

# Copula Models for Equity Portfolio Risk Estimation: A Case Study of Nairobi Securities Exchange

Maina Stanley Muchiri\*, Mung'atu Joseph Kyalo, Wanjoya Antony Kibera

Department of Statistics and Actuarial Sciences, Jomo Kenyatta University of Agriculture and Technology, Nairobi, Kenya

## Email address:

stanmaina54@gmail.com (M. S. Maina), j.mungatu@fsc.jkuat.ac.ke (M. J. Kyalo), awanjoya@gmail.com (W. A. Kibera)

\*Corresponding author

## To cite this article:

Maina Stanley Muchiri, Mung'atu Joseph Kyalo, Wanjoya Antony Kibera. Copula Models for Equity Portfolio Risk Estimation: A Case Study of Nairobi Securities Exchange. *International Journal of Statistical Distributions and Applications*. Vol. 7, No. 3, 2021, pp. 72-77. doi: 10.11648/j.ijstd.20210703.11

**Received:** July 31, 2021; **Accepted:** August 12, 2021; **Published:** August 23, 2021

---

**Abstract:** Mitigation of risk in a security-trading environment is one of the strategies that every investor would want. Stock prices keep on changing from time to time and it is difficult for participants to predict share price direction. It is therefore appropriate for any trader to use the risk minimization strategies. One of the strategies is to estimate the amount of money a trader is willing to lose in a given period. This project uses the concept of copula to describe the dependency structure of portfolio returns selected from the Nairobi Securities Exchange. The Copula concept works towards modeling the dependency structure of high volatile data by separating the univariate distribution of their respective marginals. Dependencies in a volatile data enables one to assess the relationship that appears in the extreme values of historically sampled data. The study started with selecting an optimal portfolio from the Nairobi Securities Exchange using the Capital Asset and Pricing Model. The chosen optimal portfolio involved companies from different sectors, thus implying minimization of unsystematic risk. The next step involved estimating the systematic risk based on the historical data of the optimal portfolio selected. Different types of copula-based models were fitted then compared to each other to assess the dependencies measures, the goodness of fit of the model, and backtest of the Value-at-Risk. The main findings were that the Tawn type 1 copula was the best copula-based model for modeling the dependence structure of the portfolio returns. Additionally, the backtesting results also show that this model had the highest coverage of the exceedances.

**Keywords:** Copula, Dependency, GARCH, Value-at-Risk, Volatility, Heteroscedasticity

---

## 1. Introduction

### 1.1. Background of the Study

The Nairobi Securities Exchange (NSE) is an approved institution charged with the responsibility of developing and monitoring securities trading activities. Its inception dates back to 1954 when it was registered as an association of stockbrokers under the Societies Act.

After the 2007-2008 financial crises in Kenya there was an urgent need to create regulations in the finance industry. Financial researchers have since then tried to use historical data to estimate the risk measures. There are commonly used risk measures such as Value-at-Risk (VaR) and Expected Shortfall (ES). This research worked towards using the assessing of dependencies of share prices using the concept

of copula. Copula concept as introduced by Sklar in 1959, helps to separate a multivariate cumulative distribution into the individual marginals. The copula concept is based on the nonlinear relationship between variables hence it was able to depict the true picture of relationship in clustered data points. Wali, Greene, Khattak, and Liu (2018) [1], noted that the marginals help to model the extreme losses incurred based on historically sampled data. Share prices are recorded frequently hence there are chances that the data points may cluster at a single point at a given time. Financial researchers have converged to one assertion that, high frequent data leads to a fat tailed distribution. At the lower tails of the distribution, we are able to estimate the minimum value a trader may lose at a given time horizon. Fama and French, (2017) [2] argued that semi parametric and non-parametric methods could help in assessing the risk measure in high

frequent data like the stock prices. The researchers claimed that parametric distribution makes assumption in the tail distribution of frequent data. The research study therefore adopts a non-parametric approach in measurement of portfolio risk by applying the copula, which does not make assumption on the distribution of assets returns.

