873	-0.245	5.53	1	1	1	-6.02	6.89
Utilities	0.0392	1.15	0.511	12.9	1	1	1	-7.44	12.1
Health Care	0.0447	1.08	-0.0602	11.5	1	1	1	-9.78	12.1
Consumer Discretionary	0.0496	1.41	-0.206	6.74	1	1	1	-11.6	9.78

Figure 2: Summary Statistics of Data

Additionally, we do the summary statistics of these indices and statistical hypothesis testing. From the result of Augmented Dickey-Fuller test, we can tell that the log returns of all indices are not stationary. On the other hand, the distributions of all returns are non-normal based on the Jarque-Bera test. Finally, the results of Engles autoregressive conditional heteroskedasticity test show that the indices returns present conditional heteroscedasticity at 5% confidence level.

3.2 Result for Marginal Models

We fit the ARMA-EGARCH model with log return of each index repeatedly to obtain the best parameters of p and q ranging from 0 to 5 by selecting the model with the least AIC values. Then Jarque-Bera test and ARCH test are implemented with the residuals. The results show that all degree of freedom are smaller than 18 and Jarque-Bera test rejects the null hypothesis, which refers to that using the skewed Students t distribution for residuals is appropriate.
We are intended to compare the VaR prediction performance between ARMA-EGARCH and Copula-based ARMA-EGARCH. So before implemented the Copula function, we utilize fitted models to predict the 10-day rolling VaR and plot the actual returns from out-sample dataset and the predicted VaR curve together. Based on the report, ARMA-EGARCH based predictions are precise under 5% confidence level.

	^GSPC	VOX	XLF	XLK	XLB	XLI	XLU	XLP	XLY	XLE	XLV
p	3.0000000000	4.0000000000	0.0000000000	3.000000000000	4.0000000000	2.0000000000	1.0000000000	3.0000000000	2.0000000000	1.0000000000	0.0000000000
q	2.0000000000	4.0000000000	1.0000000000	4.000000000000	5.0000000000	2.0000000000	1.0000000000	2.0000000000	2.0000000000	0.0000000000	2.0000000000
ar1	1.6185759366	0.3207407532	NA	1.182236954261	0.1259885834	-0.6149551571	0.7961047342	1.5227558655	1.9228509699	-0.0426604808	NA
ar2	-0.5876721994	1.1696680843	NA	0.451416597760	-0.1069540310	-0.9684590648	NA	-0.5318109152	-0.9327830654	NA	NA
ar3	-0.0332180663	0.4189040592	NA	-0.648570224302	0.0766589207	NA	NA	0.0028796695	NA	NA	NA
ar4	NA	-0.9105040469	NA	NA	0.8940832838	NA	NA	NA	NA	NA	NA
ar5	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
ma1	-1.6949557747	-0.3356440991	-0.0784785352	-1.205000007504	-0.1621737124	0.6052089864	-0.8319548296	-1.6037692486	-1.9401826231	NA	-0.0462891990
ma2	0.6954016744	-1.1858105000	NA	-0.439652269033	0.1139606175	0.9562092927	NA	0.6056724181	0.9519303234	NA	-0.0398235699
ma3	NA	-0.4004920010	NA	0.686944411448	-0.0884246134	NA	NA	NA	NA	NA	NA
ma4	NA	0.9221796606	NA	-0.020974453776	-0.8882805607	NA	NA	NA	NA	NA	NA
ma5	NA	NA	NA	NA	0.0221908356	NA	NA	NA	NA	NA	NA
mu	0.0003769124	0.0003549661	0.0004069398	-0.000003274173	0.0003270501	0.0003124675	0.0004681162	0.0005237583	0.0002802002	0.0003209431	0.0004393508
kappa	0.7647797626	0.7751374885	0.7561214276	0.765664447532	0.7776742869	0.7756961400	0.7724099336	0.7784680263	0.7714836150	0.7867979973	0.7731761947
alpha	-0.2283592689	-0.0956956531	-0.1107922180	-0.174554017474	-0.1358236785	-0.1290508520	-0.0425180557	-0.1200577672	-0.1358291486	-0.0826557274	-0.1402248390
beta	0.9640274353	0.9800815225	0.9894133066	0.985453590511	0.9835220563	0.9844603901	0.9824691429	0.9616314372	0.9878525860	0.9867921481	0.9619779823
omega	-0.3415769764	-0.1829260571	-0.0923333144	-0.131191742357	-0.1385119628	-0.1387265765	-0.1632233065	-0.3764389111	-0.1096093736	-0.1107595717	-0.3576261639
gamma	0.0918062919	0.1190516094	0.1783824642	0.080230390905	0.0970293028	0.1357257986	0.1589114027	0.1522509021	0.1337148247	0.1353326518	0.1597157057
skew	0.8169849277	0.8587160725	0.9332386047	0.841799341668	0.8141674639	0.8681641587	0.8891436514	0.8596422386	0.8608662979	0.8655293215	0.8793380497
D.F	7.7145584093	10.3784364254	6.6019325246	7.905308648828	11.2165827170	10.6165068663	9.5613469592	11.7880773843	9.2366753734	18.8428739039	9.7736114374
LLH	8696.5860169713	8427.0486087253	7685.6753770302	8375.295508565487	7786.2163443988	8235.1457785148	8662.2047047149	9310.4521487400	8279.7454784508	7415.3753859130	8787.6711557159
AIC	-6.5322196442	-6.3272272348	-5.7748592531	-6.289052657815	-5.8444650954	-6.1858937785	-6.5086157990	-6.9939467083	-6.2194399988	-5.5715497450	-6.6029869543
J-B test	1.0000000000	1.0000000000	1.0000000000	1.000000000000	1.0000000000	1.0000000000	1.0000000000	1.0000000000	1.0000000000	1.0000000000	1.0000000000
ARCH test	0.0000000000	0.0000000000	0.0000000000	0.000000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000

