International Journal of Energy Economics and Policy

ISSN: 2146-4553

available at http: www.econjournals.com

International Journal of Energy Economics and Policy, 2018, 8(5), 193-201.

Analyzing Extreme Comovements in Agricultural and Energy Commodity Markets Using a Regular Vine Copula Method

Kunlapath Sukcharoen[1]*, David Leatham[2]

[1]West Texas A&M University, United States, [2]Texas A&M University, United States. *Email: ksukcharoen@wtamu.edu

ABSTRACT

Using a regular vine copula approach, this paper analyzes the dependence structure and tail dependence patterns among daily prices of three agricultural commodities (corn, soybean, and wheat) and two energy commodities (ethanol and crude oil) from June 2006 to June 2016. Our findings suggest that the prices of corn and crude oil are linked through the ethanol market, which are consistent with the results from previous studies. We also find that crude oil and agricultural commodity prices are statistically dependent during the extreme market downturns but independent during the extreme market upturns. In addition, the results from our sub-sample analysis show that both the upper and lower tail dependence between crude oil and other commodity markets become weaker in the recent years when the ethanol market became more mature.

Keywords: Agricultural Markets, Energy Markets, Price Dependence, Tail Dependence, Vine Copulas
JEL Classifications: C53, C58, G11, G17, Q13, Q40

1. INTRODUCTION

It is widely believed that the expansion of the ethanol production in the United States has reshaped the linkages between agricultural and energy commodity markets. Traditionally, agricultural and energy commodity markets have been linked through the input channel (in terms of production and transportation costs). Now, due primarily to the very rapid expansion of crop-based ethanol production in the United States, the agricultural and energy commodity markets are increasingly connected through the demand channel - mainly through the policy-driven demand for corn as an ethanol feedstock - rather than the input channel[1]. In particular, various studies have reported a tighter linkages between agricultural and energy commodity markets since the boom of the U.S. ethanol industry took off in 2006.

For instance, Muhammad and Kebede (2009) find that from 2005 to 2008 oil price movements could explain >60% of the change in corn prices, whereas from 1990 to 2004 only about 2% of the change in corn prices could be explained by oil prices. In addition, Tyner (2010) shows that, for the 1988-2005 period, the correlation between oil and corn prices is low and negative (−0.26). However, since 2006 there appears to be a strong and positive correlation between the prices of oil and corn. Indeed, Tyner (2010) reports a 0.80 (0.95) correlation between oil and corn prices for the period 2006-2008 (2008-2009). Hertel and Beckman (2012) also document a similar change in the correlation pattern between oil and corn prices.

Increased connection between agricultural and energy commodity markets raises the need for deeper understanding of the links and comovements between agricultural and energy commodity price returns. Over the last decade, a number of empirical studies have examined the interrelationship between agricultural and energy commodity markets[2]. Motivated by concurrent swings in agricultural and energy commodity prices experienced after the

1 Specifically, the Renewable Fuel Standard (RFS) program, created under the Energy Policy Act of 2005 and later expanded under the Energy Independence and Security Act of 2007, is the root of cause of the rapid growth in corn-based ethanol production in the United States. Compared to 3.9 billion gallons of biofuel produced in 2005, the act requires that 36 billion gallons be produced in 2022. Of the 36 billion gallons of biofuels, at least 21 billion gallons must come from advanced biofuels and the remainder, at most 15 billion gallons, can come from conventional biofuels such as corn-based ethanol (Schnepf and Yacobucci, 2013).

2 See Serra and Zilberman (2013) and Zilberman et al. (2013) for a literature review on price linkages and transmission patterns in biofuel-related markets.

change in U.S. biofuel policies in 2006, some of these studies focus on the question of whether the ethanol/biofuel boom has caused a stronger dependence between agricultural and energy commodity prices. However, the results from these studies are rather mixed.

For example, Campiche et al. (2007) show that the prices of corn and soybean - the key agricultural commodities used for ethanol/biofuel production - are cointegrated with crude oil prices over the period of 2006-2007 but not during the period 2003-2005. Similarly, Du and McPhail (2012), Kristoufek et al. (2012), and Lucotte (2016) examine the connections among the prices of agricultural and energy commodities before and after the food crisis of 2007/2008, and find that their prices are much more closely linked after the crisis. On the contrary, Gilbert (2010) reports that the 2007-2008 agricultural price spikes could mostly be explained by macroeconomic and monetary factors, and that the biofuel demand growth is not the main cause of the agricultural price booms. Consistent with Gilbert (2010), Reboredo (2012) investigates extreme market dependence between oil and agricultural commodity prices using copulas, and finds that price spikes in the corn and soybean markets during the period 2007 to 2011 are not caused by extreme upward oil price movements. In addition, Baumeister and Kilian (2014) show that there is no compelling evidence that the change in U.S. biofuel policies in May 2006 has created a tight link between oil and agricultural commodity markets.

Despite a number of empirical studies on the agriculture-energy nexus, relatively little attention has been paid to the dependence structure between agricultural and energy commodity prices and their extreme comovements. Reboredo (2012) and Han et al. (2015) are among the few recent authors who analyze tail dependence patterns (or extreme comovements) between the prices of agricultural and energy commodity during the last decade. Considering weekly data from January 1998 to April 2011, Reboredo (2012) employs several bivariate copulas to study the extreme market dependence between oil prices and agricultural commodity prices (namely, corn, soybean and wheat prices). The results from his study indicate that agricultural commodity prices are independent of extreme upward price movements in the oil market even in the last 3 years of the sampling period. Han et al. (2015) investigate tail dependence between the returns on agricultural and energy commodity indices using a time varying symmetrized Joe-Clayton copula. Using daily data from January 2000 to January 2014, their results suggest that both lower and upper tail dependence are strongest during the financial crisis of 2008. Similar to Reboredo (2012), they find that lower tail dependence is in general stronger than upper tail dependence. While these studies provide useful information on the dependence structure as well as tail dependence between two commodity markets, little is still known on the multivariate dependence structure of agricultural and energy commodity markets.

