

# Evolving Relationship between Oil and Corn under Structural Changes

Final Honors Thesis

Zijun Tian

March 06, 2017

Economics Department

University of North Carolina at Chapel Hill

Thesis Advisor: Dr. Peter Hansen

Faculty Advisor: Dr. Klara Peter

Approved:

---

Dr. Peter Hansen

## Abstract

This paper focuses on the better estimation of the correlation between oil and corn with Realized Beta GARCH model, the first moment Granger Causality test on the prices of oil and corn, and the second moment causality test on the volatilities and returns of oil and corn. Since the past decade has witnessed the ever-fast development in biofuel production and increasing volatilities in the markets, I further perform Bai and Perron Procedure to detect possible structural breaks from 2004 to 2014 and use impulse response functions to help draw insight into the changing causal relationships in the price levels. I find that the correlation between oil and corn nearly doubled during the financial crisis and then resumes back to the normal level about 0.2. However, the statistically significant causal relationship from oil prices to corn prices detected in the prior crisis period cannot be found later as a result of higher volatilities and nonexistent arbitrage. The second moment causality test on the volatilities confirms our hypothesis that knowing the information on the returns of oil at time  $t$  can help predict the volatilities of corn at time  $t+1$ .

## **Acknowledgement**

I would like to sincerely thank my advisor, Dr. Peter Hansen, for his guidance and support throughout the entire research process, for consistently challenging and questioning me to become a better student of economics, for helping me to apply statistical theories with economic motivations, for providing me with necessary data and important tools, and for inspiring me to further pursue econometrics in my later life. I would also like to thank Dr. Klara Peter for providing the entire class with the clearly organized and well-prepared lectures, the timely help from topic selection, data cleaning, draft proposals to final thesis, and most importantly the encouragement necessary to write excellent theses. Finally, I would like to thank my classmates who both help me perfect my thesis during the discussion after every presentation and after classes. The year with my advisor, Dr. Klara Peter, and all my classmates is really enjoyable and memorable and it helps me confirm my determination to further explore and study econometrics.

## **I. Introduction**

The relationship between crude oil and agricultural commodities has changed dramatically over the past decade, increasingly affected by the development of renewable energy, and turbulent financial market. Since biofuel, substitutable for crude oil, has been more widely accepted by consumers and supported by government, we may expect evolving relationship between oil and agricultural commodities, especially for those used in the production of biofuel. Based on the primary and wide use of oil in the agriculture production and the most technically obtainable and mature substitution between oil and ethanol that heavily depends on the corn as the input, this paper addresses the recently accelerated trend of corn's emergence as an energy crop and the complicated oil-ethanol-corn relationship.

According to United States Department of Agriculture Economic Research Service, US has a dramatically rapid growth in the share of corn use for fuel ethanol in the last few years. Three or four years ago, the share only accounted for 12 to 14 percent of total U.S. corn use. However, under current conditions, ethanol share in the total corn use accounts for about one third of the total demand for corn in U.S. and is expected to further increase in the next few years with government mandates that call for increased ethanol use. With the help of increasing demand of corn in the energy market that is second in size only to the demand in the domestic feed market, the agricultural production of corn in U.S. has survived from excessive production capacity, low prices, and heavy dependence on government income supports and becomes an ever-growing sector with frequent periods of tight supplies and food price surge. On the one hand, the sizeable increase in ethanol use of corn tends to significantly strengthen corn prices. On the other hand, however, the prices and demand of corn are now more closely associated with energy market and more frequently involved in the fluctuations of oil prices, which may have major implications for various parties in

various sectors such as consumers, crop and livestock farmers, renewable energy industry, and traditional energy market.

