Chapter 2

Theoretical Foundation and Literature Review

2.1 Extreme Value Theory

Extreme value theory (EVT) has been one of the most significant statistical disciplines for the applied sciences for a long period of time. EVT is mainly used for modeling and analysis of the extreme values of the data of interest with a very low probability to happen (Alves and Neves, 2010). There are two approaches for finding the extremes of the data of interest.

1. Block Maxima (BM).
2. Peaks-Over Threshold (POT).

BM and POT are the statistical analyses of maxima or minima, and exceed over a upper or a lower threshold (Lai and Wu, 2007). Pokrivcak and Rajcaniova (2011) analyzed the statistical relationship between ethanol, gasoline, and crude oil prices by evaluating the relationship between the variables in the Impulse Response Function (IRF) and the Vector Autoregression (VAR). The result shows that oil prices have no co-integration between ethanol, and ethanol and gasoline, but that oil and gasoline prices have co-integration in the relationship. Price of oil has a shock effect on the price of gasoline.

This paper aims to study the tail dependence between the price of crude oil and the price of ethanol in the spot market by using bivariate extreme value copulas.

2.2 Generalized Extreme Value (GEV) Distribution

For a single margin, abstract is the maxima sequence, which is the same as defined before, and “i” is the number of blocks. F is the general price distribution, and G is the asymptotic extreme value distribution. The EVT shows that by founding a series of a_n and b_n, the maxima can be converted to be the general extreme value distribution (GEV) G (Coles, 2001; Beirlant, 2004):

$$G(x; b, a, \xi) = exp\left\{-\left[1 + \xi\left(\frac{x-b}{a}\right)\right]^{-1/\xi}\right\}. \quad (2.1)$$

Where ξ is the shape parameter explaining the behavior of the tail of the distribution. When $\xi < 0$ the distribution is the Weibull, $\xi > 0$ the Fréchet, and $\xi = 0$ the Gumbel.

2.3 Bivariate block maxima

The bivariate block maxima model could be investigated using the non-parametric and parametric approaches. We use the parametric approach in this study to investigate the tail dependence between the prices of crude oil and ethanol with the bivariate BM, which is provided below (Chuangchid et al, 2012):

Let (X, Y) be a bivariate random vector of the maxima which is an i.i.d. sequence for a certain time period. Let the distribution of (X, Y) be the bivariate extreme value distribution (BEVD) that is approximated with the cumulative distribution function (cdf) G. The BEVD is investigated by margins G_1 and G_2, which are necessarily EVD, by using its Pickands dependence function A (Rakonzai and Tajvidi, 2010):

$$G(x, y) = \exp\left\{\log\big(G_1(x)G_2(y)\big)\, A\left(\frac{\log\big(G_2(y)\big)}{\log\big(G_1(x)G_2(y)\big)}\right)\right\}. \quad (2.2)$$

Let, A(t) be the stracture of the dependence of the margins. The Pickands dependence function A is necessarily convex and is inside the triangle assigned by the points (0, 1), (1, 1), and (1/2, 1/2). A(t) shows the following:

1) $A(t)$ is convex;
2) $max\{(1-t), t\} \le A(t) \le t$; and
3) $A(0) = A(1) = 1$.

In the second property of A, $G(x,y) = min\{G_1(x), G_2(y)\}$, whereas the upper bound corresponds to (complete) independence.

2.4 Parametric Models of Copulas

2.4.1 Gumbel copula (logistic copula)

Invented by Gumbel (1960), the Gumbel or logistic, copula is the oldest of the EVC models. It belongs to both the extreme value and the Archimedean copulas. The dependence function A(w) is given as follws:

$$A(w) = [(1-w)^r + w^r]^{1/r}, \tag{2.3}$$

Where $r \geq 1$. The parameter r is the degree of dependence, ranging from complete independence ($r=1$) to complete dependence ($r=\infty$). Therefore, the Gumbel extreme value copula is given as

$$C(u_1,u_2) = \exp\left\{-[(-\ln u_1)^r + (-\ln u_2)^r)^{1/r}\right\}. \tag{2.4}$$