Accordingly, this paper attempts to fill the gap in the literature by analyzing the dependence structure among daily returns of three agricultural commodities - corn, soybean, and wheat - and two energy commodities - ethanol and crude oil - using a regular vine (or R-vine) copula methodology. The major advantage of the R-vine copula approach is that it allows us to capture potentially complex dependence structure and tail dependence patterns in a multivariate framework. Therefore, it allows us to uncover not only information regarding the upper and lower tail dependence between any two commodity markets but also information regarding the overall connections among multiple commodity markets. Furthermore, we add to the literature by examining whether and how the dependence structure and the degree of tail dependence change between the two periods of ethanol production: The rapid growth period (June 2006-June 2011) and the slowing growth period (June 2011-June 2016). Our findings should provide valuable information for practitioners, academics and policy makers regarding the linkages between the agricultural and energy commodity markets. In addition, as agricultural and energy commodity markets are often thought of as an alternative market for risk diversification purposes, the results from this study should also provide useful information for investors about portfolio diversification and risk management.

The remainder of this paper is organized as follows. Section 2 describes the data used in our analysis. Section 3 is devoted to explaining the regular vine copula methodology. Section 4 presents the results of the empirical analysis, and Section 5 concludes the paper.

2. DATA

Our empirical analysis is based on daily prices for three agricultural commodity futures: Corn, soybean and wheat futures; and two energy commodity futures: Ethanol and crude oil futures. All prices are obtained from the Datastream database. The price data span from June 1, 2006 to June 30, 2016, from which a sample of daily log return series are constructed using the nearest futures contracts. At the rollover date, care has been taken to ensure that the same futures contract is used to calculate the daily log returns. This yields a total of 2536 observations for each return series. Apart from examining the dependence structure of the five commodity markets for the whole sample period, we also investigate how the dependence structure and the degree of tail dependence change over the two sub-periods: June 2, 2006-June 16, 2011 and June 17, 2011-June 30, 2016. The first sub-period corresponds to the period of rapid expansion of ethanol production in the United States, whereas the second sub-period corresponds to the period of slowing growth in ethanol production (U.S. Energy Information Administration, 2011). Each sub-period has a total of 1268 observations.

Table 1 reports summary statistics (Panel A) and correlation matrix (Panel B) for the daily log returns on the futures contracts of corn, soybean, wheat, ethanol, and oil for the entire sample period. For each return series, the mean is very small relative to its standard deviation. As expected, oil returns are the most volatile series among the five commodity returns. All returns series are only slightly skewed, but have a high excess kurtosis (especially for the soybean, ethanol, and oil return series). The significant Jarque-Bera (JB) test statistics indicate that all daily log returns are not normally distributed. The Augmented Dickey-Fuller tests show that all commodity returns are stationary.

Table 1: Summary statistics and correlation analysis on daily log returns on the futures contracts of corn, soybean, wheat, ethanol, and oil

Commodity	Corn	Soybean	Wheat	Ethanol	Oil
Panel A: Summary statistics					
Mean (%)	−0.003	0.062	−0.035	0.094	−0.051
Standard deviation (%)	1.974	1.654	2.191	1.874	2.390
Skewness	−0.102	0.385	−0.012	−0.561	0.082
Excess kurtosis	1.813	10.684	1.503	5.786	3.904
Minimum (%)	−10.409	−8.141	−9.973	−16.990	−`13.065
Maximum (%)	8.6618	20.3209	8.7943	9.7525	14.5464
JB	353.30*	12149.00*	239.93*	3678.90*	1618.20*
ADF	−36.05*	−36.11*	−36.23*	−33.283*	−36.90*
ρ(1)	0.019	0.010	0.000	0.129*	−0.061*
Q (5)	4.98	3.65	3.22	49.66*	16.35*
Q_2 *(5)*	167.50*	50.79*	219.80*	285.40*	699.50*
Panel B: Correlation matrix					
Corn	1.000				
Soybean	0.610	1.000			
Wheat	0.647	0.469	1.000		
Ethanol	0.546	0.406	0.413	1.000	
Oil	0.291	0.342	0.247	0.308	1.000

Summary statistics (Panel A) and correlation matrix (Panel B) are presented for daily log returns on the futures contracts of corn, soybean, wheat, ethanol, and oil for the period June 2, 2006-June 30, 2016. The total number of observations is 2536 for each return series. JB is the JB test statistic, where * denotes the rejection of the null hypothesis of normality at the 1% significance level. ADF is the Augmented Dickey-Fuller test statistic, where *denotes the rejection of the null hypothesis that the respective return series follows a unit root process at the 1% significance level. ρ(1) is the first-order autocorrelation, where *denotes the rejection of the null hypothesis that the first-autocorrelation of the respective return series is equal to zero at the 1% significance level. Q (5) is the Ljung-Box test statistic for the return series, where *denotes the rejection of the null hypothesis that there is no serial correlations in the return series up to order 5. Q^2 (5) is the Ljung-Box test statistic for the squared return series, where *denotes the rejection of the null hypothesis that there is no ARCH effect in the return series up to order 5.

Both first-order autocorrelation (ρ(1)) and Ljung-Box (Q(5)) tests indicate that there are serial correlations in the two energy commodity returns but not in the three agricultural commodity returns. In addition, the Ljung-Box test statistics for the squared return series (Q^2 (5)) suggest that ARCH effects (or volatility clustering) are present in all return series. Unconditional correlations provide evidence of weak dependence between oil and other commodities. More specifically, the linear correlation coefficients between oil and other commodities range between 0.247 (for the pair of oil and wheat returns) and 0.342 (for the pair of oil and soybean returns). The highest linear correlation is found between corn and wheat returns (0.647).

3. METHODOLOGY

In this paper, we apply the R-vine copula approach to study the dependence structure and tail dependence (or extreme comovements) among returns of three agricultural commodity futures (corn, soybean, and wheat futures), and two energy commodity futures (ethanol and crude oil futures). The approach consists of three stages. The first stage involves modeling the marginal distributions for the individual commodity returns. In the second stage, the R-vine copula is estimated using the standardized residuals obtained from the first stage. The third stage involves calculating the upper and lower tail dependence coefficients.

