The Empirical Econometrics and Quantitative Economics Letters
ISSN 2286 – 7147
Volume 3, Number 3 (September 2014), pp. 13 - 23.

An analysis of the dependence between crude oil price and ethanol price using bivariate extreme value copulas

Aujcharapran Rojmaneebunpot, Chukiat Chaiboonsri and Jirakom Sirisrisakulchai

Faculty of Economics, Chiang Mai University

ABSTRACT

This paper studies the dependence structure between the returns of ethanol prices and crude oil prices. Since our focus was on the dependence behavior based on component-wise maxima, we applied the bivariate extreme value copulas for analysis. The risk measures considered in this paper are Value at Risk (VaR) and Expected Shortfall (ES). Our empirical study used two energy spot prices, the Chicago Ethanol Spot data and the North Sea (Forties) spot Crude Oil at the daily base. The results showed that after the middle of the year 2009, the dependence between these two energy sources in the U.S. was weak. This analysis could benefit those who are planning to invest in crude oil and ethanol. Moreover, it can guide government investors and private companies that manage power sector portfolios.

Keywords: Dependence, crude oil, ethanol, extreme value, copulas

1. Introduction

The world oil price plays an important role in the global economy as being the central source of energy. Therefore, the price of oil impacts government planning and policy decisions. In the recent past, international crude oil price has become extremely volatile, rising from about 30 dollars per barrel in 2003 to more than 100 dollars per barrel in 2008.

Ethanol is very popular in many countries that are currently experiencing high crude oil prices. However, the dynamics of the ethanol market were determined in the past by the agricultural commodity market, particularly the market for corn (Eidman 2005, USDA 2006). Since the year of 2006, the United States has been the biggest producer of ethanol, with more than 50% of the global production. Bioethanol is mainly produced by commodities such as sugar, cereals such as corn, and oilseed. Therefore, the growing demand for biofuels can stimulate an even higher demand for feedstock such as cereals. When this happens, the cost of ethanol production depends on the feedstock prices (Natanelove et al., 2011). Ethanol is now a notable source of motor-fuel. This leads to a heavy concern about the threat of enery crops to replace food crops (Kunanopadon and Suriya, 2012). When crude oil is still the main input for traditional gasoline production, there may be a relationship, or co-movement, between the expected prices of ethanol and crude oil.

The copula method is a powerful tool for linking the different margins and for measuring dependence compared to the classical linear correlation, and it is carried out by using the multivariate normal distribution. As for analyzing non-linear dependence, the copula can measure dependence for heavy-tail distributions and is flexible in the cases of parametric, semi-parametric, or non-parametric models. Moreover, copulas provide more details as compared to other tools. In the domain of finance, joint extreme events can bring about tremendous losses for investors. Therefore, this is the ideal field for applying the extreme value copula to measure the dependence structure between the extremely high and the extremely low price returns. Economists have implemented the application of the extreme value copulas in their studies. The joint behavior of extreme returns in the foreign exchange rate market has been analyzed by Starica (1999) and Lu, Tian and Zhang (2008). Also, the co-movement of equity markets that have been characterized by high volatility levels was investigated by Longin and Solnik (2001). Chuangchid et al (2012) analyzed the application of Extreme Value Copulas (EVT) in palm oil price. The data are from the futures prices of Singapore, Malaysia, and Dalian commodities by using the extreme value copula of Husler-Riess and Gumbel for estimation. The result showed that the EVT can illustrate the dependence structure for the palm oil futures prices of the Singapore, Malaysia, and Dalian commodities. Pokrivcak and Rajcaniova (2011) analyzed the statistical relationship between the ethanol, gasoline, and crude oil prices by evaluating the relationship between the variables in the Impulse Response Function (IRF) and the Vector Auto-regression (VAR). The result shows that oil has no co-integration between ethanol, and ethanol and gasoline, but that oil and gasoline prices have co-integration in the relationship. Shocks in oil price affect the cost price of gasoline.