Figure 3: First Calibration of Parameters

3.3 Results for the Copula VaR and Copula VaR Ratio

Before implementing the copula functions, we transformed the standardized residuals z_t to be uniform distributed residuals u_t which would be used as the input of copula functions. However, to confirm that the transformation of distribution of standardized residuals would not change the correlation coefficients between each index, we plot the correlation matrix of both type of residuals and the result indicates the transformation does not change the correlation coefficients and all transformed residuals pass the Kolmogorov-Smirnov test.

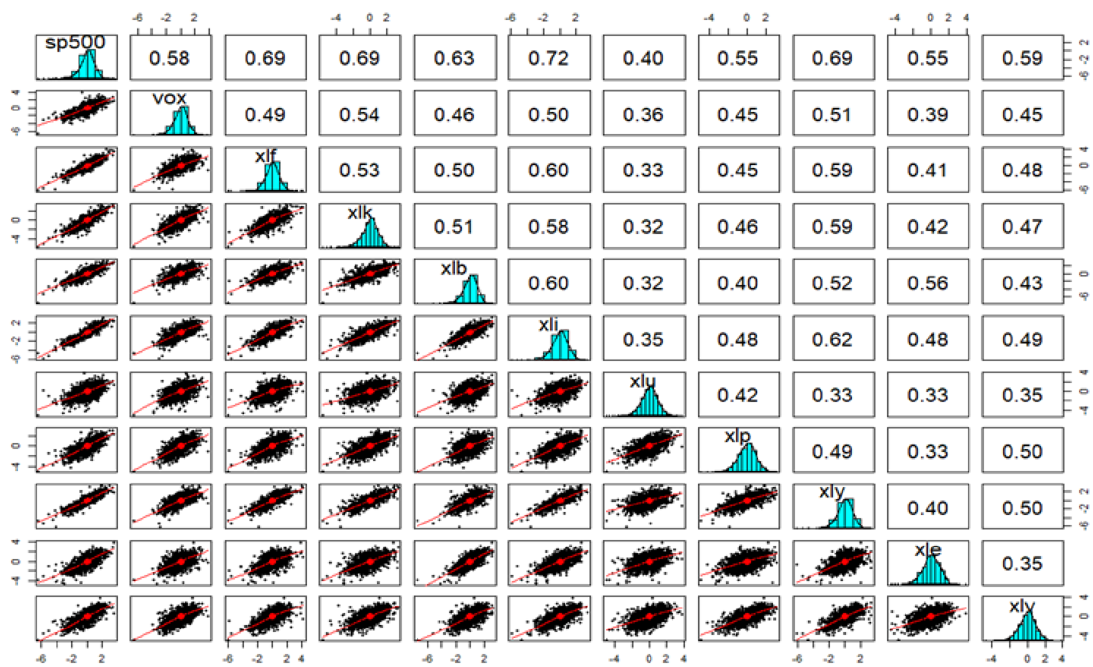


Figure 4: Standardized Residuals Correlation Matrix

Moreover, by comparing the AIC values of fitted Gaussian Copula, t Copula and R-Vine Copula in high-dimensional modelling, the R-Vine Copula outperforms other two copulas based on the AIC value and the log-likelihood.