2.4.2 Galambos copula (negative logistic model)

Let, $\hat{C}_\phi$ be the distribution of the (1-U1,…, 1-Ud) random vector. The tail dependence function could be written as follow:

$$C_*(u_1,\dots,u_k) = exp\left[-\sum_{\substack{J\subset\{1,\dots,k\}\\ |J|\geq 2}} (-1)^{|J|}\left\{\sum_{j\in J}(-logu_j)^{-\alpha}\right\}^{-1/\alpha}\right]\prod_{j=1}^{k} u_{j,} \quad \alpha > 0,. \tag{2.5}$$

2.4.3 Tawn copula (asymmetric logistic copula)

The Tawn copula, or the (asymmetric logistic copula,) is much more flexible and combine several existing models such as the logistic($\emptyset = \theta = 1$), a mixture of logistic and independence models. Complete dependence corresponds to $\emptyset = \theta = 1$ and $r = \infty$, whereas complete independence corresponds to $\emptyset = 0$ or $\theta = 0$ or r=1. The dependence function is as follows:

$$A(w) = \ [\theta^r(1-w)^r + \emptyset^r w^r\ ]^{1/r} + (0-\emptyset)w + 1 - \theta, \qquad (2.6)$$

with $\emptyset \leq 1$ or $\theta \geq 0$ and $r \geq 1$, and the copula function

$$C(u_1,u_2) = \exp\left\{\ln u_1^{1-\theta} + \ln u_2^{1-\phi} - [(-\theta \ln u_1)^r + (-\phi \ln u_2)^r]^{1/r}\right\}. \qquad (2.7)$$

2.4.4 Husler-Reiss (HR) copula

The drawbacks of the logistic and the negative logistic copulas are that they are too limited for large dimensional problems since the dependence is described only by a single parameterθ. However, the HR copula does not have this problem; we give the corresponding distribution of the bivariate case:

$$C_*(u_1,u_2) = exp\left[\Phi\left\{\frac{a}{2} + \frac{1}{a}\log\left(\frac{logu_2}{logu_1}\right)\right\}\log u_1 + \Phi\left\{\frac{a}{2} + \frac{1}{a}\log\left(\frac{\log u_1}{\log u_2}\right)\right\}\log u_2\right], \qquad (2.8)$$

where Φ is the standard normal cumulative distribution function.

In our case, specifically, let u_1 be the ethanol price return marginal and "v" be the crude oil price marginal. We apply from the above mentioned discussion the four EV copulas to calculate the dependence of the two energy prices.

2.5 Kendall tau Dependence Measure

The Kendall tau can be expressed uniquely in terms of the copula; it is in the range [-1, 1].

$$\tau = 4\int_0^1\int_0^1 C(u,v)\,dC(u,v) - 1. \qquad (2.9)$$

Especially, in terms of the dependence function A, the particular Kendall tau is given as follows:

$$\tau = \int_0^1 \frac{t(1-t)}{A(t)} A''(t)dt. \tag{2.10}$$

For the EV copula, the Kendall tau can be explicated to be the best choice for the dependence function and the linear correlation coefficient.

2.6 Extreme Value Copulas

Extreme value copulas could be analyzed to find suitable models to obtain the dependence structure of the extreme values, with the presence of the component wise maxima. Here, we consider the bivariate case for our specific problem. Let $X_i = (X_{i1}, X_{i2}), i \in \{1,...,n\}$ be an i.i.d. sample random vectors with general distribution function F, margins F_1, F_2, and copula C_F. F is assumed to be continuous. Consider the vector of the component wise maxima:

$$M_n = (M_{n,1}, M_{n,2}), \qquad where\ M_{n,j} = \bigvee_{i=1}^{n} X_{ij}. \tag{2.11}$$

Because the joint functions of M_n are given by F^n and the marginal distributions are expressed by F_1^n, F_2^n, the copula is C_n of M_n:

$$C_n(u_1,u_2) = C_F(u_1^{1/n}, u_2^{1/n})^n. \tag{2.12}$$

It is clear that the extreme value copula is the same as the Generalized Extreme Value (GEV) distribution, which shares the max-stable property (Gudendorf and Segers, 2009). Therefore, the simple of extreme- value copulas could be obtained by employing the max-stability. Also, we can see from the literature studies that copula is max-stable if and only if it is an extreme- value copula. The understanding of extreme value copula is when we know the maxima distribution; here, we know the joint maxima distribution. This is the point at which the extreme value copula is different from other copulas, and also gives the evidence to use the GEV as the margin.