This paper studies the tail dependence between the price of ethanol and the price of crude oil in the spot market by using bivariate extreme value copulas. The bivariate extreme value copula method can define and examine the extreme value, or extreme price, and dependence structure between two variables as opposed to the classical bivariate value copula which cannot define the extreme value in abnormal instances. Under abnormal circumstances, it can reasonably determine the price, and can help in maximizing profit as well as in effectively limiting risks. Limiting is important for governments, investors, researchers, and private companies as all of these are involved in the power sector. VAR and ES are the most common measures of risk. The new nonparametric measurement of Conditional Value at Risk (CVaR) and ES has been analyzed by ZongwuCai and Xian Wang (2008).

The remainder of this paper is systemized as follows: Section 2 reviews the concept of copulas, extreme value copulas and the Generalized Extreme Value distribution and also VaR and ES. Section 3 discusses the data used in the empirical analysis. Section 4 explains the empirical results. Finally, Section 5 offers a conclusion.

2. Methods

2.1 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:

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

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:

$$C_n(u_1,u_2) = C_F(u_1^{1/n}, u_2^{1/n})^n . \tag{2}$$

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.

2.2 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):

$$G(x; b, a, \xi) = exp\left\{-\left[1 + \xi\left(\frac{x-b}{a}\right)\right]^{-1/\xi}\right\}. \tag{3}$$

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. Parametric Models of Copulas

2.3.1 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 follows:

$$A(w) = [(1-w)^r + w^r]^{1/r} , \tag{4}$$

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

$$C(u_1, u_2) = \exp\left\{-[(-\ln u_1)^r + (-\ln u_2)^r)^{1/r}\right\}. \tag{5}$$

2.3.2 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 the follows:

$$C_*(u_1,\dots,u_k) = exp\left[-\sum_{\substack{J\subset\{1,\dots,k\}\\|J|\geq 2}}(-1)^{|J|}\left\{\sum_{j\in J}(-log u_j)^{-\alpha}\right\}^{-1/\alpha}\right]\prod_{j=1}^{k} u_{j,} \quad \alpha > 0,. \tag{6}$$

2.3.3 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:

$$A(w) = \ [\theta^r(1-w)^r + \emptyset^r w^r\ ]^{1/r} + (0-\emptyset)w + 1 - \theta, \tag{7}$$

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

$$C(u_1,u_2) = \exp\left\{\ln u_1^{1-\theta} + \ln u_2^{1-\phi} - [(-\theta \ln u_1)^r + (-\phi \ln u_2)^r\,]^{1/r}\right\}. \tag{8}$$

2.3.4 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:

$$C_*(u_1,u_2) = exp\left[\Phi\left\{\frac{a}{2}+\frac{1}{a}\log\left(\frac{log u_2}{log u_1}\right)\right\}\log u_1 + \Phi\left\{\frac{a}{2}+\frac{1}{a}\log\left(\frac{\log u_1}{\log u_2}\right)\right\}\log u_2\right], \tag{9}$$

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.

2.3.5 Kendall tau Dependence Measure

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

$$\tau = 4\int_0^1\int_0^1 C(u,v)\, dC(u,v) - 1. \tag{10}$$

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

$$\tau = \int_0^1 \frac{t(1-t)}{A(t)} A''(t)dt. \tag{11}$$

2.4. The Value at Risk and Expected Shortfall

2.4.1 The Value at Risk

The maximum aggregate loss which can occur with some given probability α . For a univariate risks, Value-at-Risk (VaR), which is the α quantile of the loss distribution function F-1(α) can measure this value, if F is increasing function, the VaR can achieve the unique value t, that is

$$F(t) = \alpha, \text{ where } VaR_\alpha = t. \tag{12}$$

For a multivariate risks, the multivariate VaR is difficult to define, since, for any continuous df F, there are possibly infinitely combinations (x, y), such that

$$\mathrm{F}(\mathrm{x},\ \mathrm{y}) = \alpha. \tag{13}$$

2.4.2 Expected Shortfall

The expected shortfall is exactly with the expected excess loss. However, the application of this concept to the bivariate context is the core of this section. For the univariate margins, the expected shortfall is

$$Es_\alpha = \frac{1}{\alpha}\int_0^\alpha VaR_\gamma\,(X)d\gamma. \tag{14}$$

Where VaR_γ is the γ quantile Value at Risk as last section.