	Number of parameters	AIC	LogLik
Gaussian	55	-32168.2	16139.12
t-student	110	-33456.8	16838.41
Vine	98	-33614.5	16905.25

Figure 5: Performance of each kind of copula

Using the R-Vine copula, the tree structure in level one and two is shown as follows:

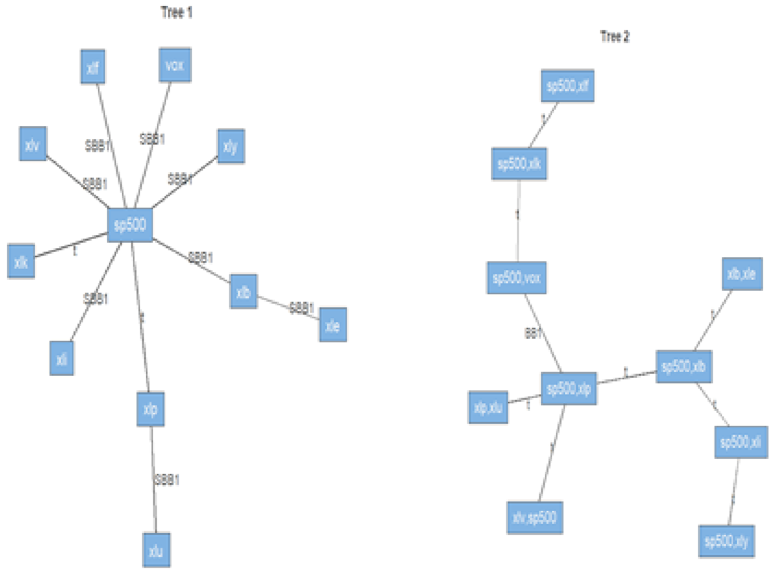


Figure 6: Vine Copula Trees

What is more, we simulate the u_t for each index then all u_t are quantile transformed back to z_t. Following the formula of ARMA-EGARCH model with estimated parameters, the log return X_t of each index can be calculated. However, the above procedure can only output 1-days log return. So, all procedure should be implemented repeatedly since the models have to refit every day so that the new parameters can be obtained. Finally, all 360-day ahead log returns are calculated and the 10-day VaR can be estimated. For the next step, we compute the VaR ratio with respect to the overall time range we used and 2008 financial crisis. Consequently, we can notice that the Energy sector contributed most risk to the entire market from 2006 to 2017 while the Financial sector has the largest VaR ratio during financial crisis.

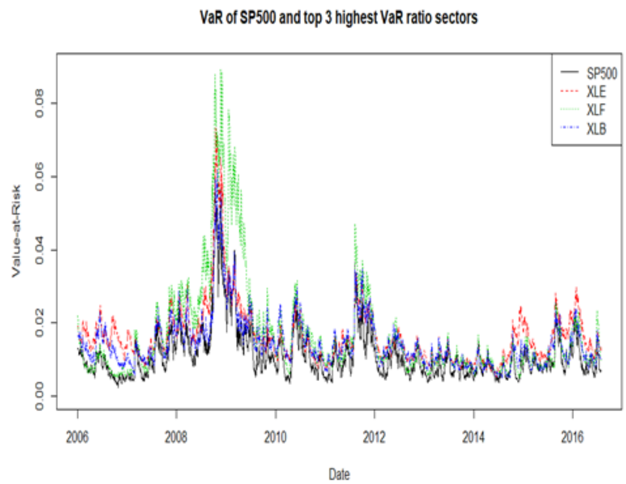


Figure 7: VaR of S&P500 and Top 3 VaR Ratio Risk Contributors

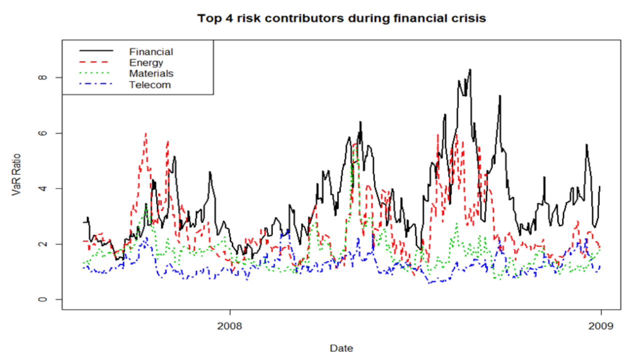


Figure 8: VaR Ratio of Top 4 Risk Contributors During the Financial Crisis of 2008

