ANKARA YILDIRIM BEYAZIT UNIVERSITY

AYBU Business Journal, 2(1), 35-45
https://dergipark.org.tr/tr/pub/abj
https://aybu.edu.tr/isletme/tr/sayfa/5218/AYBU-BUSINESS-JOURNAL

AYBU Business Journal

Submitted: 15 May 2022
Accepted: 27 June 2022
Published Online: 30 June 2022

RESEARCH ARTICLE

Modeling The Dependence Structure of Financial Data with a Copula: Electricity Index – An Example of The Dollar Exchange Rate

Tuğba Ünal[a*]

[a]*Institute of Science, Department of Statistics, Ankara University, Ankara, Turkey*

Abstract

The Copula method is used to reveal the dependency structure between random variables. In both parametric and non-parametric cases, the measurement of dependency with copula functions, alternatives to many methods, and methods that allow much simpler calculation of these calculations have been proposed. In this study, the dependency structure between the electricity index and the dollar rate was examined using the copula function. The relationship between the two indices was compared with MSE, AIC and BIC values. As a result of these calculations, it was determined that the most suitable modeling according to MSE could be done with Clayton, and the most suitable modeling according to AIC and BIC could be done with Gumbel.

Keyword

Copula, dependency structure,index, Electricity Index

1. INTRODUCTION

Copulas were proposed by Abe Sklar in 1959. The history of the copula theory dates back to the work of Hoeffding in the 1940s. Copulas reveal the dependency structure between random variables. Copulas are functions whose univariate marginals are uniformly distributed within the range [0,1] and relate multivariate distributions to their univariate marginals. Capulas are a powerful method used to model the distributions of variables with a common marginal distribution. The main purpose of the capula functions is to obtain the most appropriate multivariate distribution for the studied data by revealing the dependency structure. In the first stage of this study, the definitions of Copulas, their basic features, Archimedean Copula families and the characteristics of these Copula families are explained. In the next section, the estimates of the copulas explaining the dependency structure between the BIST electricity index and the dollar rate between 01.10.2015 and 01.10.2019 and the results obtained from the applications are included. "Copula" is a Latin word and was first used by Abe Sklar in 1959. As a word meaning, copula means to relate (associate), commitment. Copulas, in a broader sense, are the explanation of multivariate distribution functions with the help of univariate marginals. Copulas do this by using a common marginal distribution function to model the distribution of

[*] Corresponding author.

Contact: Tuğba Ünal tubaunal1991@gmail.com

To cite this article: Ünal T. (2022). Modeling The Dependence Structure of Financial Data with a Copula: Electricity Index – An Example of The Dollar Exchange Rate. *AYBU Business Journal*, 2(1), 35-45.

variables. The basis of copulas is that they can be used in non-parametric dependency measures, that is, the assumption of normality is not sought and a large number of marginal distributions are determined. The main purpose of the capula functions is to reveal the multivariate distribution, which is suitable for the observed data, with the dependency structure. That is, it expresses the dependency structure between random variables. Copulas are widely seen in finance, economics, actuarial, financial time series, survival analysis and medical statistics, and its popularity is increasing day by day.