3. Data

The data have been gathered from the Chicago Ethanol Spot data (USD per barrel) and the North Sea (Forties) spot Crude Oil (USD per barrel). The data are collected from EcoWin. The data span is from November 4, 2005, to December 26, 2013, at a daily frequency, which amounts to a total of 1,188 observations. The daily return was computed as $R_{i,t} = \ln(p_{i,t} / p_{i,t-1})$, where $P_{i,t}$ and $P_{i,t-1}$ are the daily spot prices for days “t” and “t-1” for market i.

4. Empirical Results

4.1 Statisticals Summary

Figure 1 shows the close relationship between the changes in ethanol prices and the changes in crude oil prices from late 2005 through 2009, when rapid expansion of the ethanol industry was taking place. In 2008, international crude oil price registered an incidence of extreme volatility and became a hot issue because the price had risen up to more than 100 dollars per barrel. In the second half of 2008, the crude oil price came tumbling down from the peak to just 33 dollars per barrel due to the severe financial crisis and economic recession that was caused by the American subprime crisis. Since we used the block maxima method, in which the block length is a calendar month (around 22 days), we received 38 maxima. Table 1 shows the descriptive statistics of the two energy commodity prices returns maxima.

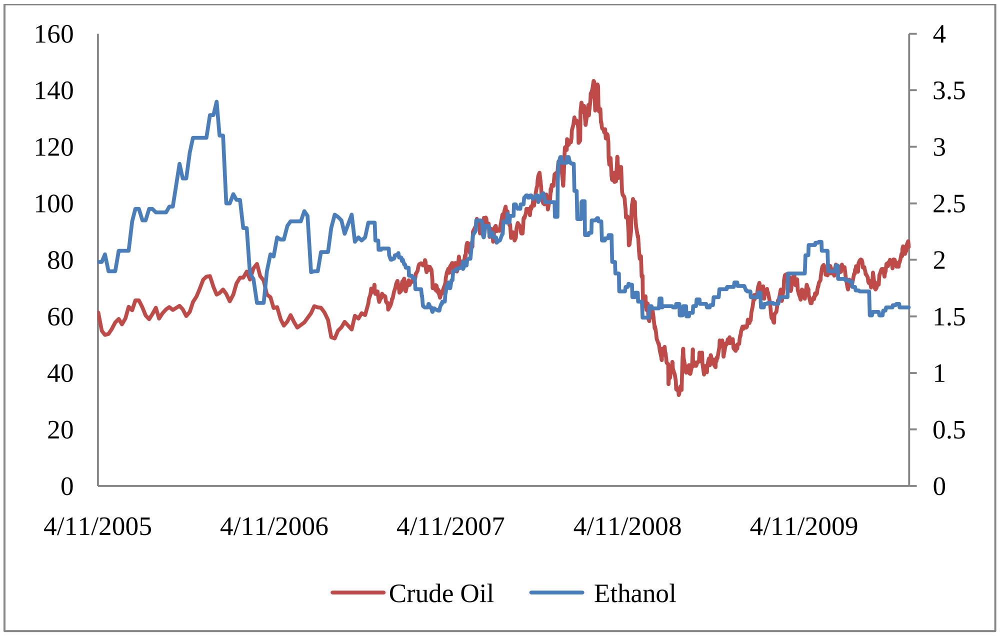

Figure 1. The co-movement of crude oil and ethanol price
Note: The primary axis is for the crude oil prices, and the secondary axis is for the ethanol prices.