There is high demand among decision makers in the investment field in order to make appropriate decisions for meeting the ever-increasing investment appetite among investors (Evans, 2016) [3]. Stockbrokers would like to reap heavily from profits earned by their clients/shareholders. The Government also would want to increase financial activities in trading of financial assets in an effort to improve the economy. If risk measurement is not managed well, investors will shy off leading to slow economic growth and no employment opportunities for stock market facilitators.

### 1.2. Statement of the Problem

The problem at hand was to quantify the amount of investment an investor intends to lose at a given time horizon by selecting an array of financial assets. It is the wish of every investor to minimize the risk as much as possible while at the same time maximize the investment profit. This study intended to measure the value that partitions the distribution of portfolio returns at 0.05 lower quartile. The single value that partitions the lower tails of a distribution is called lower quartile (Ye, Luo, & Liu, 2017) [4]. The value of this lower quartile forms the basis of quantifying the value-at-risk. With the value-at-risk, the investor will be at a position to estimate the average amount of investment to lose at a given time horizon.

### 1.3. Justification of the Study

This research study used the historical data for the share prices of Companies trading in the Nairobi Securities Exchange. Portfolio returns were assessed at the lower quartile to yield the Value-at-Risk. These financial risk measures played a key role in helping investors and traders become risk averse thus contributing a lot in risk measurement quantification.

A comparison of the copula-based models enabled the research to determine the best model for this kind of data. The subjection of the models to a backtesting technique formed the basis of hypothesis testing. The model with a significant backtesting results would help NSE participants in understanding investment risks and ensure risk averseness.

The findings of this research study will benefit any investor willing to make investment for future use and those aspiring to work as financial analysts in the Nairobi Securities Exchange. It will also be a starting point for researchers who wish to carry out further research studies in the field of risk management.

### 1.4. Objectives

#### 1.4.1. General Objective

The research study will work towards measuring the investment risk of a portfolio returns selected from

Companies trading in the Nairobi Securities Exchange using the concept of copula.

#### 1.4.2. Specific Objectives

- To select an optimal portfolio from the companies listed on the Nairobi Securities Exchange.
- To fit a copula-based model to returns of the optimal portfolio selected from the companies listed on the Nairobi Securities Exchange.
- To get the best type of copula-based model that effectively measures the portfolio risk measures.

## 2. Literature Review

Agarwal, Mullally, and Naik, (2015) [5] pointed out that assets risk and returns are inseparable when trying to make investment decisions. Returns are calculated as the amount of money a price holder gets when the prices go either up or down. Erel, Myers, and Read Jr, (2015) [6] denoted the computation of total share price returns in securities trading environment as

$$r_t = \frac{(p_1 - p_0)}{p_0} \quad (1)$$

Where  $p_1$  is the ending price of the share price,  $p_0$  is the initial price of the share price.

Erel, Myers, & Read Jr, (2015) [7] further noted that, for the price holder to earn profit the ending prices must go higher than the initial price. On the other hand, loss occurs when the ending prices are lower than the initial price. Losses in any business environment imply a risk and this is what any investor would want to minimize. Barnes, (2016) [8], is of the opinion that it is beyond human cognition to predict future movement of share prices hence prompting researcher to invent ways of solving this problem. Financial risk analysts and researchers have been digging deeper into strategies that would minimize losses incurred by investors. One of the ways is to estimate the amount of money an investor could lose based on historical data. With a collection of historical data of the prices, financial analysts have recommended fitting of statistical distribution functions to historical data. The distribution functions would then form the basis of predicting statistical estimates such as the quartile that would help in quantifying risk measures. Parametric methods that assume the distribution of historical data for share prices have been recommended in the estimation of risk measures.

Copula modeling has become popular in the modeling of the dependence structure of financial time series data. Omari, Mwita, and Gichuhi, (2018) [9], did a study to measure the risk of portfolio consisting of currency exchange rates. The authors noted that volatility clustering is best modeled using a GARCH model. A GARCH model is able to model the variance of the innovation of a return series as a function of previously time period error.