2. LITERATURE REVIEW

According to Hafner and Reznikova (2010), in their study on the portfolio they created from stocks based on the 30 Dow Jones index, showed that the stock model effectively measures the asymmetric correlation and is at least as successful as the parametric methods in creating a portfolio. Paletier (2006) used the dynamic conditional correlation method as Engle and Sheppard (2001) applied to the data set in the regime variable structure, and the model showed successful performance especially in detecting sudden and high volatility. Cuaresma and Wojcik (2006) used dynamic conditional correlation, which takes into account the joint correlation of both variables, to test the effect of interest rates on the monetary policies of the Czech Republic, Poland and Hungary. Bautista (2006) examined the change in time between exchange rates and interest rates. In addition, the performance of the model, which takes into account the varying correlation, was tested with data from six different Asian countries. Wang and Thi (2007) used a conditional correlation model to dynamically detect the time-varying correlation between Taiwan and USA stock markets. Lanza et al. (2006) used a conditional correlation model to model the volatility of prices in oil markets. Lee (2006) used the dynamic conditional GARCH model to determine the relationship between inflation and production before the Second World War. The assumption of the time-varying correlation between the variables was established as the basis for the studies on detection. Sklar (1959) used the term copula to indicate the dependence between variables. It is based on a common marginal distribution of variables expressed by the copula method. Copula in the literature is a very powerful tool for modeling. The most basic feature underlying the copula function is related to the common distribution among the variables. The Copula method is very useful for determining the distribution of the marginal number for the assumption of normality and the n-dimensional joint distribution. In this way, the marginal distributions of a copula function can be multidimensional. It allows detecting the varying co-correlation volatility by combining it with the distribution over time and calculating the dependence between the variables. Due to its listed features, the copula method is used in modeling the portfolio distribution. stated that it can be used. In this study, MSE, AIC and BIC values were calculated to determine the most suitable copula family in modeling the dependence between the BIST electricity index and the dollar rate. As a result of these calculations, it was determined that the most suitable modeling was with Clayton according to MSE, and with Gumbel according to AIC and BIC.

3. SKLAR THEOREM

The theorem at the beginning of this chapter is central to the theory of the copula. The theorem that relates the multivariate distribution functions and their one-variable marginal distribution functions is called the sclar theorem. Let $H(x, y)$ be a common distribution function with marginals $F(x)$ and $G(y)$. For each x,y ϵ $\overline{\mathrm{R}}$

$$H(x, y) = C(F(x), G(y))$$

It has a C copula. This equation can be converted to an equation expressing the copula using the inverse of the common distribution function and marginal distribution functions. But if the marginal distribution function is not strictly increasing, the function cannot be inverted. In this case, the quasi-inverse concept will be examined.

3.1. Semi-Inverse Function

Let F be a distribution function.

i) If t F is member of the value set of F, then $F_{(t)}^{-1}$ is equal to any value of x in $\overline{\mathrm{R}}$ such that $F(x) = t$

$F(F_{(t)}^{-1}) = t$

ii)If t is not in the value set of F

$F_{(t)}^{-1} = ebas\{x \backslash F(x) \geq t\} = ek\ddot{u}s\{x \backslash F(x) \leq t\}$

If F is a strictly increasing function, then F has only one known semi-inverse. This is shown as F^{-1}, which is known as the inverse of F (Nelsen, 1999).

Common distribution function $H(x, y)$ with marginals $F(x)$ and $G(y)$ and C

Let $H(x, y) = C(F(x), G(y))$ be a copula function. Let F^{-1} and G^{-1} F ve G be semi-inverses of F and G, respectively. Then for any (u, v) in the domain of C'

$C'_{(u,v)} = H(F_{(u)}^{-1}, G_{(v)}^{-1})$

3.2. Frechet - Hoeffding Limints

Suppose we have any variable m with its univariate marginals. Let the distribution function $F(y_1, y_2, \dots, y_m)$ and its marginals be in the range $F_1, F_2, \dots, F_m$ [0,1].

If we denote the Frechet-Hoeffding lower and upper bounds with F_u and F_v

$F_u(y_1, y_2, \dots, y_m) = max\,[\sum F_{j-m+1,}\, 0] = W$

$F_v(y_1, y_2, \dots, y_m) = min\,[F_1, F_2, \dots, F_m] = M$

$W = max[\sum F_{j-m+1,}\, 0] \leq F(y_1, y_2, \dots, y_m) \leq min[F_1, F_2, \dots, F_m] = M$

In the case of univariate marginals, in the term Frechet - Hoeffding class, m denotes the class of variable distributions$F(F_1, F_2, \dots, F_m)$ marginals are constant and given. (F_{12}, F_{13}) refers to the classes $F(F_{12}, F_{13}, F_{23})$ when it is bivariate or given in higher dimension.
More precisely, the Frechet-Hoeffding lower bound copula is smaller than any copula, and the Frechet-Hoeffding upper bound copula is greater than any copula.