TABLE 1. Statistical Summary of Ethanol Price and Crude Oil Price Returns

	Ethanol	Crude Oil
Minimum	-0.234	-0.180
Median	0.000	0.001
Mean	0.000	0.000
Maximum	0.159	0.220
Std. Dev.	0.026	0.030
Skewness	-1.751	0.064
Kurtosis	22.638	5.512
Jarque-Bera	18443.560	1070.874
Probability	0.000	0.000
Observation	839	839

The descriptive statistics test (see Table 1) showed that the returns of the ethanol prices are skewed to the right tail while the returns of the crude oil prices are almost symmetric, and that the excess Kurtosis is larger than zero. Therefore, the two distributions are in a higher peak when they are compared to the normal distribution. By using the Jarque-Bera test, we rejected the null hypothesis of a normal distribution at the 5% level for both the distributions. Therefore, we concluded that using an extreme value distribution was suitable for our study. The correlation between the two spot price returns is -0.035, which is very weak. But the dependence may be greater when we measure the dependence in the tail.

4.2 Empirical Results

The empirical results can be checked in Table 2. In the margin estimation, as we all know, the shape parameter ξ governs the tail behavior of the distribution. The sub-families defined by $\xi = 0$, $\xi > 0$, and $\xi < 0$ correspond, respectively, to the Gumbel, Fréchet, and Weibull families. In our case, the shape parameters are both greater than zero; therefore, the two margins are heavy-tailed: this again justifies the usefulness of the GEV distribution. In the dependence part, we compared the four copulas by using the AIC criterion. The HR copula is the best among them. The dependence parameter is equal to 0.465 and also significant. The dependence structure can be seen in Figure 4.2.

TABLE 2. The Estimation Results of Four Extreme Value Copulas

	Gumbel Copula		Galambos Copula		Tawn Copula		Husler-Reiss Copula	
	Ethanol	Oil	Ethanol	Oil	Ethanol	Oil	Ethanol	Oil
μ	0.034	0.042	0.034	0.041	0.034	0.042	0.034	0.042
	(0.005)***	(0.003)***	(0.005)***	(0.003)***	(0.005)***	(0.003)***	(0.005)***	(0.003)***
σ	0.029	0.0159	0.029	0.0159	0.029	0.016	0.029	0.015
	(0.005)***	(0.003)***	(0.005)***	(0.003)***	(0.005)***	(0.003)***	(0.005)***	(0.003)***
ξ	0.433	0.39	0.432	0.378	0.415	0.409	0.426	0.371
	(0.147)***	(0.194)***	(0.158)***	(0.195)***	(0.139)**	(0.203)**	(0.171)***	(0.193)**
r	1.076(0.157)**		0.239(0.33)		0.246(0.352)		0.465(0.164)***	
AIC	325.978		325.778		325.902		325.738	