Therefore, the degree of stability and strength of the relationship between oil and corn plays an important role in appropriate risk-management strategies as well as an insight into the consumption behaviors and investment plans. Four major research papers have explored the correlation between oil and corn. Tang and Xiong (2012) analyzed the increasing correlation between oil price and other non-energy commodity prices by focusing on the index investment and financialization in commodity markets. Along the similar line, Thorp (2015) also argued that the correlation between oil and most agricultural commodities returns has been markedly increased and become more volatile and conducted impulse response function analysis among crude oil and three major agricultural commodities including corn, wheat, soybeans. In addition to an integrated partial equilibrium framework that showed the impact of ethanol promotion on the correlation between oil and corn (Tyner and Farzad, 2008), more advanced time series econometrics may draw new insights into this relationship. By applying both discrete and smooth threshold error correction models, Merkusheva and Rapsomanikis (2012) explored the price relationships in the food-ethanol-oil nexus and found that oil prices are the long run drivers of ethanol and grains prices. In this research paper, I aim to first offer a more accurate estimation of the correlation between oil and corn based on the Realized Beta GARCH model (Hansen, Lunde, and Voev, 2013), which enables us to have a contemporaneous sense of the performance of one commodity in the market with the information of the other. Moreover, I also expand upon previous works by addressing the first moment Granger causality between oil prices and corn prices under structural breaks, and the second moment causal relationship between volatilities and returns of oil and corn from 2004 to 2014. The analysis of causality, different from that of correlation in the sense of

having less attention and exploration, can provide us with a step-ahead forecast of the prices and volatilities of oil and corn. In order to draw more insight into the causal relationship between oil and corn and explain certain phenomena happening in more economically disturbed periods that we cannot easily interpret based only on Granger causality test results, I also conduct an analysis of impulse responses to examine how the change in the percentage change of prices of each commodity in the system responds to one standard deviation shock of the other. Finally, I also explore possible explanations for the observed changes in the correlation and causalities between oil and corn corresponding to the detected structural breaks including policy, agricultural, and financial causes. Due to the more volatile financial market after financial crisis in 2008, I hypothesize that the correlation between oil and corn significantly increases afterwards. Additionally, based on the more exogenous supply of oil and the high volatility of its prices, I may expect that the first moment causality of price levels goes from oil to corn. However, it is highly possible that the causality is spurious and the market prices of these two commodities are not causally related since both prices are volatile and not very predictable. Under this consideration, I may rather hypothesize that the causal relationship from the returns of oil to the volatilities of corn is significant, corresponding to Merkusheva and Rapsomanikis' analysis using error correction models (2012).

Based on my results, my initial hypotheses are partially confirmed and further enriched. The correlation between oil and corn nearly doubled in 2008 but then started to decline to its original level around 0.2 in 2011. No significant causalities in prices are detected under full dataset. However, oil prices are shown to Granger cause corn prices prior to April 2008. It seems like the financial crisis disturbs the financial market as well as the causal relationship between oil and corn. However, the causality in price levels does not resume back to the prior

crisis period even after financial crisis, which corresponds to the lack of arbitrage in the financial market. With regards to the corn price shocks presented in Figure 5 b), 6 b), 7 b), 8 b), the initial response of the percentage change of oil prices is not significant under the full dataset as well as in each regime, corresponding to no significant causal relationship from corn prices to oil prices. The abnormality of Figure 9 b) will be explored in details later. However, Figure 5 c) and 6 c) confirms the inverse causality since the initial response of the percentage of change of corn prices to the oil price shocks is positive while significant and the latter, or the first sub-period, has larger magnitude and smaller variances. If we look at the causality from the perspective of volatilities and returns where nonstationarity in the prices are eliminated, we can detect a significant causal relationship that the returns of oil at time  $t-1$  can help predict the volatilities of corn at time  $t$  at significance level 0.05.

This paper unfolds as follows. Section 2 presents a literature review of the past theoretical techniques and empirical analysis. It also contains the potential contribution of my research based upon its extension from the previous works. Section 3 introduces the basic theoretical model I plan to use. Empirical model and general methodology are outlined in Section 4. Section 5 contains data selection and modification, and we discuss our results and findings in section 6, followed by conclusions in Section 7. All tables and figures are collected in the appendix after Section 7.