Engle & Manganelli, (2004) [10] noted that a distribution of highly frequently sampled data leads to a fat tailed distribution, which is different from the normal distribution.

They further pointed out that a method that does not make assumptions on the tails of a distribution works well in assessing the dependence structure of a time series. The model consists of order  $p$  for the GARCH terms, order  $q$  for the ARCH terms is denoted as follows.

$$\delta_t^2 = \omega + \alpha_1 \varepsilon_{t-1}^2 + \dots + \alpha_{t-n} \varepsilon_{t-n}^2 + \beta_1 \delta_{t-1}^2 + \dots + \beta_n \delta_{t-p}^2 \quad (2)$$

This can be summarized as follows;

$$\delta_t^2 = \omega + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^p \beta_j \delta_{t-j}^2 \quad (3)$$

Where the  $\sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2$  is the ARCH component of order  $q$  and  $\sum_{j=1}^p \beta_j \delta_{t-j}^2$  is the GARCH component of order  $p$ .

### 3. Research Methodology

#### 3.1. Introduction

The data set was sourced from the Nairobi Securities Exchange inform of share prices of Companies trading in the Market. The period was between the year 2012 and 2018. An optimal portfolio was selected from the companies using the Capital Assets and Pricing Model. It is one of the best models for dealing with unsystematic risk by assigning different weights to the Companies share prices.

#### 3.2. Portfolio Return Series

The portfolio returns time series was given by;

$$r_t = \mu_t + \varepsilon_t \quad (4)$$

Where  $r_t$  is the portfolio return series at time  $t$ ,  $\mu_t$  is the Mean equation of the portfolio return series and  $\varepsilon_t$  is the innovation/Error term of the time series model.

On getting the Portfolio return series, the study worked towards understanding the distribution of the portfolio returns series. A histogram was used to check whether the distribution is normally distributed or not. Additionally, summary statistics of the portfolio returns series were generated to give the values of skewness and kurtosis. The descriptive statistics helped a lot in determining the nature of the tails of the portfolio returns.

#### 3.3. Fitting the GARCH Model

The portfolio return series data was subjected to the Ljung Box test for determining the best autoregressive model to fit on the data. Ljung box test helps in determining the autocorrelation present in the time series data. A GARCH model was fit after determining the order of the ARCH and the GARCH terms in the overall model.

From the equation, the  $GARCH(p, q)$  model will be of the form;

$$\varepsilon_t = \delta_t * \varepsilon_t \quad (5)$$

$$\delta_t^2 = \beta_0 + \beta_1 \delta_{t-1}^2 + \dots + \beta_n \delta_{t-n}^2 + \alpha_1 \varepsilon_{t-1}^2 + \dots + \alpha_{t-n} \varepsilon_{t-n}^2 \quad (6)$$

This can be summarized as follows;

$$\delta_t^2 = \beta_0 + \sum_{i=1}^p \beta_i \delta_{t-i}^2 + \sum_{j=1}^q \alpha_j \varepsilon_{t-j}^2 \quad (7)$$

Where the  $\sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2$  is the ARCH component of order  $q$  and  $\sum_{j=1}^p \beta_j \delta_{t-j}^2$  is the GARCH component of order  $p$ .

$\delta_t^2$  is the variance of the innovation of the return series which is very sensitivity due the squared property. This property makes the series to overshoot depending on the movement of the time series data. This property enables the GARCH model to describe the volatility clustering experienced in the highly sampled data.

With the GARCH model in place, the study worked towards estimating the lower tails of the distribution of the portfolio return series. In Copula-based modeling the first step involves estimation of the marginal distribution of the innovations of the return's series. According to Sklar theorem there exists a function  $C$  that separates the joint distribution into univariate marginal.

The study applied the suggestion of (Kole, Koedijk, & Verbeek, 2007) [11] in determining the best type of copula to fit on the series. In this case, Kolmogrov and Anderson Darling test was used to fit the copula in the data.