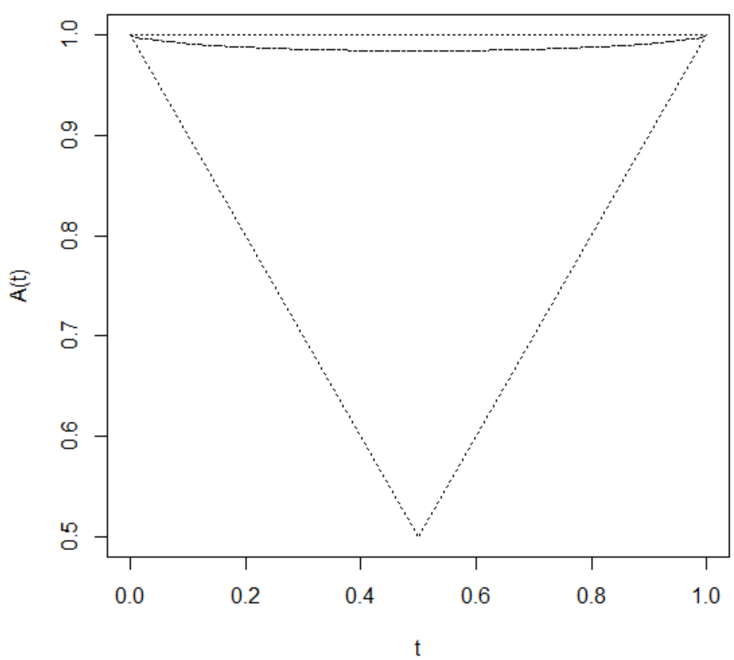

Figure 2. The dependence function of the returns between ethanol price return and crude oil price return, using the HR copula.

The dependence function of A(t) of the HR copula is presented in Figure 2. Based on the table, we see that the results are consistent with the Kendall tau measure (= 0.02), which we have calculated and presented in Table 2. The dependence is close to the upper bound of A(t), and the dependence structure between the oil price and the ethanol price is weak, as we expected.

4.3 Goodness of Fit of Extreme Value Copula

For this part of the study, we used the Cramér–von Mises statistics to test whether or not our data were fit for the selected EV copula: the details of the statistics can be found in the work by Genest, Kojadinovic, G. Nešlehová, and Yan (2010). The statistic is as follows:

$$S_n = \int_0^1 n|A_n(t) - A_{\theta n}(t)|^2 \, dt. \tag{12}$$

The results are demonstrated in Table 4.2. Since the null hypothesis of Cramér–von Mises is the data fit for the specific copulas, all of the copula, except for the Tawn copula, do not reject the null hypothesis.

TABLE 3. Cramér–von Mises Statistics

	Gumbel Copula	Galambos Copula	Tawn Copula	Husler-Reiss Copula
the statistics	0.042	0.029	0.150	0.023
p-value	0.348	0.462	0.016	0.490

Note: The p-value was obtained by using a boots tapping process.

4.4 The Value at Risk and Expected Shortfall

4.4.1 The Value at Risk

For risk managers, the important thing to determine is that the maximum aggregate loss which can occur with some given probability α . For a univariate risks, Value-at-Risk (VaR), which is the α -quantile of the loss distribution function F-1(α) can measure this value, if F is an increasing function, the VaR can achieve the unique value t, that is $F(t) = \alpha,$, such that $VaR_\alpha = t$. For a multivariate risks, the multivariate VaR is difficult to define, since, for any continuous df F , there are possible infinitely combinations (x, y), such that $\mathrm{F}(\mathrm{x},\ \mathrm{y}) = \alpha$. Therefore the solutions to the multivariate VaR is a set.