3.3. Survival (Life) Copula

Whether the random variables of interest represent the lifetimes of individuals or objects in a heap, the probability that an individual or object will live longer than x time is called the survival function or life function (or reliability function).

$\bar{F}(x) = P[X > x] = 1 - F(x)$

form is also displayed. The value range is $[0, \infty)$. Survival function for pair of random variables (X, Y) with common distribution function H

$\bar{H}(x, y) = P[X > x, Y > y]$ is expressed in the form.

$\bar{H}(x, y) = 1 - F(x) - G(y) + H(x, y)$

$= \bar{F}(x) + \bar{G}(y) - 1 + C\big(F(x), G(y)\big)$
$= \bar{F}(x) + \bar{G}(y) - 1 + C(1 - \bar{F}(x), 1 - \bar{G}(y))$

It is written in the form. $\hat{C}: I^2 \to I$

$\hat{C}(u,v) = u + v - 1 + C(1-u, 1-v)$ survival function
$\bar{H}(x,y) = \hat{C}(\bar{F}(x), \bar{G}(y))$
The form is also obtained. $\hat{C}$ is the survival copula X and Y

4. THE PROPERTY OF SYMMETRY IN THE COPULA

The joint distribution function of X and Y is H, the marginal distribution functions are F and Let G be continuous random variables with capula C. If C(u,v)=C(v,u) for all (u,v) pairs in I^2, then the random variables X and Y are interchangeable, and C is simply called symmetric (Nelsen, 2006)

4.1. Archimedean Copulas

It was first put forward by Kimberling. Archimedes' copulas are very preferred as they provide us mathematical convenience and flexibility and are used to model the insufficiency structure between two variables.
This approach allows a multivariate capulan to be reduced to a simple univariate function. $C: [0,1]^2 \to [0,1]$ let's consider a bivariate capula.
$\varphi: [0,1] \to [0,\infty)$ let be a continuous, decreasing and convex function such that
$\varphi(1) = 0$
$u, v \in [0,1]$ for $C(u,v) = \varphi^{-1}\big(\varphi(u) + \varphi(v)\big)$

Your format can also be expressed. φ is called the generator of the gate C. For this, let's first define the function $\varphi^{(-1)}$

$\varphi: [0,1] \to [0,\infty)$ ‘let a be a continuous, strictly decreasing function such that
$\varphi(1) = 0$. The function $\varphi^{(-1)}$,defined as the set of definitions of the inverse function of φ $[0,\infty)$ and the set of values [0,1] is as follows,

$$\varphi^{(-1)} = \begin{cases} \varphi^{(-1)}, & 0 \le t \le \varphi(0) \\ 0 \quad , & \varphi(0) \le t < \infty \end{cases}$$

the function $\varphi^{(-1)}$ and not increasing on $[0,\infty)$ and $[0,\varphi(0)]$ is the exact reduction on Moreover, in the range of [0,1], $\varphi^{(-1)}(\varphi(u)) = u$, and

$$\varphi\big(\varphi^{(-1)}(t)\big) = \{\begin{matrix} t, & 0 \le t \le \varphi(0), \\ \varphi(0), & \varphi(0) \le t < \infty, \end{matrix}$$

$$= min(t, \varphi(0))$$

Finally, if $\varphi(0) = \infty$, then $\varphi^{(-1)} = \varphi^{-1}$. (Nelsen, 2006)

Let C be an archimedean copula whose manufacturer is φ. The properties of archimedean copulas are given below.There:
i)Let $C(u,v) = C(v,u)$ be symmetric if and only if
ii) If u,v,w ∈ I, then $C(C(u,v),w) = C\big(u, C(v,w)\big)$ such that
iii) any constant c>0 is the generator of C' at c φ.

4.2. Some Archimedean Copulas

Ali-Mikhail-Haq Copula Family

$$C_\theta(u,v) = \frac{uv}{1-\theta\,(1-u)(1-v)}\ , \theta\ \epsilon\ [-1,1)$$

Where $\theta = 0$ and $C_\theta(u,v) = uv = \prod(u,v)$ forms the independence copula. This kapula also belongs to an Archimedean copula family.