#### 3.4. Estimation Procedure of the Copula

The estimation involved the usage of the maximum likelihood method to estimate the marginal and the copula. The fit of the copula was fitted using the distance measure between the mean and maximum absolute error and the empirical/hypothesized copula. According to (Kole, Koedijk, & Verbeek, 2007) [12], the Anderson-Darling test is a scaled version of the Kolmogrov Simonov test that works towards giving the weights to the tails thus enabling the capturing of the tail dependence of the data.

The next step involved simulating the return series shocks where a test was done to test whether the distance measure provide evidence for rejecting the fitted copula. The hypothesis involved construction of the distribution of distances where a sample size  $T$  will be used. The copula parameters were calculated for the new estimate of  $\hat{\theta}$  with distribution estimates. In testing steps, the criterion for selecting the lower tail distribution threshold was based on the P-Value.

#### 3.5. Goodness of Fit of the Model

Different families of copula were used to estimate the copula function from the portfolio return series. The families included Frechet Family, Gaussian distribution, Students  $t$  distribution and the Archimedean Copula Family. All the fitted models were subjected a model selection criterion for determining the best model. The Akaike's Information Criteria was used. The model with the lowest value of the AIC was selected as the best fit for the portfolio returns.

#### 3.6. Estimation of the Risk Measures

Risk measures of this study involved the estimation of the Value-at-Risk at 95%. The model with the best goodness of fit was used to calculate the VaR based on training data set.

### 3.7. Backtesting the VaR

A backtest analysis was done on the test data to establish the rate at which the model correctly predicts the Value-at-Risk.

## 4. Results and Discussion

### 4.1. Introduction

Secondary data was sourced from the Nairobi Securities Exchange data center for a period of six years. Out of the 63 Companies listed in the NSE, 31 Companies were considered in this particular study because 32 companies had insufficient trading data. The NSE as mandated by the Capital Markets Authority (CMA) has the right to terminate trading activities for inactive Companies.

The collected data was the closing share prices  $P_t$  for the considered Companies. These prices formed the basis of computing the log daily returns.

### 4.2. Optimal Portfolio Selection

The Capital Assets Pricing Model (CAPM) was used to select an optimal portfolio from the selected Companies. The CAPM enabled the study to select Companies that were from different industries, thus minimizing unsystematic risks as shown in Table 1.

Table 1. List of optimal portfolio companies.

Trading Symbol	Company Name	Industry
ARM	ARM Cement	Industrials
BERG	Crown Paints Kenya	Basic Materials
KUKZ	Kakuzi	Consumer Goods
KQ	Kenya Airways	Consumer Services
BBK	Barclays Bank of Kenya	Financials
KENO	KenolKobil	Oil & Gas

### 4.3. Portfolio Returns

The portfolio returns were split into training data and test data. Training data (in sample data) consisted of 1326 returns, which was used to train the Copula-based model. The test data consisted of 332 returns, which were used to test the predictability of the built copula model. Descriptive statistics analysis was done on the sample data and reported in table 2.

Table 2. Summary statistics of portfolio returns.

Statistics	Value
Observations	1326
Mean	2.749
Standard Deviation	1.524
Kurtosis	8.348
Skewness	-0.457
Jarque Bera	106110
P-Value	$2.2 \times 10^{-16}$

The percentage mean portfolio returns was 2.749 while the standard deviation of the portfolio returns was 1.524. The portfolio returns show a negative skewness of -0.457 and excess kurtosis of 8.348. This is in line with the characteristics of a heavy tailed distribution of high sampled data.

The portfolio returns were then subjected to a normality test. Jarque Bera test was carried out to assess whether the portfolio returns data conformed to a normal distribution. A probability value of  $2.2 \times 10^{-16}$  was reported. This value is less than 0.05 level of significance resulting to the rejection of the null hypothesis that the returns conform to a normal distribution. The conclusion was that the data of the portfolio returns was non-normal hence; normality assumption could not be used to analyze the data (Koizumi, Okamoto, & Seo, 2009) [13].