## **II. Literature Review**

Over the past few years, there have been a number of theoretical and empirical analyses on the correlations between market returns and asset returns, the changing magnitude of the correlation between oil and corn specifically, and the factors that can impact this correlation. Firstly, Hansen, Lunde, and Voev (2013) have looked into the improvement of conventional

GARCH model by incorporating realized measures and leverage functions. In order to better estimate the dynamics between the market return and an individual asset return, they proposed Realized Beta GARCH model that extracts information from realized variances and covariances to calculate conditional correlation. Secondly, Tang and Xiong (2012) analyzed the increasing correlation between oil price and other non-energy commodity prices by focusing on the contemporaneous increasing trend of index investment in commodity markets. In particular, they demonstrated that the recent higher correlation between oil price and other indexed commodities prices are concurrent with more investment in these commodities. Thorp (2015) also argued that the correlation between oil and most agricultural commodities returns has been markedly increased and become more volatile. She admitted the existence of structural changes in conditional correlation that can occur at the time of the introduction of biofuel policies. Additionally, based on her analysis of impulse responses, Thorp noted that the adjustments to the oil price shocks is stronger for biofuel feedstocks, like corn, and negligible for many other agricultural commodities, which suggests that the substitutability between biofuel and oil, and the reliance of the production of biofuel on agricultural commodities play an important role in driving the binding relations. With one step ahead, Merkusheva and Rapsomanikis (2012) analyzed the nonlinear linkage between oil, ethanol and other grains prices in the short run. In the long run, as they suggested, oil prices are the drivers of ethanol and grains prices, and they adjust to the small deviations of oil price very quickly.

However, while most of the literatures have been focused on the correlation between oil and corn, the models by which they calculate the correlation are less accurate and less information-inclusive than the Realized Beta GARCH model. Therefore, I plan to extend Hansen, Lunde, and Voev's model of market return and one asset return by using two asset



returns that are the return of oil and the return of corn to calculate their correlation.

Additionally, the economic analysis on the causality between oil price and corn price is much more scarce and usually absent, not to mention the causality between the volatilities and returns of oil and corn. Without being aware of the causal direction between the prices of these two goods, policy makers may not be able to determine what the impact of biofuel policies on the oil market might be or whether these policies can bring potentially adverse fluctuations in the oil market. Without having sensible and relatively accurate prediction of the range of changes in which these two goods' prices can bring, governments may not be able to adopt an effective risk management approach. Firms may only have limited planning horizons, leading to investments postponement and expensive reallocation of resources. Economists also need forecasts on macroeconomic trends and risk exposure of energy and agricultural commodity markets. For example, it is crucial for them to know if an oil shock will influence biofuel market at first, the demand for corn subsequently, and the price of food that is relevant to each individual consumer eventually.

### **III. Theoretical Model**

Under the assumption of competitive retail energy and agricultural commodities markets, my theoretical models are the basic economic models of supply and demand, and substitutes and complements proposed by Alfred Marshall in his work *Principles of Economics* in 1890. Since the number of sellers is sufficiently large and there are generally no barriers for new firms to enter the market, we can reasonably assume that the markets are perfectly competitive and the firms are generally price takers. These models offer channels that the demand and supply of corn and oil can interact with each other, thus leading to co-movement of their prices and volatility spillover. They raise uncertainty and questions toward causalities between these two commodities because we can apparently observe their concurrent price

and volatility changes but cannot easily determine the underlying sequential dynamics between them. There is no denying that the casual relationship between oil and corn can take many forms.

Modern agriculture uses oil products to fuel farm machinery, to transport raw input to the farm, and to transport farm output to the ultimate consumer. Additionally, oil is often used as input in production of agricultural chemicals. Therefore, oil price increases put pressure on the supply of agricultural commodities, like corn. Decreasing supply of corn drives up its price in the market, and thus reflects causality from oil price to corn price. This causal direction may also build upon the relationship between corn and ethanol that is the most common type of biofuel. Since biofuels are the only non-fossil liquid fuels able to replace petroleum products in existing combustion engines and motor vehicles, (Thorp, 2015), as oil prices rise, the demand for these alternative fuels will increase, leading to an increase in the demand of corn that is the major feedstock for ethanol. Therefore, corn prices are forced upward. However, the above causality can be reversed. For example, the technological and government mandated expansion of ethanol industry is believed to have contributed to the recent increase in commodity prices. As the price of corn increases, marginal cost of biofuel production also rises. If we assume the market of biofuel is perfectly competitive, then the increase in the biofuel price makes consumers substitute away from now relatively more expansive energy source back to oil. As the demand of oil increases, the price of oil increases correspondingly. Moreover, the causality can also be bidirectional. Some shocks in the macroeconomic environment may impact both prices. One notable instance is the financial crisis in 2008. As a result of wealth effect, when people have less purchasing power, they consume less. Thus, the demand for both energy sources may decrease, leading to simultaneous changes in their prices, though in the same direction. It may also be interesting

to note that the causality can be asymmetric depending on the current prices of oil and corn, and their respective change rate. In Figure 1, we can observe both similar moving direction and discrepancies between the oil and corn prices. However, when both corn and oil have high prices or their prices are being driven up, their price co-movement is relatively stronger and more obvious.