It is the copula of the joint distribution function, the marginals of which have a standard logistic distribution, $\theta\ \epsilon\ [-1,1)$. As is known as the Ali-Mikhail-Haq copula, it is also known as the copula of the Gumbel bivariate logistics distribution.

Clayton Copula Family

The Clayton capula family is an asymmetric archimedean capula. The Clayton kapula is defined as follows:

$$C_\theta(u,v) = max([u^{-\theta} + v^{-\theta} - 1])^{-1/\theta}, 0)\ , \theta\ \epsilon\ [-1,\infty)\backslash\{0\}$$

Manufacturer function,

$$\varphi = \frac{1}{\theta}(t^{-\theta} - 1)$$

Where $\theta \to 0$, the $\lim_{\theta\to 0} C_\theta(u,v) = uv = \prod(u,v)$ forms the independence capula. It implies perfect dependence, while $\theta \to \infty$ (Nelsen, 1999).

Frank Copula Family

The Frank capula family is a symmetric Archimedean kapula. The Frank capula is defined as follows:

$$C_\theta(u,v) = -\frac{1}{\theta} In[1 + \frac{(e^{-\theta u} - 1)(e^{-\theta v} - 1)}{e^{-\theta} - 1}], \theta\ \epsilon\ (-\infty,\infty)\backslash\{0\}$$

Manufacturer function,

$$\varphi(t) = -ln\frac{e^{-\theta t} - 1}{e^{-\theta} - 1}$$

Where $\theta \to 0$ iken, $\lim_{\theta\to 0} C_\theta(u,v) = uv = \prod(u,v)$ forms the independence copula. When $\theta \to \infty$, Frenchet-Hoeffding becomes equal to its upper limits.
The Frank capula family is preferred more than other capulas in practice because it allows negative dependence among marginals, unlike other capulas (Nelsen, 1999).

Gumbel Hougaard Copula Family

The Gumbel capula family is an asymmetric Archimedean capula. The Gumbel hood is defined as follows:

$$C_\theta(u,v) = \exp\{-[(-lnu)^\theta + (-lnv)^\theta]^{1/\theta}\}$$

Manufacturer function,

$$\varphi(t) = (-lnt)^\theta$$

with an increase in the parameter θ, the deciency between observations increases.
When $\theta \to 1^{+}$, that is, the $\lim_{n\to\infty} C_\theta\,(u,v) = uv = \prod(u,v)$ becomes equal to the independence capula. While $\theta \to \infty$, it shows perfect dependence (Joe,1997).

5. COPULA ESTIMATION METHODS

The problem of statistical modeling of kapulas is also studied in two stages:
i) determination of marginal distributions
ii) identification of the appropriate copula function (Cherubini et al., 2004).

The Exact Method of Maximum Likelihood (MLE) ve The Method of Inference on Marginals (IFM).

5.1. The Exact Maximum Likelihood Method (MLE)

Multidimensional density function,

$$f(x_1, x_2, \dots, x_n) = c(F_1(x_1), F_2(x_2), \dots, F_n(x_n) \prod_{j=1}^{n} f_j\left(x_j\right)$$

is also expressed.

Multidimensional capula function,

$$C(F_1(x_1), F_2(x_2), \dots, F_n(x_n)) = \frac{\partial^n C(F1^{(x1)}\ , F2^{(x2)}, \dots\ , Fn^{(xn)})\ )}{\partial F1^{(x1)}\,.\partial F2^{(x2)} \dots \partial Fn^{(xn)}}$$

The order partial derivative, the c-capula density, and the fi's are univariate marginal probability density functions. We can trace the solutions of the statistical model problem of capulas in two steps as mentioned below:
i) determination of marginal distributions
ii) identification of the appropriate copula function.

let be our sample data matrix. From here, the log-likelihood function

$$l\ (\theta) = \sum_{t=1}^{T} ln\ c\ (F_1\ (x_{1t}), F_2(x_{2t}), \dots, F_n(x_{nt}) + \sum_{t=1}^{T} \sum_{j=1}^{n} In\ fj^{(x_{jt})}$$