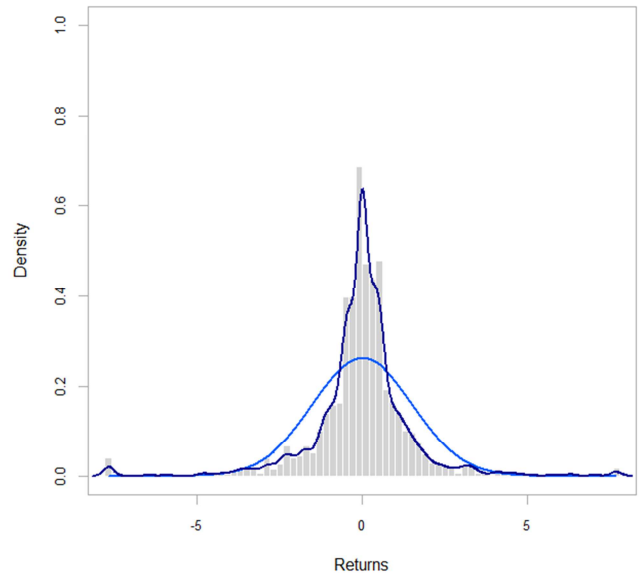


Figure 1. Portfolio Returns Distribution.

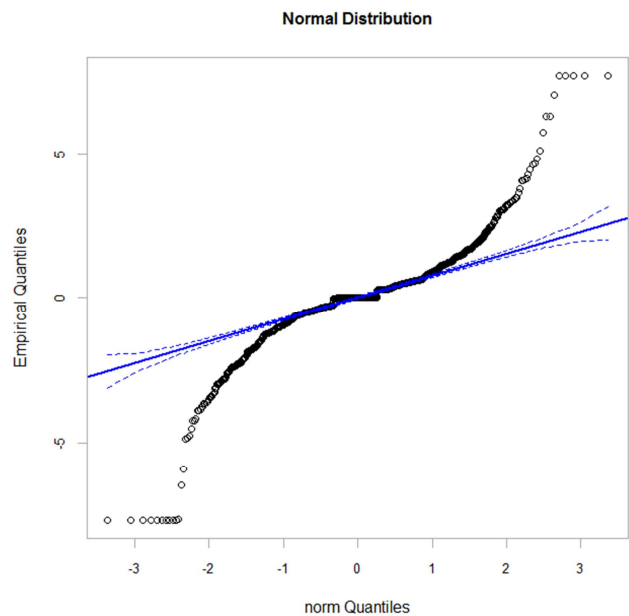
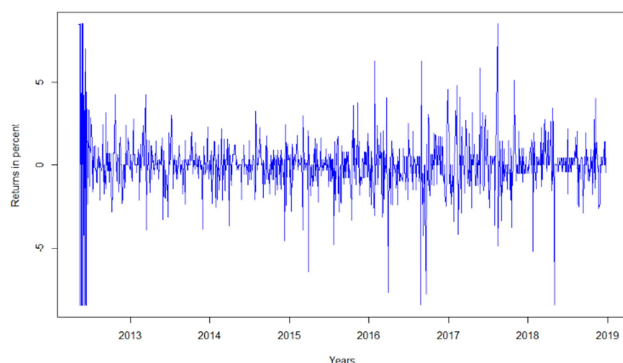


Figure 2. Portfolio returns QQ Plot.

### 4.4. GARCH Model

The log portfolio returns were plotted against the years as shown in figure 3.



**Figure 3.** Plot of portfolio returns against years.

The characteristics observed from the plot above indicated presence of heteroscedasticity. The observations from the plots are further backed by the results from the Box-Ljung test.