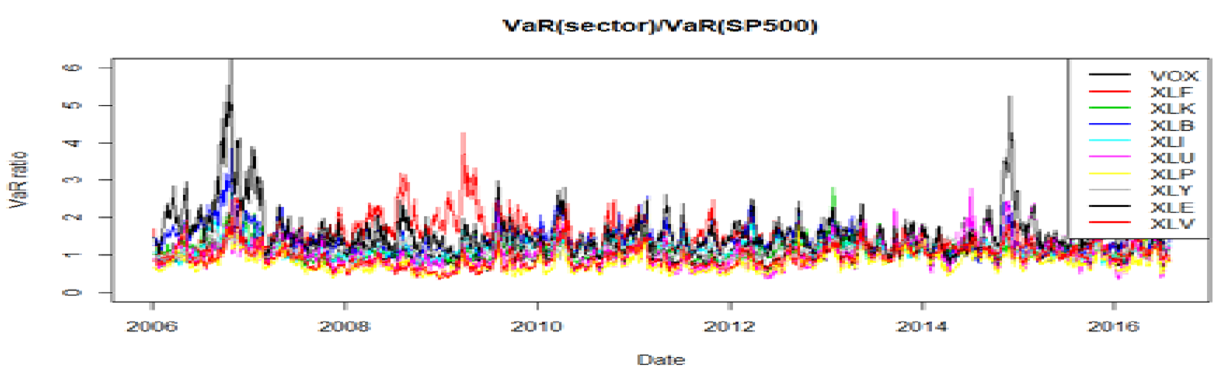


Figure 9: In-Sample VaR Ratio

For the purpose of distinguishing the difference between VaR Ratio and CVaR Ratio, we made VaR Ratio barplot 10 and CVaR Ratio barplot 11 by year separately, as well as the comparison graph 12. From the observation, most of CVaRs are smaller than VaR, which makes sense since the distribution of financial return data are usually fat-tailed.

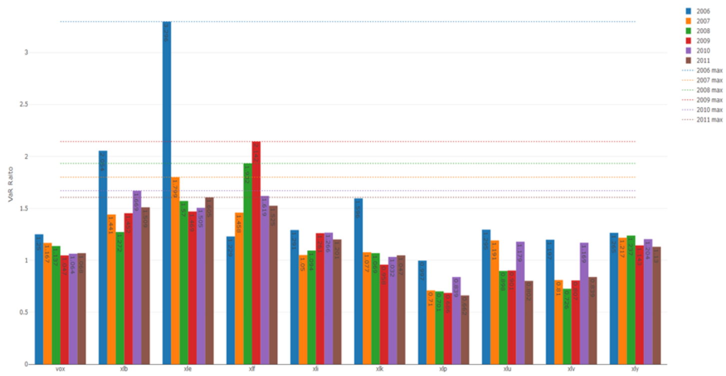


Figure 10: VaR Ratio by Year

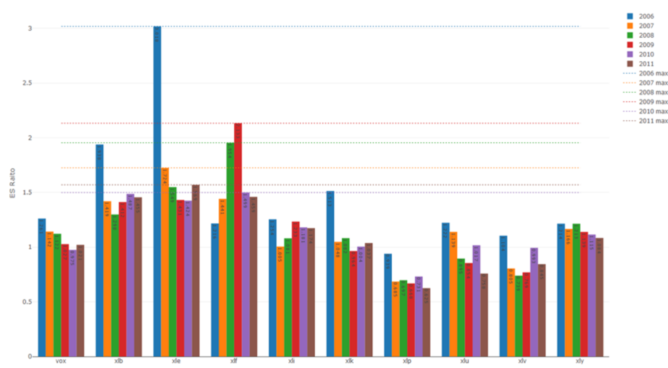


Figure 11: ES Ratio by Year

Figure 12: VaR vs. ES for Each Sector

Furthermore, we implemented five machine learning algorithms and selected the optimal model with minimal error for doing the forecast as well. From the result we can recognize that the Support Vector Machine(SVM) is the best model for our data. However, by comparing to the ARMA-EGARCH model and Copula-based ARMA-EGARCH model, the forecasting performance is terrible. The forecast curve fluctuates randomly and the actual exceeds are usually larger than the expected exceeds with 1% confidence level.