3.1. Modeling Marginal Distributions

In the first stage, the marginal distributions for all commodity returns are modeled. To account for possible serial correlation and volatility clustering in commodity returns, we consider four alternative GARCH models: A GARCH(1,1) model with a constant unconditional mean, an AR(1)-GARCH(1,1) model, an MA(1)-GARCH(1,1) model, and ARMA(1,1)-GARCH(1,1) model. Let $y_{i,t}$ denote the daily log return for commodity *i*. The ARMA(1,1)-GARCH(1,1) model is specified as follows:

$$y_{i,t}=\mu_i+\phi_i y_{i,t-1}+\theta_i\varepsilon_{i,t-1}+\varepsilon_{i,t} \tag{1}$$

$$\sigma_{i,t}^2=\omega_i+\alpha_i\varepsilon_{i,t-1}^2+\beta_i\sigma_{i,t-1}^2 \tag{2}$$

$$z_{i,t}=\frac{\varepsilon_{i,t}}{\sigma_{i,t}}\sim i.i.d.D(0,1) \tag{3}$$

Where D(0,1) is a zero mean and unit variance probability distribution[3]. For each return series, the mean model with the lowest Bayesian Information Criterion (BIC) is chosen. The series of standardized residuals, $z_{i,t}$, is then transformed into a standard uniform variable or copula data (denoted as $\mu_{i,t}$) using an empirical distribution function (EDF). The series $z_{i,t}$ is also referred to as filtered returns. Several goodness-of-fit tests are performed to ensure that the marginal distributions are appropriately specified.

3.2. Selecting and Estimating Regular Vine Copula

The second stage involves estimating the R-vine copula using the standard uniform variables obtained from the first stage. Simply put, an R-vine copula is a multivariate distribution for which the marginal distribution of each variable is standard uniform. Let $z=(z_1,z_2,z_3,z_4,z_5)$ be a five-dimensional random vector of filtered commodity returns (or standardized residuals in equation (3)) with a joint distribution function $F(z)$ and a joint density function $f(z)$. According to the Sklar's theorem (Sklar, 1959), the joint distribution of **z** can be expressed as:

$$F(z)=C(u_1,u_2,u_3,u_4,u_5) \tag{4}$$

Where $C: [0,1]^5\rightarrow[0,1]$ is a copula function, and $u_i=F_i(z_i)$ is the marginal distribution function of z_i for $i=1,2,\ldots,5$. Suppose that **C** and F_i are differentiable. Then, the joint distribution function of **z** can be written as:

$$f(z)=f_1(z_1)\, f_2(z_2)\ldots f_5(z_5)c(F_1(z_1),F_2(z_2),\ldots,F_5(z_5)) \tag{5}$$

Where *C* is the density of the copula and $f_i=(f_i z_i)$ is the density of $Fi=F_i(z_i)$.

In particular, the copula function represents the dependence structure of a multivariate random vector of filtered commodity returns. Thus, we can use multivariate copulas to analyze tail dependence among multiple commodity returns. Two most obvious

3 For the GARCH (1,1) model with a constant unconditional mean, both ϕ_i and θi are set to zero. θ_i is set to zero for the AR(1)-GARCH (1,1) model, whereas ϕ_i is set to zero for the MA (1)-GARCH (1,1) model.

choices of multivariate copulas are the Gaussian and Student's t copulas. However, the Gaussian copula cannot capture non-linear dependence between random variables. In other words, it assumes that the dependence pattern between each pair of variables does not change with market conditions. In addition, it unrealistically imposes independence in the tails or during extreme market movements. On the other hand, the Student's t copula allows us to capture tail dependence. However, it requires all pairs of random variables to have exactly the same degree of tail dependence, which seem to be unrealistic. Therefore, both Gaussian and Student's t copulas are too restrictive, especially when modeling the dependence structure of more than two random variables.

This paper exploits the more flexible multivariate copula construction method ("pair-copula construction (PCC) method"). The PCC method was first proposed by Joe (1996) and further extended by Bedford and Cooke (2001; 2002) and Kurowicka and Cooke (2006). The idea of the PPC method begins by factorizing a joint density function into marginal and conditional density functions. For example, a five-dimensional density function can be factorized as:

$$f(z)=f_1(z_1)f_{2|1}(z_2|z_1)f_{3|1,2}(z_3 \mid z_1,z_2)f_{4|1,2,3}(z_4 \mid z_1,z_2,z_3)f_{5|1,2,3,4}(z_5|z_1,z_2,z_3,z_4) \quad (6)$$

Using the Sklar's theorem, any conditional marginal distributions in the right hand of (6) can be expressed as:

$$f(z_i \mid v)=c_{z_i,v_j|v_{-j}}(F(z_i \mid v_{-j}),F(v_j \mid v_{-j}))f(z_i|v_{-j}) \quad (7)$$

With

$$F\left(z_i|v\right)=\frac{\partial C_{z_i,v_j|i_{-j}}\left(F\left(z_i|v_{-j}\right),F(v_j\,|v_{-j})\right)}{\partial F(v_j\,|v_{-j})} \quad (8)$$

Where v is the conditioning set of marginal distribution of z_i, v_i is a variable in the set v, and $v\text{-}j$ is the set of variables in v excluding v_i. For example, $f_{2|1}(z_2|z_1)$ an be written as $c_{1,2}(F_1,F_2)f_2$, and $f_{3|1,2}(z_3|z_1,z_2)$ an be expressed as $c_{2,3|1}(F_{2|1},F_{3|1})f_{3|1}$. Accordingly, the joint density function, $f(z)$, can be decomposed as products of bivariate copula densities and marginal density function of z_i

Obviously, the factorization in (6) is not unique. This suggests that there are a large number of possible PCCs from which to choose. Bedford and Cooke (2001) introduce a graphical structure called regular vine (or R-vine) structure to help organize different decompositions. In particular, a five-dimensional R-vine structure is defined by a sequence of four trees: T_1, T_2, T_3, T_4. T_1 has five nodes and four edges. Edges in T_1 then become nodes in T_2. The two nodes in T_2 are connected by an edge only if they share a common node in T_1 (proximity condition). Edges in T_2 then become nodes in T_3. Again, the two nodes in T_3 are connected by an edge only if they share a common node in T_2 (proximity condition). Then, edges in T_3 become nodes in T_4, and the two nodes in T_4 are connected by an edge only if they share a common node in T_3. The joint density function of a five-dimensional R-vine copula is given by[4]:

4 Kurowicka and Cooke (2006) provide the derivation of the joint density of a general R-vine copula.

$$f\left(z\right)=\prod_{k=1}^{5}f_k\prod_{i=1}^{4}\prod_{e\in E_i}c_{j(e),k(e)|D(e)}\left(F_{j(e)|D(e)},F_{k(e)|D(e)}\right) \quad (9)$$

Where E_i is a set of edges in T_i J(e)and $k(e)$ are the two (conditional) nodes associated with each edge e, $D(e)$ is the conditioning set associated with edge e. An example of a five-dimensional R-vine structure is illustrated in Figure 1, and its corresponding joint density function is:

$$f(z)=f_1f_2f_3f_4f_5c_{1,2}c_{1,4}c_{2,3}c_{2,5}c_{1,3|2}c_{2,4|1}c_{3,5|2}c_{1,5|2,3}c_{3,4|1,2}c_{4,5|1,2,3} \quad (10)$$

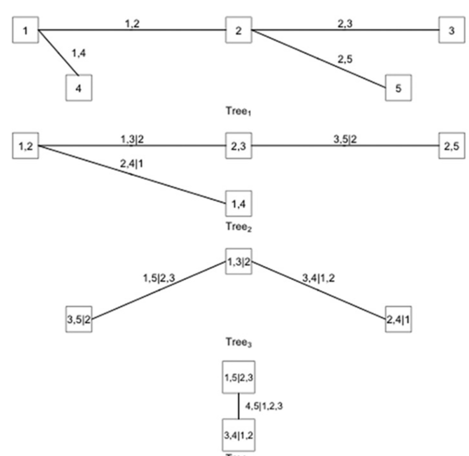
According to Morales-Nápoles et al. (2010), there exist $(d!/2)\times 2(\frac{d-2}{2})$ different R-vine decompositions for the d-dimensions. Thus, there are 480 possible R-vine structures for a five-variate density. In this study, we employ a sequential estimation procedure proposed by Dißmann et al. (2013) to select and estimate an R-vine structure as well as its corresponding bivariate copulas. This procedure begins in the first tree of the R-vine structure. The structure of the first tree is formed by maximizing the sum of the absolute values of pairwise Kendall's tau coefficients. For the first tree, this is done using the standard uniform variables obtained from the first stage. Given the selected structure, the pair copulas are chosen from a range of 39 different parametric bivariate copula families by minimizing the BIC[5]. Copula parameters are estimated using the maximum likelihood estimation (MLE) method. Once the first tree is specified and the pair-copula families are chosen, the same is done for the second, the third, and forth trees using the transformed observations, $F(z_i \mid v)$, calculated from equation (8).

3.4. Measuring Tail Dependence Coefficients

In the third stage, the upper and lower tail dependence coefficients are calculated to measure the degree of comovements between two

5 The 39 bivariate copula families include Gaussian, Student's t, Clayton, Gumbel, Frank, Joe, BB1 (Clayton-Gumbel), BB6 (Joe-Gumbel), BB7 (Joe-Clayton), BB8 (Joe-Frank), Tawn type 1, Tawn type 2, and the rotated versions (90°, 180° and 270°) of Clayton, Gumbel, Joe, BB1, BB6, BB7, BB8, Tawn Type 1, and Tawn Type 2 copulas.

Figure 1: Estimated R-vine copula structure with 1 = Ethanol, 2 = Corn, 3 = Soybean, 4 = Oil, and 5 = Wheat



commodity markets at the extreme events. The upper and lower tail dependence coefficients for commodities i and j are defined, respectively, as:

$$\lambda_U = \lim_{u \to 1^-} \Pr\left[Z_i > F_i^{-1}(u) | Z_j > F_j^{-1}(u) \right] = \lim_{u \to 1^-} \frac{1-2u+C(u,u)}{1-u} \tag{11}$$

$$\lambda_L = \lim_{u \to 0^+} \Pr\left[Z_i < F_i^{-1}(u) | Z_j < F_j^{-1}(u) \right] = \lim_{u \to 0^+} \frac{C(u,u)}{u} \tag{12}$$

In this study, we follow Loaiza et al. (2015) and calculate the non-parametric tail dependence coefficients for all pairs of commodities through a simulation exercise. Specifically, the estimated R-vine copula density obtained from the second stage is used to generate S=10,000 draws of the five standard uniform variables, $\{u_{(1,s)}, u_{(2,s)}, u_{(3,s)}, u_{(4,s)}, u_{(5,s)}\}_{(s=1)}^{(s=10,000)}$. This simulation exercise is replicated $R = 1000$ times. For each replication r, the upper and lower tail dependence coefficients are respectively estimated using the following non-parametric estimators:

$$\hat{\lambda}_U^r = \lim_{k_U \to S^-} \frac{1 - 2\frac{k_U}{S} + \hat{C}(\frac{k_U}{S}, \frac{k_U}{S})}{1 - \frac{k_U}{S}} \tag{13}$$

$$\hat{\lambda}_L^r = \lim_{k_L \to 0^+} \frac{\hat{C}(\frac{k_L}{S}, \frac{k_L}{S})}{\frac{k_L}{S}} \tag{14}$$

Where k_U/S and k_L/S are the thresholds used in the estimation of tail dependence coefficients. Similar to Loaiza et al. (2015), we set $k_U/S = 0.99$ and $k_L/S = 0.01$. $\hat{C}(k_1/S, k_2/S)$ is the empirical copula, which can be estimated using:

$$\hat{C}\left(\frac{k_1}{S}, \frac{k_2}{S}\right) = \frac{1}{S}\sum_{s=1}^{S} 1\left(F_i\left(z_{i,s}\right) \le \frac{k_1}{S}, F_j\left(z_{j,s}\right) \le \frac{k_2}{S} \right) \tag{15}$$

The upper and lower tail dependence coefficients for each pair of commodities are calculated as $\hat{\lambda}_U = (1/R)\sum_{r=1}^{R} \hat{\lambda}_U^r$ and $\hat{\lambda}_L = (1/R)\sum_{r=1}^{R} \hat{\lambda}_L^r$, respectively. Confidence intervals for $\hat{\lambda}_U$ and $\hat{\lambda}_L$ are then constructed by computing the associated percentiles of their empirical distributions.

4. EMPIRICAL RESULTS

This section first reports the estimation results for the marginal distributions for the individual commodity returns. The section then proceeds to present the R-vine copula estimation results. Finally, we discuss the tail dependence results.