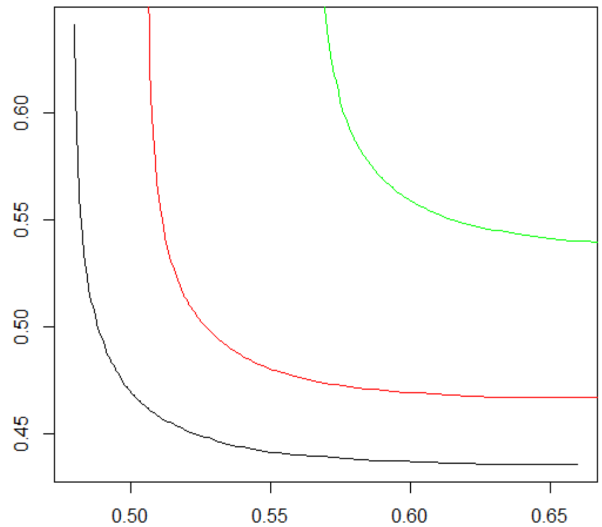

Figure 3. The multivariate VaR for ethanol price and crude oil price

To apply the idea of multivariate VaR to our study, we used the results of Husler-Reiss Copula since it has the smallest AIC. The steps of our analysis are the following, to achieve the VaR0.90 for the bivariate distribution, First, fix the first margin by establishing ethanol price as the 90% quantile of F1, with the GEV distribution parameters of ethanol price, that is, 0.480.
Second, given the first margin, make the quantile of bivariate distribution be equal to 90%, use the numerical method. We get the second margin 0.641, which is 99.9% quantile of F2.
Third, repeat the above steps 100 times with accumulated 0.01 quantile of F1 each time. This can be implemented by trying a 91% quantile, 92% quantile,…, etc., we make 100 points and then draw the curves.
Fig.3 shows that the black curve is VaR0.90 for bivariate risk, and the red curve is bivariate VaR0.95. The Green curve is the bivariate VaR0.99. VaR of the bivariate distribution is the sum of two margins. For the bivariate VaR0.90, it is between [0.96, 1.121]. For the bivariate VaR0.95, it is between [1.025, 1.177], and the last for VaR0.99 of the bivariate distribution is between [1.158, 1.312]. Therefore, we receive a range of VaR, which has the worst and best situation. In our case, the VaRs is not much different than the independent copula, since the dependency parameter is quite small, the dependence is weak.

4.4.2 Expected Shortfall
If the VaR ask a question like "how bad is it when the loss happens", the expected shortfall ask "if the loss really happens, then what is our expected loss?" The expected shortfall also has several names, such as Conditional Value at Risk (CVaR) which is defined as the Excess Loss and Shortfall. Every name provides some different explanations to the expected shortfall. The most understandable one is the last one, since the expected shortfall is exactly the expected excess loss. However, it is the application of this concept to the bivariate context which is the core of this section. For the univariate margins, the expected shortfall is $Es_{\alpha} = \frac{1}{\alpha}\int_0^{\alpha} VaR_{\gamma}\,(X)d\gamma$, where VaR_{γ} is the γ quantile Value at Risk as the last section.
According to the different points in the VaR line, we get different expected shortfall for different combinations. For example, for the first point of the bivariate VaR0.90, they are the 90% quantile for the first margin, and 99.9% quantile for the second margin. To get the bivariate shortfall, we calculate the two shortfalls for each and sum them up. The shortfall of first margin is 0.498 and the

second margin is 0.661, which is slightly higher than the VaR. and also the bivariate shortfall is 1.160. According to this logic, we also can draw the bivariate shortfalls as:

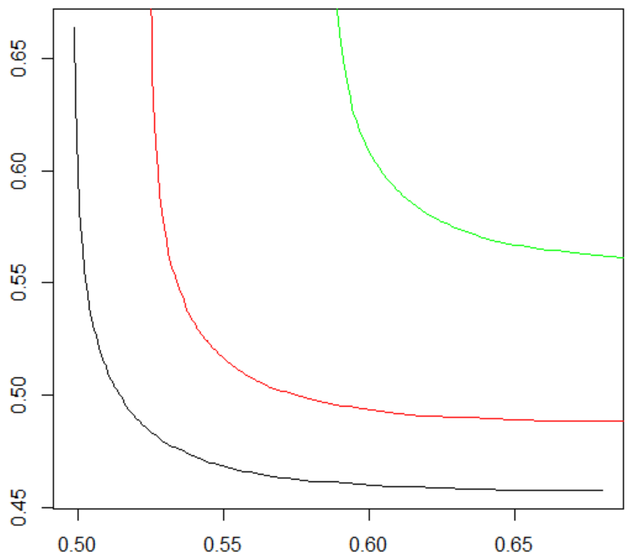

Figure 4. The expected Shortfall of ethanol price and crude oil price

Fig.4 shows that the black curve is 0.90% Shortfall for bivariate risk, and the red curve is bivariate 0.95% Shortfall. The Green curve is the bivariate 0.99% Shortfall. Shortfall of the bivariate distribution is the sum of two margins. For the bivariate Shortfall 0.90, it is between [1.008, 1.162]. For the bivariate Shortfall0.95, it is between [1.066, 1.219], and the last for Shortfall0.99, it is between [1.200, 1.354]. This is the same as the multivariate VaR, which has a "worst and best situation".