	Linear	Ridge	Bayes	Tree	SVM	min error	min model
sp500	171.059376	170.413953	170.452997	256.065926	120.539505	120.539505	SVM
vox	283.996136	283.891205	283.920718	400.567816	247.247825	247.247825	SVM
xlf	299.399637	296.529079	296.543668	380.586151	250.531667	250.531667	SVM
xlk	232.531402	232.491872	232.533860	283.438582	175.081985	175.081985	SVM
xlb	255.674190	256.038791	256.075294	347.472234	185.031670	185.031670	SVM
xli	226.803736	226.756790	226.797881	319.399657	165.268515	165.268515	SVM
xlu	254.927106	254.864597	254.914683	376.850857	211.767394	211.767394	SVM
xlp	188.252173	188.093685	188.164995	274.350107	152.016060	152.016060	SVM
xly	203.400636	203.463169	203.490330	286.470826	156.749098	156.749098	SVM
xle	329.679684	329.668597	329.692545	520.103999	260.932353	260.932353	SVM
xlv	235.551096	235.614893	235.672567	338.162302	174.276179	174.276179	SVM

Figure 13: Machine Learning Algorithms Comparison

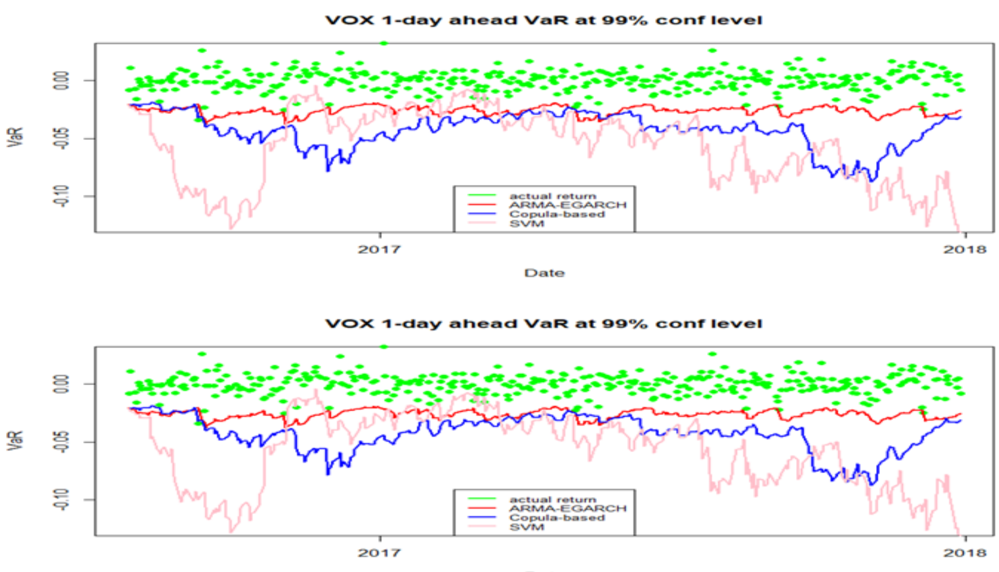


Figure 14: 1 Day Ahead VaR Prediction Comparison

	Advantage	Shortcoming
ARMA-EGARCH	1: Easier to implement and forecast VaRs and CVaRs with given packages 2: The predicted VaRs are more close to the actual VaRs	1: The interdependence of returns of multiple assets cannot be calculated when they are non-normal 2: Actual exceeds are sometimes larger than the expected exceeds
Vine-Copula ARMA-EGARCH	1: Good to handle the skewed or asymmetric distributions 2: Actual exceeds are always smaller than the expected exceeds	1: Extremely long time for debugging in R 2: The predicted VaRs are sometimes way more larger than the actual VaRs

Figure 15: Comparison of ARMA-EGARCH and Copula-based ARMA-EGARCH

Last but not least, we tried to construct yearly portfolios from 2006 to 2017 and compared the VaRs of portfolios to the VaRs of S&P500 index. Theoretically, they should be the same. However, since we used fixed weight for each year, there are some different between their VaRs, but it's pretty subtle.