4.1. Marginal Distribution Estimation

Table 2 presents the parameter estimates and standard errors of the selected marginal distribution models for the whole sample period (Panel A), first sub-period (Panel B), and second sub-period (Panel C). For all sample periods, a GARCH(1,1) model with a constant unconditional mean is selected for all commodity return series except for the ethanol return series. In other words, the mean of these series is simply characterized by a constant. For the ethanol return series, the AR(1)-GARCH(1,1) model is chosen for both the whole sample period and the second sub-period, whereas the ARMA(1,1)-GARCH(1,1) model is selected for the first sub-period. This implies that it is necessary to include at least the autoregressive part to capture the strong serial correlation in ethanol return series (Table 1).

Table 2: Results for the marginal distributions

Commodity	Corn	Soybean	Wheat	Ethanol	Oil
Panel A: June 2, 2006 to June 30, 2016					
μ_i	−0.000081	0.000673	−0.000700	0.000899	0.000061
	(0.000346)	(0.000268)	(0.000472)	(0.000354)	(0.000346)
ϕi				0.122495	
				(0.021330)	
ω_i	0.000005	0.000003	0.000003	0.000010	0.000004
	(0.000004)	(0.000002)	(0.000012)	(0.000001)	(0.000003)
α_i	0.057551	0.069248	0.051514	0.097183	0.072531
	(0.008766)	(0.012372)	(0.056051)	(0.007071)	(0.016776)
β_i	0.929596	0.921669	0.943409	0.874143	0.923088
	(0.010211)	(0.013564)	(0.063835)	(0.009700)	(0.018226)
Panel B: June 2, 2006 to June 16, 2011					
μ_i	0.000758	0.000999	-0.000272	0.001240	0.000634
	(0.000599)	(0.000437)	(0.000662)	(0.000642)	(0.000568)
ϕ_i				0.819652	
				(0.095655)	
θ_i				−0.743033	
				(0.111961)	
ω_i	0.000018	0.000003	0.000030	0.000010	0.000010
	(0.000013)	(0.000003)	(0.000014)	(0.000002)	(0.000004)
α_i	0.052156	0.070081	0.067055	0.085517	0.069174
	(0.021722)	(0.016065)	(0.018115)	(0.009920)	(0.004581)
β_i	0.911599	0.923528	0.885697	0.886500	0.914133
	(0.044301)	(0.017053)	(0.036075)	(0.013089)	(0.011326)
Panel C: June 17, 2011-June 30, 2016					
μ_i	−0.000574	0.000507	−0.001067	0.000579	−0.000287
	(0.000409)	(0.000340)	(0.000458)	(0.000496)	(0.000431)
ϕ_i				0.156408	
				(0.029782)	
ω_i	0.000009	0.000005	0.000004	0.000010	0.000003
	(0.000001)	(0.000002)	(0.000003)	(0.000001)	(0.000004)
α_i	0.079758	0.070717	0.059687	0.105519	0.082122
	(0.006739)	(0.007999)	(0.016452)	(0.009609)	(0.029272)
βi	0.890516	0.901056	0.929859	0.864721	0.915080
	(0.011189)	(0.011742)	(0.019396)	(0.013999)	(0.030061)

μ_i, ϕ_i, θ_i, ω_i, α_i, and β_i are the parameters of an ARMA-GARCH model (refer to equations (1)-(3) in Section 3.1). Figures in parentheses are standard errors of the coefficient estimates

It is crucial that the marginal distribution models are well specified as marginal distribution misspecification can result in copula misspecification (Fermanian and Scaillet, 2005; Patton, 2006). Hence, we apply several goodness-of-fit tests to confirm the adequacy of the chosen marginal distribution models. These tests include the Ljung-Box tests of lack of autocorrelation in the standardized residuals and the squared standardized residuals, the Engle's (1982) Lagrange Multiplier (LM) test of lack of the ARCH effect in the standardized residuals, the LM tests of serial independence (Patton, 2006) of the first four moments of transformed standardized residuals or copula data, and the Kolmogorov-Smirnov test of uniformity of the copula data. The

p-values of these tests are reported in Table 3. All selected model pass all the tests at the 5% significance level, confirming that the marginal distribution models are appropriately specified.

4.2. Regular Vine Copula Estimation

The selected R-vine copula structure for the whole and two sub-sample periods is presented in Figure 1, and the estimated parameters of the corresponding bivariate copulas are given in Table 4. While different pair-copula families (i.e., dependence patterns) are chosen for the two sub-sample periods, the R-vine structure – the connection structure between agricultural and energy commodity markets – remains the same during and after the period of rapid growth of U.S. ethanol production[6]. In particular, we find that the ethanol market has established a link between the corn and crude oil markets (see the first tree in Figure 1). This result is consistent with the findings of Tyner (2010) who uses price correlations between (1) corn and crude oil, and (2) corn and ethanol during different time periods to show that the prices of corn and crude oil are connected through the ethanol market. The interaction between corn and crude oil markets through the ethanol market is likely explained by the increased use of corn as ethanol feedstock induced by the Renewable Fuel Standard (RFS) mandate (Schnepf and Yacobucci, 2013).

As can be seen from Table 4, almost all the parameters of the conditional and unconditional bivariate copulas are statistically significant at the 5% level. The only exception is the parameter of the conditional pair-copula $c_{Wheat}, Oil|Corn, Soybean, Ethanol$ for the case of the whole sample period. Given the selected unconditional copulas, we find that dependence patterns between the returns of (1) corn and soybean, (2) corn and wheat, and (3) corn and ethanol are all captured by two-parameter copula families. These results indicate that there are strong co-movements between corn and these commodity markets during both extreme market downturns and upturns. In particular, the heavy-tailed Student's t copula is chosen for the three pairs of commodity returns during the period of rapid growth in U.S. ethanol production (June 2, 2006-June 16, 2011). This implies that the degree of tail dependence is the same in both the upper and lower tails for these commodity pairs. During the period of slowing growth in ethanol production (June 17, 2011-June 30, 2016), the dependence pattern between corn and wheat is still best characterized by the Student's t copula. However, the BB1 and Rotated BB1 (180°; "Survival BB1") copulas are selected for the corn-soybean and corn-ethanol pairs, respectively. Given the estimated parameters, the upper (lower) tail appears to be somewhat heavier (lighter) than the lower (upper) tail for the corn-soybean (corn-ethanol) pair during the second sub-period.[7]

For the unconditional dependence patterns between ethanol and crude oil, the Rotated Gumbel (180°; "Survival Gumbel") copula is selected for the first sub-period whereas the Gaussian copula is chosen for the second sub-period. This implies that, during

6 Recall that the R-vine structure is selected based on maximum spanning trees with the absolute values of pairwise Kendall's tau coefficients as weights.