It can be expressed as. Where θ is the set of all parameters related to the capula and the marginals. If the set of marginal probability density functions and the capula are known, it can be written to the log-likelihood function mentioned above, which maximizes this equality $\hat{\theta}_{MLE} = max\ l\ (\theta)$ the maximum likelihood estimator can be obtained. (Cherubini et al., 2004)

5.2. The Inference Method for Marginals (IFM)

Since it is not always possible to use the most likelihood estimation method or it requires very intensive calculations, the two-stage IFM estimation method is preferred in multivariate kapula estimation methods. These stages are:

Step 1: The parameters of the marginals were estimated by estimating the marginal distributions with a variable.

$$\theta_1 = ArgMax\theta_1 \sum_{t=1}^{T} \sum_{j=1}^{n} log\ fj^{(xjt\ ;\ \theta 1)}$$

is also obtained.

Step 2: The obtained θ1 estimate is substituted into the log-likelihood function and the capula parameter θ_2 is estimated.

$$\theta_2 = ArgMax\theta_2 \sum_{t=1}^{T} ln\ c\ (F_1\ (x_{1t}), F_2(x_{2t}), \dots, F_n(x_{nt}); \theta_2, \theta_1)$$

This method is called the Inference Method for Marginals (IFM).
The estimator,

$$\hat{\theta}_{IFM} = (\theta_1, \theta_2)'$$

is also expressed.

IFM estimator can be calculated more easily than MLE estimator. Therefore, it will be easier to compare the asymptotic covariance matrices of the IFM estimator compared to the MLE estimator.

It is proved that IFM estimators, like MLE estimators, which gave the IFM estimator for the first time (joe, 1997), have the property of asymptotic normality under favorable conditions.

$$\sqrt{T}(\theta_{IFM} - \theta_0) \rightarrow\ N(0, G^{-1}(\theta_0))$$

Where $G(\theta_0)$ is the Godambe cognitive Matrix. l Let the logability function . Here is the Skar fan

$$s(\theta) = (\frac{\partial l_1}{\partial\theta_{11}}, \frac{\partial\ l_2}{\partial\theta_{12}}, \dots, \frac{\partial l_n}{\partial\theta_{1n}}, \frac{\partial l_c}{\partial\theta_2})'$$

The Godambe information matrix is;

$$G(\theta_0) = D^{-1}V(D^{-1})'$$

Here,

$$D = E[\frac{\partial\ s(\theta)}{\partial\theta}]\ \text{ ve }\ V = E[s(\theta)s'(\theta)]$$

Many derivative calculations are required for the estimation of this covariance matrix. (Joe, 1997), It has been stated that IFM method is very effective in comparing with MLE method (Cherubini et al.,2004).

6. Application

In this study, between 1.09.2015 and 01.09.2019, the electricity index of the Istanbul Stock Exchange and the dollar exchange rate received from the Central Bank were used. Here, an analysis has been made using the returns of the data. The return is calculated by the formula:

$$ln\left(\frac{P_t}{P_{t-1}}\right)$$

The following 3 methods were used in the selections of models for copulas:

- MSE (Mean Error Squared)
- AIC (Akaiki Information Criterion)
- BIC (Bayesian Information Criterion)

Mean Error Squared Criterion (MSE)

$$MSE = \frac{1}{N}\sum(F_{ei} - F_i)^2$$

$F_{ei}: Empirical\ \ Probability$

$F_e: Theoretical\ Probability$

$N: Number\ of\ \ Observations$

Akaiki Information Criterion (AIC)

The Akaiki Information Criterion is one of the measures of goodness of fit that is often used in practice.

$$\mathrm{AIC} = 2\mathrm{k} - 2\ln\,(\mathrm{L})$$

- k is the number of parameters in the model,

- ln L the log likelihood value of the model

shows.

When choosing a model, the smallest AIC value is preferred.

Bayesian Information Criterion (BIC)

Bayesian Information Criterion is another measure of goodness of fit that is mostly used in practice.

$$BIC = k\ln(N) - 2\ln\,(L)$$

• k is the number of parameters in the model
• ln(L), the likelihood value of the model
• N is the sample size.
Descriptive statistics are given in Table 1.