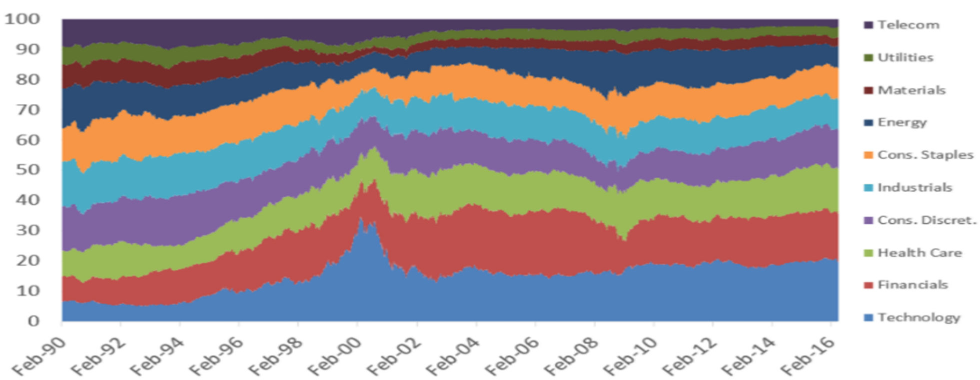


Figure 16: S&P500 Sector Weights by Year

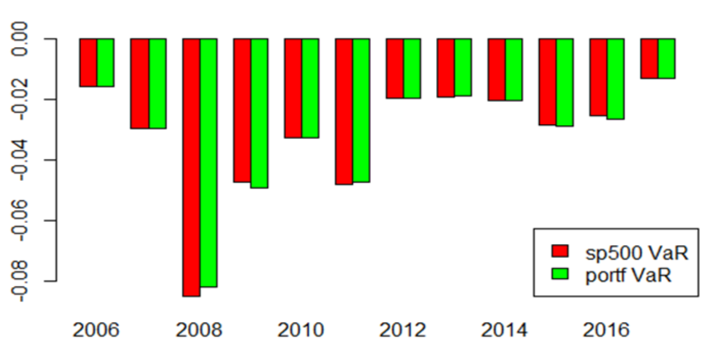


Figure 17: S&P500 VaR vs. Sector Portfolio VaR

Moreover, we estimated which sector had risk contribution percentage that was larger than their weights. From the graph below, we can tell that the contribution percentage of financial sector were always larger than their weights. On the other hand, the contribution percentage of energy sector is larger than its actual weights before 2016 but began to contribute less risk with respect to its weights.