7 Refer to Example 5.1 in Joe and Hu (1996) for the relationship between copula parameters and tail dependence coefficients of two-parameter families of Archimedean copulas.

Table 3: Tests of the marginal distribution specifications

Commodity	Corn	Soybean	Wheat	Ethanol	Oil
Panel A: June 2, 2006-June 30, 2016					
LBQ (10) on standardized residuals	0.4082	0.8349	0.6927	0.0743	0.8354
LBQ (10) on squared standardized residuals	0.4045	0.8970	0.7161	0.5112	0.4621
LM on squared standardized residuals	0.4674	0.9999	0.9988	0.8637	0.9999
1st moment LM test on transformed standardized residuals	0.9671	0.9985	0.9875	0.6815	0.9925
2nd moment LM test on transformed standardized residuals	0.8132	0.9988	0.9906	0.9604	0.9992
3rd moment LM test on transformed standardized residuals	0.9950	0.8702	0.9986	0.7620	0.9999
4th moment LM test on transformed standardized residuals	0.8224	0.9999	0.9993	0.9801	0.9996
Kolmogorov-Smirnov test	0.6822	0.7262	0.9980	0.7979	0.9999
Panel B: June 2, 2006-June 16, 2011					
LBQ (10) on standardized residuals	0.6057	0.7866	0.8894	0.2533	0.9619
LBQ (10) on squared standardized residuals	0.9428	0.9389	0.6479	0.6073	0.9208
LM on squared standardized residuals	0.9985	0.9999	0.8941	0.9715	0.9999
1st moment LM test on transformed standardized residuals	0.9814	0.9999	0.9999	0.8763	0.9987
2nd moment LM test on transformed standardized residuals	0.9996	0.9906	0.8359	0.9222	0.9831
3rd moment LM test on transformed standardized residuals	0.9987	0.9933	0.9998	0.9801	0.9999
4th moment LM test on transformed standardized residuals	0.9975	0.9979	0.9833	0.9001	0.9973
Kolmogorov-Smirnov test	0.8678	0.9077	0.9999	0.9434	0.9999
Panel C: June 17, 2011 to June 30, 2016					
LBQ (10) on standardized residuals	0.6295	0.1879	0.6947	0.3067	0.2957
LBQ (10) on squared standardized residuals	0.6066	0.9984	0.9683	0.2099	0.3586
LM on squared standardized residuals	0.7322	0.9999	0.9999	0.8232	0.9999
1st moment LM test on transformed standardized residuals	0.9481	0.9080	0.9646	0.9073	0.9593
2nd moment LM test on transformed standardized residuals	0.9506	0.9064	0.9868	0.8612	0.9985
3rd moment LM test on transformed standardized residuals	0.9457	0.6591	0.9957	0.9543	0.9808
4th moment LM test on transformed standardized residuals	0.9274	0.9978	0.9976	0.8337	0.9983
Kolmogorov-Smirnov test	0.8678	0.9077	0.9999	0.9434	0.9999

The table reports the -values from the Ljung-Box *(LBQ)* tests on standardized residuals and on squared standardized residuals, the Engle's (1982) Lagrange Multiplier (LM) tests on squared standardized residuals, the LM tests of serial independence (Patton, 2006) of the first four moments of transformed standardized residuals or copula data, and the Kolmogorov-Smirnov test of uniformity of the copula data

Table 4: Results for the regular vine copula models

Commodity pair	Pair-Copula	Para1	SE1	Para2	SE2
Panel A: June 2, 2006-June 30, 2016					
(Corn, Soybean)	t	0.601	0.013	7.035	1.150
(Corn, Wheat)	t	0.651	0.011	7.775	1.241
(Corn, Ethanol)	t	0.602	0.014	4.218	0.451
(Corn, Oil|Ethanol)	Gaussian	0.074	0.020	-	-
(Soybean, Wheat|Corn)	Survival Gumbel	1.072	0.014	-	-
(Soybean, Ethanol|Corn)	Frank	0.680	0.121	-	-
(Soybean, Oil|Corn, Ethanol)	Clayton	0.222	0.027	-	-
(Wheat, Ethanol|Corn, Soybean)	Survival Clayton	0.074	0.021	-	-
(Wheat, Oil|Corn, Soybean, Ethanol)	Frank	0.064	0.119	-	-
(Ethanol, Oil)	Gaussian	0.319	0.017	-	-
Panel B: June 2, 2006 to June 16, 2011					
(Corn, Soybean)	t	0.649	0.017	5.991	1.240
(Corn, Wheat)	t	0.646	0.016	8.437	2.170
(Corn, Ethanol)	t	0.583	0.021	3.091	0.374
(Corn, Oil|Ethanol)	Gaussian	0.137	0.027	-	-
(Soybean, Wheat|Corn)	Survival Gumbel	1.104	0.021	-	-
(Soybean, Ethanol|Corn)	Gaussian	0.117	0.027	-	-
(Soybean, Oil|Corn, Ethanol)	Frank	1.346	0.170	-	-
(Wheat, Ethanol|Corn, Soybean)	Clayton	0.081	0.032	-	-
(Wheat, Oil|Corn, Soybean, Ethanol)	Frank	0.371	0.168	-	-
(Ethanol, Oil)	Survival Gumbel	1.351	0.029	-	-
Panel C: June 17, 2011-June 30, 2016					
(Corn, Soybean)	BB1	0.240	0.056	1.409	0.044
(Corn, Wheat)	t	0.663	0.015	8.410	1.958
(Corn, Ethanol)	Survival BB1	0.260	0.058	1.547	0.050
(Corn, Oil|Ethanol)	Gumbel	1.005	0.015	-	-
(Soybean, Wheat|Corn)	Gaussian	0.080	0.028	-	-
(Soybean, Ethanol|Corn)	Gumbel	1.054	0.017	-	-
(Soybean, Oil|Corn, Ethanol)	Survival Gumbel	1.102	0.021	-	-
(Wheat, Ethanol|Corn, Soybean)	Survival Clayton	0.098	0.031	-	-
(Wheat, Oil|Corn, Soybean, Ethanol)	Rotated Gumbel (90 degrees)	1.022	0.018	-	-
(Ethanol, Oil)	Gaussian	0.220	0.026	-	-