Table 1. Descriptive statistics are given in Table 1.

DESCRIPTIVE STATISTICS

	N	Minimum	Maximum	Mean	Std. Deviation	Skewness		Kurtosis	
	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic	Std. Error
BIST_elektrik	996	-,097030	,070207	,00028637	,014312605	-,476	,077	3,762	,155
USD	996	-,064757	,147066	,00068965	,010509434	3,197	,077	47,399	,155
Valid N (listwise)	996								

The graphs of BIST Electricity Index and USD Logarithmic Difference Series are given below.

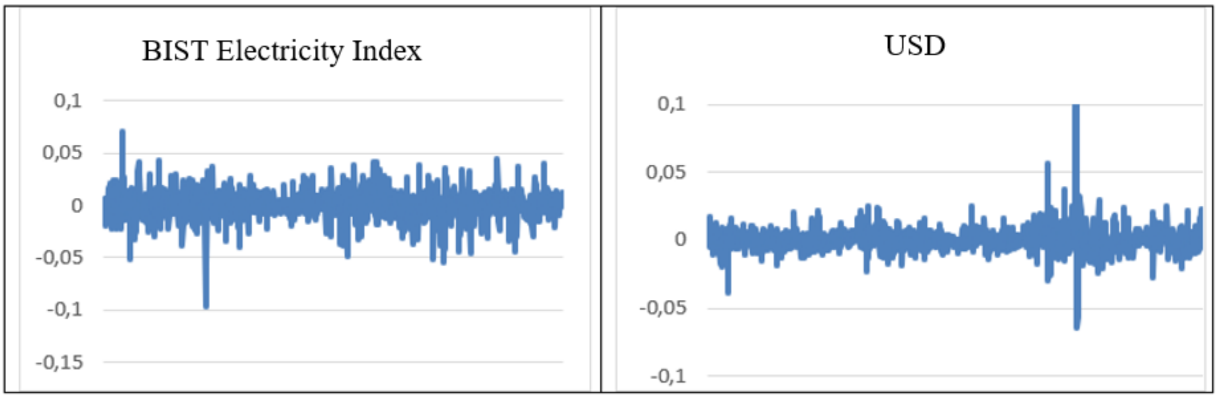

Figure 1. BIST Electricity Index and USD Logarithmic Difference Series

AIC, BIC and MSE values are given in Table 2. According to the results given in Table 2, it was determined that the most suitable modeling according to MSE could be done with Clayton, and the most suitable modeling according to AIC and BIC could be done with Gumbel.

Table 2. AIC, BIC and MSE values

	$\hat{\theta}$	LL	AIC	BIC	MSE
Frank	-0.2489	-0.8433	3.6866	8.5903	0.000062362
Clayton	-0.0110	-0.0695	2.1390	7.0427	**0.000013402**
Gumbel	0.9987	-0.0103	**2.0206**	**6.9243**	0.000159600

AIC, BIC and MSE values are given in Table 2. According to the results given in Table 2, it was determined that the most suitable modeling according to MSE could be done with Clayton, and the most suitable modeling according to AIC and BIC could be done with Gumbel. Frank, Clayton and Gumbel BIST Electricity Index and USD Distribution Diagram are given in Figure 2.