**Table 3.** ARCH Effects – Box Ljung Test.

X-Squared	DF	P-value
592.83	12	$2.2 * 10^{-16}$

A test for the ARCH effects was done to check whether the observations from the plot were true. Table 3 reports the Ljung Box test results. A P-value of  $2.2 * 10^{-16}$  was reported which led to the rejection of the null hypothesis that there are no ARCH effects. The rejection of the null hypothesis was due to the fact that the P-value is less than 0.05 level of significance (Aloui, Aïssa, & Nguyen, 2013) [14]. The GARCH model was the best model for analyzing time series data characterized by conditional heteroscedasticity (Omari, Mwita, & Gichuhi, 2018) [15].

**Table 4.** GARCH(1,1) parameter estimates.

Parameter	Estimate	Std. Error	t Value	P-Value
Mu	-0.0139	0.0357	-0.389	0.697
Omega	0.2852	0.0471	6.055	0.000
Alpha 1	0.4127	0.0734	5.6194	0.000
Beta 1	0.5385	0.0547	9.83476	0.000

Table 4 above shows that the parameters of GARCH(1,1) are significant at 0.05 level of significance hence the model can be used to describe conditional heteroscedasticity.

**Table 7.** Kupiec Unconditional Test.

	Tawn Type 1	Gaussian	t – Student	Frank	Gumbel	Clayton
Coverage	0.9337	0.7732	0.770	0.7756	0.780	0.801
Test Statistics	1.685	327.795	334.4703	382.15	114.18	91.56
P-Value	0.1942	0.001	0.0001	0.001	0.001	0.001
Decision	Fail to Reject H0	Reject H0	Reject H0	Reject H0	Reject H0	Reject H0

Table 7 above shows the backtesting result of the Value-at-Risk for the different copula-based models at 95% confidence level. The Tawn type 1 Copula model had the

#### 4.5. Copula Models Estimation

The dependence structure of the GARCH model was fitted using the Copulas. Different types of Copula-based models were carried out of the sigma and the residuals series of the GARCH.

**Table 5.** Estimated Parameters for the different copulas.

Copula	Estimated Parameters	Test Statistics	P-Value
Gaussian	0.0501	0.2632	0.100
t Student	0.0537	0.910	0.869
Frank	0.3274	0.9104	0.167
Gumbel	1.0814	0.3587	0.018
Clayton	0.0730	1.924	0.042
Tawn Type 1	0.3422	1.959	0.027

Table 5 above shows the different copula models that were fitted to the dependence structure of the portfolio returns. The table also includes the estimated parameters for the different copulas, test statistics and the probability values.

#### 4.6. Goodness of Fit

All the models were subjected to a goodness of fit test.

Tawn type 1 model had the lowest value of the AIC (Akaike Information Criteria) and was therefore selected as the model with the best goodness of fit compared to the copula-based models.

**Table 6.** Tawn Type 1 copula AIC value.

LogLik	AIC	BIC
40.58	-77.16	-66.78

This finding leads to the key finding in this research where they Tawn type I copula was the best copula-based model to describe the dependency structure of the GARCH model.

#### 4.7. Estimation of the GARCH (1, 1) – Copula Based VaR

The next steps involved computing the copula based VaR using the estimated GARCH model at 95% confidence level. At 95% confident level, an investor was likely to lose a maximum of 2.306% of the total investment on this portfolio.

#### 4.8. Back Testing the VaR

As a measure of performance, coverage ratios were used to compare models. The results for the unconditional coverage test (Kupiec test) were as follows;

highest unconditional coverage as compared to the other copula-based models. The decision attributes indicated the presence of correct exceedances when the Tawn type 1

copula model was checked against the test data.

The results for the conditional coverage test

(Christoffersen test) were as follows;

**Table 8.** Christoffersen Conditional Test.

	<b>Tawn Type 1</b>	<b>Gaussian</b>	<b>t – Student</b>	<b>Frank</b>	<b>Gumbel</b>	<b>Clayton</b>
Test Statistics	1.8917	327.877	335.6307	382.564	117.559	91.73
P-Value	0.3883	0.001	0.001	0.0015	0.001	0.042
Decision	Fail to Reject H0	Reject H0	Reject H0	Reject H0	Reject H0	Reject H0

Table 8 shows the results for the conditional back testing of the 95% VaR where the Tawn type 1 Copula also showed the correct exceedances when checked against the test data.