the period of rapid expansion of ethanol production, ethanol and crude oil returns are more highly correlated in periods of market downturns than in periods of market upturns. Nonetheless, after the period of rapid growth of ethanol production, ethanol and crude oil returns seem to be somewhat independent during extreme market movements. For the other pairs of commodity returns, the (conditional) dependence patterns are all modeled with one-parameter copula families. Because the estimated parameters of these conditional bivariate copulas are difficult to interpret, we derive the easier-to-interpret unconditional estimates of tail dependence coefficients using the simulation-based method described in Section 3.3. The results from the simulation exercise are discussed in the next section.

4.3. Tail Dependence Coefficients

The upper (lower) tail dependence coefficients are reported in the upper (lower) triangular parts of the matrix in Table 5. The upper (lower) tail dependence coefficient measures the probability that we will observe a large price hike (decline) in one commodity market, given that the price of another commodity also has had increased (decreased) significantly. Based on the results for the whole sample period, we find that the upper tail dependence coefficients are statistically significant at the 5% level for only four pairs of commodities: Corn-soybean, corn-wheat, corn-ethanol, and ethanol-oil. The lower tail dependence coefficients are all statistically significant at the 5% level. This indicates that

Table 5: Upper and lower tail dependence coefficients

Commodity	Corn	Soybean	Wheat	Ethanol	Oil
Panel A: June 2, 2006-June 30, 2016					
Corn		0.2844*	0.3139*	0.3435*	0.0534
Soybean	0.2877*		0.1581	0.1704	0.0445
Wheat	0.3128*	0.2175*		0.1908	0.0411
Ethanol	0.3442*	0.1737*	0.1722*		0.0665*
Oil	0.0487*	0.1150*	0.0466*	0.0600*	
Panel B: June 2, 2006-June 16, 2011					
Corn		0.3358*	0.3035*	0.3728*	0.0520
Soybean	0.3378*		0.1840	0.2080	0.0573
Wheat	0.3023*	0.2576*		0.1764	0.0447
Ethanol	0.3748*	0.2110*	0.2008*		0.0504
Oil	0.2093*	0.1615*	0.1350*	0.3326*	
Panel C: June 17, 2011-June 30, 2016					
Corn		0.3696*	0.3153*	0.2741*	0.0349
Soybean	0.2288*		0.1900	0.2063	0.0391
Wheat	0.3176*	0.1383*		0.1772	0.0243
Ethanol	0.4393*	0.1633*	0.2123*		0.0403
Oil	0.0302	0.1111*	0.0222	0.0362	

Upper and lower tail dependence coefficients are respectively reported in the upper and lower triangular parts of the matrix, where * indicates the rejection of the null hypothesis that the respective tail dependence coefficient is equal to zero at the 5% significance level

all commodity markets are significantly correlated during extreme market downswings. During both extreme market upturns and downturns, the most highly correlated markets are the corn and ethanol markets $(\hat{\lambda}_U = 0.3435; \hat{\lambda}_L = 0.3442)$, whereas the least highly correlated markets are the wheat and crude oil markets $(\hat{\lambda}_U = 0.0411; \hat{\lambda}_L = 0.0466)$.

The results for the two sub-periods show that, during the market upturns, the corn market is significantly linked with the soybean, wheat, and ethanol markets. However, the upper tail dependence coefficients are insignificant for any other pairs of commodity markets. Comparing the upper tail dependent coefficients for the two sub-periods, we find that the degree of comovements between the returns of (1) corn and soybean and (2) corn and wheat remain relatively stable. Nonetheless, the degree of upper tail dependence between corn and ethanol returns is stronger during the first sub-period than during the second sub-period $(\hat{\lambda}_U = 0.3728 \text{ versus } \hat{\lambda}_U = 0.2741)$. This indicates that the probability of simultaneous jumps in the prices of corn and ethanol has fallen as the ethanol market became more mature. While insignificant in both sub-periods, it is worth noting that the degree of upper tail dependence between crude oil and other commodity markets becomes even weaker during the period of slowing growth in U.S. ethanol production.

In addition, the lower tail dependence results indicate that all pairs of commodity markets are significantly correlated during the market downturns for the first sub-period but not for the second sub-period. For the second sub-period, the lower tail dependence coefficients are significant at the 5% level for most commodity pairs, except for the oil-corn, oil-wheat, and oil-ethanol pairs. While the lower tail dependence coefficients between oil and soybean markets are statistically significant for both sub-periods, the degree of dependence is weaker during the second sub-period than during the first sub-period. Similar to the results for the upper tail dependence, these findings suggest that the lower tail dependence between crude oil and other commodity markets starts to disappear in the recent years. Furthermore, we find that crude oil and other commodity returns are more dependent during extreme market downturns than during extreme market upturns for the first sub-period. However, during the second sub-period, we find neither asymmetric nor tail dependence between crude oil and most commodity markets (namely, corn, wheat, and ethanol). This empirical evidence regarding the change in the link between crude oil and agricultural commodity markets may be explained by the recent stability and slight drawdowns in ethanol production in the United States.

5. CONCLUSIONS

In this paper, we analyze the dependence structure and tail dependence patterns among returns of three agricultural commodity futures (corn, soybean, and wheat futures) and two energy commodity futures (ethanol and crude oil futures) from June 2, 2006 to June 30, 2016. Based on the results for the whole sample period, we find that the returns of corn and crude oil are linked through the ethanol market, and this is likely explained by the increased demand for corn as an ethanol feedstock. In addition, our empirical results indicate that crude oil and agricultural commodity prices are statistically dependent during the extreme market downturns but independent during the extreme market upturns. This evidence is consistent with Reboredo (2012) who reports that oil and agricultural commodity prices tend to move independently during market upswings.