Additionally, there are also various possibilities and ambiguities in the causal relationship between the volatilities and returns of oil and corn due to the substitutability between and the nature of perfectly competitive market of these two goods. Actually, driven by changes in uncertainty related to supply, demand, and inventories, changes in volatility are extremely crucial for speculation and market stability if we do not assume a rational expectations world. Volatility often reflects the amount of market uncertainty or risk about the size of changes in a commodity's value. A higher volatility means that a commodity's value can potentially be spread out over a larger range of values, resulting from changes in market demand, the release of new economic or supply information, changes in consumers' expectations, changes in market tastes, or unanticipated events and circumstances that can cause large price adjustments. Since the changes in market supply and demand of corn can affect those of oil and vice versa, it is hard to determine whether knowing the changes of price of corn can help predict the volatility of oil beforehand or the causality should be the other way around. On the one hand, along with more necessity and awareness of resource scarcity and environmental protection, more consumers start to increase their willingness to use biofuel instead of oil and substitute away from oil. The change in consumers' taste can drive up the demand for ethanol as well as for corn. It increases corn price as a result and making consumers substitute toward oil instead. Thus, the more fluctuated and less uniform change of the demand for oil increases its volatilities. On the other hand, since the supply of oil is

relatively more exogenous, supply shocks in oil can lead to changes in the demand of its substitutes, biofuel like ethanol, dramatically, thus leading to the changes in the demand of corn used in energy production. The corn volatility can be driven up even if the initial shock starts in the oil market. However, it may also turn out to be that other unanticipated events can cause the simultaneous increase in the volatilities of oil and corn since 2004. For example, the financial crisis, which initially disrupted the markets for mortgage-backed securities, negatively impacted the balance sheets of many financial institutions. Eventually it reduced the risk appetite of investors for seemingly unrelated assets in their strategic portfolio allocation to avoid more potential uncertainty. The sudden change in market tastes and demand may lead to the contemporary high volatilities of oil and corn.

Therefore, based on the interaction and variety of the above channels provided by basic economic models of supply and demand and substitutes and complements, it is difficult to tell from rule of thumb if there exists a causal relationship that can help us predict the volatility of one commodity using the price information of the other without further analysis. If there does exist such a way, then I aim to find out what direction and to what level of that causal relationship actually is.

## **IV. Empirical Model**

### *4.1 Bivariate Realized Beta GARCH model*

The ordinary least squares model assumes that the expected value of all squared error terms is the same at any given point. However, since the accuracy of the predictions of financial asset is important for portfolio selection, risk analysis, and derivative pricing, we need to reasonably admit the existence of heteroskedasticity, we and want to know what makes the

variance of the error terms larger for some ranges of data than for others. Therefore, we need models, like GARCH (generalized autoregressive conditional heteroskedasticity), which are capable of treating the undesired heteroskedasticity as the variance to be modeled instead of trying to avoid and fix it as what the OLS model does. From this point view, even if “robust standard errors” (Engle, 2001) makes the concern over heteroskedasticity less problematic, our requirement for model accuracy and for the insight of the variance changes of the error terms make the basic version of the least squares model relatively inapplicable and biased. Thus, I use Bivariate Realized Beta GARCH model, an improved version of GARCH model, which incorporates current levels of volatilities and correlation to calculate both the conditional variance and conditional correlation between oil and corn of a ten-year timespan. Lag one is used for both assets because it is the common way used in previous papers. We have five observable variables in the model and we define the information set to be  $(r_{oil,t}, r_{corn,t}, x_{oil,t}, x_{corn,t}, y_{oil,corn,t})$  where  $r_{oil,t}$  is the realized return on oil,  $r_{corn,t}$  the realized return on corn,  $x_{oil,t}$  the realized measure of oil volatility,  $x_{corn,t}$  the realized measure of corn volatility, and  $y_{oil,corn,t}$  the realized correlation between oil and corn. Motivated by findings in the research of Hansen, Lunde, and Voev (2013), we define  $z_{oil,t} \sim i. i. d. N(0,1)$  and  $z_{corn,t} \sim i. i. d. N(0,1)$ . Furthermore, as it states in Andersen et al. (2001b), returns standardized by realized volatility are approximately normally distributed. Therefore, if we define  $h_{oil,t}$  and  $h_{corn,t}$  to be conditional variances of oil and corn respectively at time  $t$ , then we have  $r_{oil,t} = \mu_{oil} + \sqrt{h_{oil,t}} z_{oil,t}$  and  $r_{corn,t} = \mu_{corn} + \sqrt{h_{corn,t}} z_{corn,t}$  that follow our assumptions mentioned previously. If we look at the above equations from a regression perspective, then the realized returns of oil and corn are our dependent variables and  $h_{oil,t}$  and  $h_{corn,t}$  can be interpreted as the variances of the residuals. Then, we need to specify how the information is used to forecast the variance of the return and the correlation between oil and corn, conditional on the past information set. Based on