It was therefore correct to state that the Tawn Type 1 based Copula model was the best model for describing the dependency structure of the portfolio returns volatility.

## 5. Conclusion and Recommendations

### 5.1. Conclusion

Being aware of the risk exposure is crucial in any investment that involves uncertainties. This study showed that copula-based models could be used to estimate VaR for an optimal portfolio of commodities. The parametric model was made by fitting suitable time-series models for the margins and copulas for the dependence. Some of the most common time series models have been used to order the margins, while some of the most common copulas have been used to model the dependence. The explication of the result is believed to mainly depend on the modeling. There were many types of copula-based models and the best type of Copula that would fit this type of data is Tawn Type 1 copula. This was justified using the goodness of fit and the backtesting results.

### 5.2. Recommendations

More care be given to modelling the margins and perhaps include other models or allow re-estimation of models during the predictions. Also, other copula-based models can be included and used in order to better estimate the dependence.

### 5.3. Further Research

The idea of creating new models that would offer solutions to practical business problems such as the use of transitional matrix for the estimating risk measures for treasury bonds.

## References

- [1] Agarwal, V., Mullally, K. A., & Naik, N. Y. (2015). The economics and finance of hedge funds: A review of the academic literature. *Foundations and Trends® in Finance*. 10 (1), 1-111.
- [2] Aloui, R., Aïssa, M. S., & Nguyen, D. K. (2013). Conditional dependence structure between oil prices and exchange rates: a copula-GARCH approach. *Journal of International Money and Finance*, 32, 719-738.
- [3] Barnes, P. (2016). *Stock market efficiency, insider dealing and market abuse*. Gower.
- [4] Engle, R. F., & Manganelli, S. (2004). CAViaR: Conditional Autoregressive Value at Risk by Regression Quantiles. *Journal of Business & Economic Statistics*, 22, 367-381.
- [5] Erel, I., Myers, S. C., & Read Jr, J. A. (2015). A theory of risk capital. *Journal of Financial Economics*, 118 (3), 620-635.
- [6] Evans, W. B. (2016). *Effects Of Political Risk And Macroeconomic Factors On Stock Market Returns At Nairobi Securities Exchange*. Kenyatta University.
- [7] Fama, E. F., & French, K. R. (2017). Long-horizon returns. *The Review of Asset Pricing Studies*.
- [8] Genest, C., Rémillard, B., & Beaudoin, D. (2009). GGoodness-of-fit tests for copulas: A review and a power study. *Insurance: Mathematics and economics*, 44 (2), 199-213.
- [9] Koizumi, K., Okamoto, N., & Seo, T. (2009). On Jarque-Bera tests for assessing multivariate normality. *Journal of Statistics: Advances in Theory and Applications*, 1 (2), 207-220.
- [10] Kole, E., Koedijk, K., & Verbeek, M. (2007). Selecting copulas for risk management. *Journal of Banking & Finance*, 31 (8), 2405-2423.
- [11] Omari, C. O., Mwita, P. N., & Gichuhi, A. W. (2018). Currency Portfolio Risk Measurement with Generalized Autoregressive Conditional Heteroscedastic-Extreme Value Theory-Copula Model. *Journal of Mathematical Finance*.
- [12] Pries, H. (2016). *Market risk calculations in stock-and bond prices: a garch-copula approach*.
- [13] Rao, C. a. (2012). *Handbook of statistics*. Elsevier Science & Technology.
- [14] Wali, B., Greene, D. L., Khattak, A. J., & Liu, J. (2018). Analyzing within garage fuel economy gaps to support vehicle purchasing decisions—A copula-based modeling & forecasting approach. *Transportation Research Part D: Transport and Environment*, 63-186.
- [15] Ye, W., Luo, K., & Liu, X. (2017). Time-varying quantile association regression model with applications to financial contagion and VaR. *European Journal of Operational Research*, 256 (3), 1015-1028.