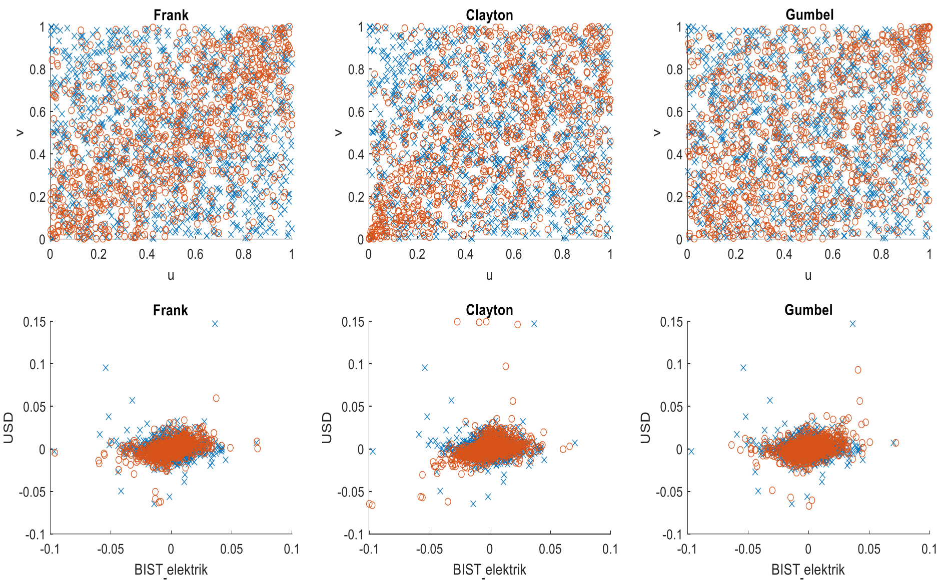

Figure 2. Frank, Clayton, and Gumbel BIST Electricity Index and USD Scatter Diagram

RESULT

In this study, the applications of the theory of copulas in finance are discussed. For this reason, Frank, Clayton, and Gumbel copulas used especially in finance have been examined. In addition, the properties of Archimedean copulas, which are the most important class of capsules, were examined and their advantages over the dependence measure were emphasized. In the application part of this study, to determine the most appropriate scapula family in modeling the dependence between the BIST electricity index and the dollar exchange rate, the following calculations were performed: MSE, AIC, and BIC. In the results of these calculations, it was determined that the most appropriate modeling according to MSE is with Clayton, and according to AIC and BIC, it is with Gumbel.

REFERENCES

Alhan, A. Independence "Copula involving Families, and copula dependence structure prediction methods on the Istanbul Stock Exchange industry" , PhD thesis, Gazi University Institute of Sciences, 2008, 1-162.Cox, D. R. and Johnson, M. E., "Analysis of Survival Data", Cahpman and Hall ,1984.

BIST electricity Rate History – Investing.com, Retrieved from https://tr.investing.com/indices/ise-electricity-historical-data, 2019.

Central Bank of the Republic of Turkey (MB), Retrieved from

https://evds2.tcmb.gov.tr/index.php?/evds/serieMarket/#collapse_2 ,2019

Cuaresma, J. C., & Wójcik, C. (2006). Measuring monetary independence: Evidence from a group of new EU member countries. *Journal of Comparative Economics*, *34*(1), 24-43.

Engle III, R. F., & Sheppard, K. (2001). Theoretical and empirical properties of dynamic conditional correlation multivariate GARCH.

Genest, C. And Rivest, L. P.," Statistical inference procedures for bivariate Archimedean copulas", J. Amer. Statist. Assoc., 1993.

Gumbel, E. J.," Bivariate Exponntial Distributions", J. Amer. Statist. Assoc., 55, pp. 698-707,1960.

Hafner, C. M., & Reznikova, O. (2010). Efficient estimation of a semiparametric dynamic copula model. *Computational Statistics & Data Analysis*, *54*(11), 2609-2627.

Joe, H., 1997. Multivariate Models and Dependence Concepts, Chapman and Hall, 1997 London.

Lee, J. (2006). The comovement between output and prices: Evidence from a dynamic conditional correlation GARCH model. *Economics Letters*, *91*(1), 110-116.

Nelsen, R., "An introduction to Copulas", Springer, 1998.

Sklar A.Fonction de re'partition a' n dimensions et leurs marges,8,Publications de L'Institute de Statistique, Universite' de Paris,Paris, 1959, 229–231.

Trivedi, PK. And Zimmer, D.M., "Copula Modelling: An Introduction for Practitioners", Now Publishers, 2007.

Wang, K. M., & Thi, T. B. N. (2007). Testing for contagion under asymmetric dynamics: Evidence from the stock markets between US and Taiwan. *Physica A: Statistical Mechanics and its Applications*, *376*, 422-432.