the basic structure of bivariate realized Beta GARCH model that consist of GARCH equation and measurement equation for each individual asset, specifically oil and corn here, we define the model as follows:

$$\log h_{oil,t} = a_{oil} + b_{oil} \log h_{oil,t-1} + c_{oil} \log x_{oil,t-1} + \tau_{oil}(z_{oil,t-1}) \quad (1)$$

$$\log x_{oil,t} = \xi_{oil} + \varphi_{oil} \log h_{oil,t} + \delta_{oil}(z_{oil,t}) + u_{oil,t} \quad (2)$$

$$\log h_{corn,t} = a_{corn} + b_{corn} \log h_{corn,t-1} + c_{corn} \log x_{corn,t-1} + d_{corn} \log h_{oil,t} + \tau_{corn}(z_{corn,t-1}) \quad (3)$$

$$\log x_{corn,t} = \xi_{corn} + \varphi_{corn} \log h_{corn,t} + \delta_{corn}(z_{corn,t}) + u_{corn,t} \quad (4)$$

According to Andersen et al. (2001a, 2001b, 2003), we assign Gaussian specification for  $u_{oil,t}$  as well as  $u_{corn,t}$  such that  $u_{oil,t} \sim i. i. d. N(0, \sigma_{oil}^2)$  and  $u_{corn,t} \sim i. i. d. N(0, \sigma_{corn}^2)$ .

$\tau_{oil}(z_{oil,t-1})$  and  $\tau_{corn}(z_{corn,t-1})$  are leverage functions that model the leverage effect of the dependence between an asset's return and its changes of volatility. Hansen, Lunde, and Voev justified in their research (2014) that they are basically second-order polynomials. Finally, the equation for the dynamic modeling of conditional correlation between oil and corn is given by

$$F(y_{oil,corn,t}) = \xi_{oil,corn} + \varphi_{oil,corn} F(\rho_{oil,corn,t}) + v_{oil,corn,t} \quad (5)$$

$F(\rho)$  is the Fisher transformation that maps the correlation originally ranging from  $-1$  to  $1$  to the whole real number line, which spreads out the distribution of correlation and enables us to observe its changes over time more easily and obviously. Additionally, it also makes  $v_{oil,corn,t} \sim N(0, \sigma_{oil,corn}^2)$  a reasonable assumption. Using log likelihood function and substituting  $h_t$  for variance in the normal likelihood, we can have a systematic way to adjust all parameters simultaneously and give the “best fit” estimation. With all the parameters estimated and known, we can construct time series of conditional variances of oil and corn,

and the conditional correlation between them. The preliminary estimation results are shown in Figure 2, 3, and 4. Based on the fact that  $z_{oil,t} \sim i. i. d. N(0,1)$ ,  $z_{corn,t} \sim i. i. d. N(0,1)$ , and the basic property of bivariate normal distribution, we have

$$E[z_{oil,t}|z_{corn,t}] = \rho_{oil,corn,t} \times z_{corn,t} \quad (6)$$

Therefore, having better estimation of the correlation and knowing  $z_{corn,t}$  which can be obtained from the above model, we can know the expectation of the conditional value  $z_{oil,t}$ . Based upon that, we can further get a contemporaneous intuition on the volatility, or the risk, of oil.