2.7 Literature Review

Martin Schlather (2001) studied those bivariate distributions that have their focus on the tail behavior by using Frechet margins, which can be distinguished by the coefficient about tail dependence and a weakened varying function. The feature is possible sometimes, and it is not denoted by the fact that the domain of attraction is focused on the distribution location of a bivariate extreme value distribution.

Results have shown that bivariate distributions, whose intensities are provided as a mixture of the densities of two Frechet variables in two, are completely independent and have two independent variables. The four members of this level have illustrated that the coefficient about tail dependence and the membership in the dominion of temptation of an extreme value distribution have two undiagnosed properties of distributions upon using unit Frechet margins.

SimlaTokgoz and Amani Elobeid (2006) studied the effect of price shocks in three output and input markets that are significant to ethanol: corn, sugar, and gasoline. The work examined the impact of shocks of these markets on ethanol and correlated agricultural markets in Brazil and the United States.

Their results show that the constituent of a country is a vehicle that defines the guidance for explaining ethanol consumption to the differentiation in gasoline price. Changes in feedstock costs influence the profitability of ethanol production and the household price of ethanol. In Brazil, commodities are influenced by sugarcane; thus, any changes in the sugar market influence the competing ethanol market.

Erik Brodin and Holger Rootzén (2008) studied hurricane and wind storm risks. This study developed extreme value methods that can be made applicable to storm insurance. The methods were used for the losses from 1982 to 2005 incurred by the biggest Swedish insurance company by using both a new bivariate and a univariate Generalized Pareto Distribution (GPD). The bivariate model led to reduced measures of risk, except for extreme events.

Their results demonstrate that the bivariate model offered a more substantive picture of real inaccuracies. Moreover, this model enabled the analysis of the effects of changes in the insurance portfolio, and showed that losses are somewhat unlinked to portfolio changes. In addition, it was observed that there was a low trend in the sizes of small personal claims.

Pál Rakonczai and Nader Tajvidi (2010) studied the modeling of extremes values, that is, multivariate peaks over threshold and extreme value distribution models, using multivariate generalized Pareto distributions. These models were compared in their ability to forecast extremes in wind speed data in many German cities. When using such extreme cases, fitting univariate margins requires some knowledge about the dependence structure.

Their results show that parametric cases of the bivariate maxima are completely developed. Another alternative way to model the dependence was gotten from non-parametric dependence functions. Investigating only the maxima can conceal the time feature inside the given period. Both exceedances and maxima have been used for bivariate datasets occurring from the wind time series that was estimated in northern Germany.

Kantaporn Chuangchid, Aree Wiboonpongse, Songsak Sriboonchitta and Chukiat Chaiboonsri (2012) studied the dependence structure of the extreme value of growth rate between factors having an effect on palm oil prices, which are crude oil price and soybean oil price, by using the Bivariate Extreme Value method. The data were taken from the daily soybean oil, crude oil, and palm oil prices ranging from July 1988 to January 2012.

Results show that the relationship between palm and soybean prices has some dependence at extreme price levels. But the growth rates of crude oil price and palm oil prices have only weak dependence at extreme price levels.

Mutita Kaewkheaw, Pisit Leeahtam, and Chukiat Chaiboosri (2012) studied the behavior of the U.S. dollar index and gold price. Their work employed bivariate extreme value copulas for illustrating the dependence structure between the return of the U.S. dollar index and gold price.

The results show that the returns on the U.S. dollar index and gold price are independent.

Ribatet Mathieu and Sedki Mohammed (2012) studied max-stable processes and extreme value copulas. The problem encountered was in modeling extreme values. The extreme value theory shows how to analyze max-stable distributions and set some limitations on the copulas to be managed. While the theory for multivariate extremes is well defined, it's normally guided outside the copula framework. Their paper focused on an application of modeling on extreme temperatures in Switzerland.

Valeri Natanelov, Andrew M. McKenzie, and Guido Van Huylenbroeck (2013) examined the relationships between ethanol, crude oil and corn prices during the period between 2006 and 2011 by using a holistic mapping of the present market situation and a contextual analytical design.

The results showed that the corn and crude oil markets have a strong relationship on one side, and the ethanol and crude oil markets on the other. Moreover, the price relationship between ethanol and corn is driven by the U.S. government fuel policy. Their analysis indicates that the corn market fluctuates in tandem with the levels of ethanol production. Therefore, when crude oil and/or corn prices are high, the situation leads to a competitive market for ethanol.