We also examine whether and how the dependence structure and the degree of tail dependence evolve over the two periods of ethanol production: The rapid growth period (June 2, 2006 to June 16, 2011) and the slowing growth period (June 17, 2011 to June 30, 2016). Based on our sub-sample analysis, we uncover several interesting results. First, crude oil and agricultural commodity markets are connected through the ethanol market during both sub-periods. Second, the connection between the corn and ethanol markets during the extreme market upturns is stronger in the period of rapid growth than the period of slowing growth. Third, during the extreme market upswings, the prices of crude oil and other four commodities tend to move more independently in the slowing growth period. Forth, all commodity prices are likely to move together when markets experience downward movements in the first sub-period, but not in the second sub-period. In particular, in the second sub-period, the lower tail dependence coefficients are statistically significant for most commodity pairs, except for the oil-corn, oil-wheat, and oil-ethanol pairs. Finally, the lower tail dependence between crude oil and other commodity markets starts to disappear in the recent years when the ethanol market became more mature. Our findings regarding the change in the degree of connectedness between crude oil and agricultural commodity markets during the extreme market upturns and downturns should provide useful information for practitioners, academics and policy makers.

REFERENCES

Baumeister, C., Kilian, L. (2014), Do oil price increases cause higher food prices? Economic Policy, 29(80), 691-747.

Bedford, T., Cooke, R.M. (2001), Probability density decomposition for conditionally dependent random variables modeled by vines. Annals of Mathematics and Artificial Intelligence, 32(1), 245-268.

Bedford, T., Cooke, R.M. (2002), Vines: A new graphical model for dependent random variables. Annals of Statistics, 30(4), 1031-1068.

Campiche, J.L., Bryant, H.L., Richardson, J.W., Outlaw, J.L. (2007), Examining the Evolving Correspondence between Petroleum Prices and Agricultural Commodity Prices. Portland, OR: Proceeding of the AAEA Meeting.

Dißmann, J., Brechmann, E.C., Czado, C., Kurowicka, D. (2013), selecting and estimating regular vine copulae and application to financial returns. Computational Statistics and Data Analysis, 59, 52-69.

Du, X., McPhail, L.L. (2012), Inside the black box: The price linkage and transmission between energy and agricultural markets. Energy Journal, 33(2), 171-194.

Engle, R.F. (1982), Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation. Econometrica, 50, 987-1007.

Fermanian, J.D., Scaillet, O. (2005), Some statistical pitfalls in copula modeling for financial applications. In: Edith, K., editor. Capital Formation, Governance and Banking. New York: Nova Science Publishers. p59-74.

Gilbert, C.L. (2010), How to understand high food prices. Journal of Agricultural Economics, 61(2), 398-425.

Han, L., Zhou, Y., Yin, L. (2015), Exogenous impacts on the links between energy and agricultural commodity markets. Energy Economics, 49, 350-358.

Hertel, T.W., Beckman, J. (2012), Commodity Price Volatility in the Biofuel Era: An Examination of the Linkage Between Energy and Agricultural Markets. In: Zivin, J.S., Perloff, J.M., edtiors. The Intended and Unintended Effects of U.S. Agricultural and Biotechnology Policies. Chicago: University of Chicago Press for NBER. p189-221. Available from: http://www.nber.org/papers/w16824.

Joe, H. (1996), Families of M-variate distributions with given margins and m (m-1)/2 bivariate dependence parameters. Lecture Notes-Monograph Series, 28, 120-141.

Joe, H., Hu, T. (1996), Multivariate distributions from mixtures of max-infinitely divisible distributions. Journal of Multivariate Analysis, 57(2), 240-265.

Kristoufek, L., Janda, K., Zilberman, D. (2012). Correlations between biofuels and related commodities before and during the food crisis: A taxonomy perspective. Energy Economics, 34(5), 1380-1391.

Kurowicka, D., Cooke, R.M. (2006), Uncertainty Analysis with High Dimensional Dependence Modelling. Sussex: John Wiley and Sons. Available from: https://www.wiley.com/en-us/Uncertainty+Analysis+with+High+Dimensional+Dependence+Modelling-p-9780470863060.

Loaiza, M., Albeiro, R., Gomez-Gonzalez, J.E., Melo Velandia, L.F. (2015), Latin American exchange rate dependencies: A regular vine copula approach. Contemporary Economic Policy, 33(3), 535-549.

Lucotte, Y. (2016), Co-movements between crude oil and food prices: A post-commodity boom perspective. Economics Letters, 147, 142-147.

Morales-Nápoles, O., Cooke, R.M., Kurowicka, D. (2010), About the Number of Vines and Regular Vines on n Nodes. Working Paper. Delft University of Technology.

Muhammad, A., Kebede, E. (2009), the emergence of an agro-energy sector: Is agriculture importing instability from the oil sector? Choices, 24(1), 12-15.

Patton, A.J. (2006), Modelling asymmetric exchange rate dependence. International Economic Review, 47(2), 527-556. Available from: https://www.jstor.org/stable/3663514?seq=1#page_scan_tab_contents.

Reboredo, J.C. (2012), Do food and oil prices co-move?. Energy Policy, 49, 456-467.

Schnepf, R., Yacobucci, B.D. (2013), Renewable Fuel Standard (RFS): Overview and Issues. CRS Report for Congress 7-5700, Congressional Research Service. Washington DC.

Serra, T., Zilberman, D. (2013), Biofuel-related price transmission literature: A review. Energy Economics, 37, 141-151.

Sklar, A. (1959), Fonctions de Répartition à n Dimensions et Leurs Marges." Publications de l' Institut Statistique de l'Université de Paris, 8, 229-231.

Tyner, W.E. (2010), The integration of energy and agricultural markets. Agricultural Economics, 41(1), 193-201.

U.S. Energy Information Administration. (2011), Growth Slows in U.S. Ethanol Production and Consumption". Today in Energy (September 14, 2011), Washington DC. Available from: https://www.eia.gov/todayinenergy/detail.php?id=3070. [Last accessed on 2017 Feb 12].

Zilberman, D., Hochman, G., Rajagopal, D., Sexton, D., Timilsina, G. (2013), The impact of biofuels on commodity food prices: Assessment of findings. American Journal of Agricultural Economics, 95(2), 275-281.