#### *4.2 Bai and Perron procedure*

For the bivariate realized Beta GARCH model established for the volatility and correlation mentioned above, it describes a dynamic process that counts for continuous changes in the market and economic environment. Since domestic policies, international trade pattern, and other supply and demand shocks can influence the relationship between oil price and corn price, it may also be appropriate to model the prices in the same way that includes changes over time. However, currently, their higher volatility and much more unpredictability prohibit us from doing so. Therefore, in order to incorporate underlying changes as well as model much as possible, we rather try to find discrete changes for oil and corn prices, analyzing different regimes instead of full dataset under the assumption of no structural changes. Even if prices are not very likely to jump from one pattern to another, it is still meaningful for us to do so because they allow us to see which event coincides with which breaks and which kind of events, political or economic, internal or external, are more likely to cause huge changes in the prices of oil and corn. The methods of detecting and estimating structural breaks are diverse and have evolved toward more complexity and more applicability. The classic Chow

Test can only test for one structural change with known break point. It simply tests for whether the coefficients before and after the break are the same, and we can conclude that the given break does exist if we have sufficient evidence to statistically significantly reject the null hypothesis. Then Andrews proposed SupW tests (Andrews, 1993) that allows for detection of a single yet unknown break point within a given time interval of all candidate break dates. Its limitation on the requirement that all the regressors in the ordinary least squares regression must be strictly stationary is later expanded by Fixed Regressor Bootstrap procedure. However, since we can neither tell the number of breaks for sure nor arbitrarily assign an exact number of breaks to the oil and corn prices from 2004 to 2014, we use more advanced Bai and Perron procedure (2003) with full structural model where all parameters are subject to shifts. It is used to detect multiple endogenous structural changes and help split data correspondingly. Basically, this procedure is an iterative method using OLS and allows for general forms of serial correlation and heteroskedasticity in the errors. It is also less restrictive on the distributions of independent variables and errors. They can have trend or even different distributions across segments. Consider the following regression with  $n$  breaks or, namely,  $(n + 1)$  regimes:

$$y_t = z_t' \alpha_j + v_t; t = T_{j-1} + 1, \dots, T_j; j = 1, 2, \dots, n + 1; T_0 = 0, T_j = T \quad (7)$$

Let  $y$  denote corn prices and  $z$  denote regressors such as oil price, U.S. ethanol price, inflation rate as a proxy for general market condition, cross-price elasticity of demand between oil and corn that approximates their substitutability, ethanol share in the total use of corn, and wheat price with  $\alpha_j$  representing the regime-specific vector of coefficients. We have  $v_t$  to be the disturbance. The method of estimation, according to Bai and Perron, is based on the least-squares principle. For each  $n$ -partition  $(T_1, T_2, \dots, T_n)$ , denoted  $\{T_j\}$ , we estimate the parameter vector  $\alpha_j$  by minimizing sum of squared residuals:



$$\sum_{j=1}^{n+1} \sum_{t=T_{j-1}+1}^{T_j} [(y_t - z_t' a_j)]^2 \quad (8)$$

over all possible partitions of the whole timespan with a maximum number of breaks  $u$  and minimum regime length  $d$ . In other words, we should satisfy  $j \leq u$  and  $T_j - T_{j-1} \geq d$ . Then let  $\hat{\alpha}(\{T_j\})$  denote the estimated coefficients and  $S_T(T_1, T_2, \dots, T_n)$  the resulting sum of squared residuals. Substituting  $\hat{\alpha}(\{T_j\})$  into the objective function of least-squares, we can obtain the estimated break points  $(\hat{T}_1, \hat{T}_2, \dots, \hat{T}_n)$  such that

$$(\hat{T}_1, \hat{T}_2, \dots, \hat{T}_n) = \operatorname{argmin}_{T_1, T_2, \dots, T_n} S_T(T_1, T_2, \dots, T_n) \quad (9)$$

After obtaining the number and the position of the breaks, we can do separate tests on different regimes and see which events coincide with which breaks.