5. Conclusion

It is important to know the dependence structure of the dependence between the two energy spot prices since the structure can have an impact on both current energy policy and future policies. Our empirical results show that after 2009, the dependence relationship between oil and ethanol prices in the U.S. was weak. There are several reasons for using the extreme value copula to characterize the dependence between the two variables: First, using the copula makes it possible to avoid the unsatisfying assumption of independence; this method can capture a nonlinear relationship. Second, the extreme value copula provides for modeling the dependence between the constitutive relation parameters of a random couple which illustrates the greatest values of two properties considered over the same period of time. This is an important issue in our study because the extreme tail dependence and co-movement of prices of traditional energy and new energy is a subject matter that calls for close attention and analysis.

6. Acknowledgments

The authors would like to express their gratitude, particularly, to Prof. Songsak Sriboonchitta, and Miss. Gong Xue and Miss Mutita Keawkheaw, Faculty of Economics, Chiang Mai University, Thailand, for their helpful suggestions and comments.

7. References

Cai, Zongwu and Xian Wang. 2008. "Nonparametric estimation of conditional VaR and expected shortfall". *Journal of Econometrics 147*: pp. 120–130.

Chaithep, K. 2012. "Value at Risk analysis of Gold price return using Extreme Value Theory". Master's Thesis of Economics Chiang Mai University.

Chuangchid, K., et al. 2012. "Application of Extreme value Copulas to palm oil prices analysis". *Business Management Dynamics 2*, (July 2012): 25-31.

Chuangchid, K., et al. 2012. "Factors Affecting Palm Oil Price Based on Extremes Value Approach". *International Journal of Marketing Studies 4, 6*; 2012.

Chuangchid, K., et al. 2013. "Predicting Malaysian palm oil price using Extreme Value Theory". *International Journal of Agricultural Management 2, 2* (January 2013):91-99.

CME Group. 2012. "Ethanol outlook report". Commodity Research Bureau.Retrivedfrom http://cmegroup.barchart.com/ethanol/archive/1340628849 CME-Weekly-Ethanol-25-Jun-2012.pdf.

Gumbel, E. J. 1960. "Bivariate exponential distribution". *Journal of the American Statistical Association 55,* 292: pp. 698-707.

Huang, J-J.et al. 2009. "Estimating value at risk of portfolio by conditional copula-GARCH method". Insurance: *Mathematics and Economics 45*: 315-324.

Kunanopadon, Sarunut and Komsan Suriya. 2012. "What would happen to the economy when energy crops replace food crops? A case of gasohol production in Thailand" *The Empirical Econometrics and Quantitative Economics Letters 1*, 2: pp. 111-122.

Lai, L., and Wu, P. 2007. "An Extreme Value Analysis of Taiwan's Agriculture Natural Disaster loss data". International Conference on Business and Information (BAI). Tokyo, Japan.

Ning, C. and Wirjanto, S.T. 2009. "Extreme return–volume dependence in East-Asian stock markets: A copula approach". *Finance Research Letter, 6*: pp. 202-209.

Segers, J. 2005. "Extreme-Value Copulas". *Medium Econometrische Toepassingen 13,* 1: pp. 9-11.

SimlaTokgoz and AmaniElobeid. 2006. "An Analysis of the Link between Ethanol, Energy, and

Zhang Y and Wang Z. 2013. "Investigating the price discovery and risk transfer functions in the crude oil and gasoline futures markets: Some empirical evidence". *Applied Energy 104*: pp.220-228.