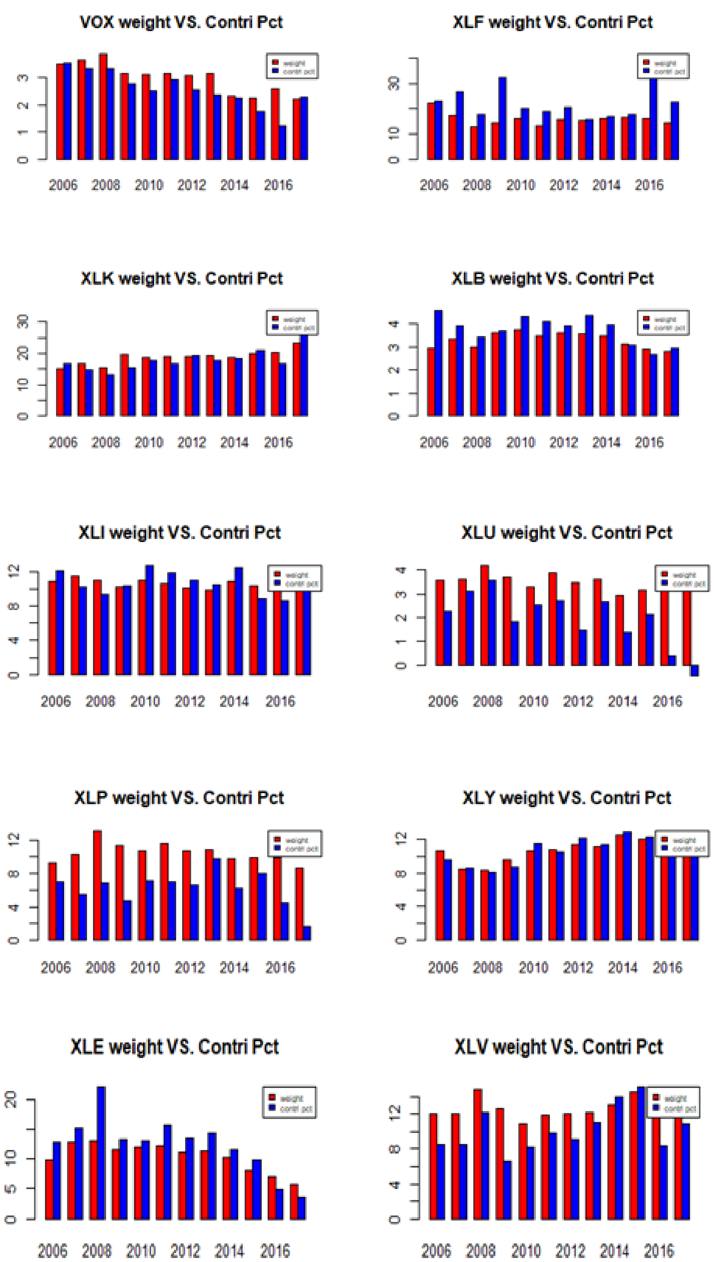


Figure 18: Risk contribution percentage VS. Weights

4 Conclusion

The Vine-Copula based ARMA-EGARCH has the advantage of handling skewed or asymmetric distributions and the actual exceeds are always smaller than the expected exceeds. However, comparing to the original ARMA-EGARCH model, the predicted VaR from Vine-Copula based ARMA-EGARCH are sometimes too larger than the actual VaR even though the actual exceeds are always smaller than the expected exceeds. By comparing to the ARMA-EGARCH model and Copula-based ARMA-EGARCH model, basic machine learning algorithms cannot well forecast the Value at Risk based on the actual exceeds and expected exceeds.

Moreover, observing the risk contribution percentage and weight of each sector, financial sector is the most risky sector since it always contributes exceed risk to the S&P500 index with respect to its weight.

Additionally, we find out an easier method to evaluate systemic risk, as well as the risk contributors by using Vine-Copula based ARMA-EGARCH model to forecast VaR and VaR ratio.

References

Aas, K., C. Czado, and H. B. Arnoldo Frigessi (2009). Pair-copula constructions of multiple dependence. *Insurance:Mathematics and economics*, 182–198.

Angelidis, T., A. Benos, and S. Degiannakis (2004). The use of garch models in var estimation. *Statistical methodology 1*(1-2), 105–128.

Bedford, T. and R. Cooke (2002). Vines: a new graphical model for dependent rando variables. *Ann Stat*, 1031–1068.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 307–327.

Brunnermeier, M. and T. Adrian (2011). Covar. *Federal Reserve Bank of New York Staff Reports 348*.

Chen, K.-H. and K. Khashanah (2014). Measuring systemic risk: copula covar.

Chen, K.-H. and K. Khashanah (2015). Analysis of systemic risk: A dynamic vine copula-based arma-egarch model. In *The World Congress on Engineering and Computer Science*, pp. 15–30. Springer.

Dissmann, J., E. Brechmann, C. Czado, and D. Kurowicka (2013). Selecting and estimating regular vine copula and application to financial returns. pp. 52–69.

Engle, R. F. (1986, July). Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica* (4).

Ghalanos, A. (2018). *rugarch: Univariate GARCH models.* R package version 1.4-0.

Hakwa, B. (2011). Measuring the marginal systemic risk contribution using copula.

Hakwa, B., M. Jäger-Ambrożewicz, and B. Rüdiger (2012). Measuring and analysing marginal systemic risk contribution using covar: A copula approach. *arXiv preprint arXiv:1210.4713*.

Joe, H. (1996). Families of multivariate distributions with given margins and m(m-1)/2 bivariate dependence parameters. *Lecture notes-monograph series*, 120–141.

Nelson, D. (1991). Conditional heteroskedasticity in asset returns: a new approach. *Econometrica*, 347–370.

Schepsmeier, U., J. Stoeber, E. C. Brechmann, B. Graeler, T. Nagler, and T. Erhardt (2018). *VineCopula: Statistical Inference of Vine Copulas*. R package version 2.1.4.

Sklar, M. (1959). Fonctions de repartition an dimensions et leurs marges. *Publ. Inst. Statist. Univ. Paris 8*, 229–231.

Smaga, P. (2014). The concept of systemic risk.