#### 4.3.1 Granger Causality Test for Price

Granger causality tests on whether oil price is able to increase the accuracy of the prediction of corn price with respect to a forecast based on only past values of corn price and vice versa. It is based on two properties of causality. First, the future cannot cause the past. Second, a cause contains unique information about an effect that is not available or replicable elsewhere. Therefore, we can establish the following inequality:

$$F(X_{t+h} | \Omega_t) \neq F(X_{t+h} | \Omega_t - Y_t) \quad (10)$$

where  $F$  is the conditional distribution;  $\Omega_t$  is the information set used to predict  $X_{t+h}$  at time  $t$ ;  $\Omega_t - Y_t$  is all the information in the universe used to predict  $X_{t+h}$  except for series  $Y_t$ . If the above relation holds, then  $Y_t$  is said to “granger cause”  $X_t$  since it can help predict future  $X$ . For the simplicity of use and interpretation, I establish the vector autoregressive

(VAR) model and first consider bivariate Granger Causality test where oil price and corn price are the only two variables included in the test, which can be expressed as following:

$$P_{oil,t} = \alpha_t + \sum_{i=1}^m \beta_i P_{oil,t-i} + \sum_{j=1}^n \gamma_j P_{corn,t-j} + \mu_t \quad (11)$$

$$P_{corn,t} = \theta_t + \sum_{i=1}^p \varphi_i P_{corn,t-p} + \sum_{j=1}^q \pi_j P_{oil,t-j} + v_t \quad (12)$$

Based on this test and particularly straightforward for the bivariate case, we can have four possible outcomes that depend on  $\sum_{j=1}^n \gamma_j$  and  $\sum_{j=1}^q \pi_j$ , which are unidirectional Granger-causality from price of oil to price of corn, unidirectional Granger-causality from price of corn to price of oil, bidirectional causality, and independence between price of oil and price of corn. For example, if  $\sum_{j=1}^q \pi_j$  is statistically significantly different from zero and  $\sum_{j=1}^n \gamma_j$  is not, then we can conclude that it is the first outcome. However, the causality may not be limited only to bivariate case since it is hard to define and measure  $\Omega_t$  in practice. The unobserved common factors are always a potential problem for any finite information set. Other factors, like the cross-price elasticity of demand between oil and corn, the fraction of total consumption of corn used for energy use, the price of wheat, and the financial market fluctuations, may also influence the causality between oil price and corn price. Therefore, many efforts have been done on extending the Granger Causality test to a multivariate case. Caines, Keng and Sethi (1981) proposed a reasonable procedure. If a process  $X_t$  has more than one causal variable, say  $Y_1, Y_2, \dots, Y_n$ , we construct bivariate VAR model with each pair of  $(X_t, Y_i)$ , for  $i = 1, 2, \dots, n$  with order selected to minimize prediction error. Then we rank these causal variables according to a decreasing order of their specific gravity, the prediction error. For each causal variable, the lag order is determined by establishing the optimal univariate AR model with minimal prediction error. Finally, we add the causal variable one at a time based on their causal rank obtained before with orders optimized at each step. The final result would be an optimal ordered multivariate AR model of  $X_t$  including all the causal

variables. However, while we respect all the changes and improvement conducted on the conventional Granger Causality Test, we do believe that simplicity is also important.

Actually, I highly doubt the existence of the first moment (price levels) causality between oil and corn, at least after around 2008 when commodities have been more financialized and more frequently traded in the financial market. Otherwise, there would be huge amount of arbitrage from which speculation can make a lot of money. It is also the reason that I perform second moment (volatility and returns) causality test as well since it may offer more stable and accurate insight on the causality between oil and corn.

#### *4.3.2 Impulse Response Function based on VAR Model*

Impulse response functions describe the dynamic response of one variable to a one-period shock from the other variable based on a VAR system that includes stationary variables or time series with time invariant expected values, variances, and covariances. Following the above Granger causality test for the price levels of oil and corn, we want to further analyze how one commodity adjusts to the exogenous shocks of the other. If the response is statistically significantly different from zero, meaning the 95% confidence interval does not contain zero response, then we can, on the one hand, conclude the unidirectional causality. On the other hand, we can get a quantitative measure of the speed of adjustment of a good to its own shock or to the other good's shock. Similarly, we consider the whole timespan as well as the four sub-periods corresponding to the three structural breaks we identified via Bai and Perron Procedure previously.

In order to assure stationarity of the time series used in the VAR model, I perform the Augmented Dickey and Fuller's (1979) ADF unit root test on the oil and corn prices, as

Merkusheva and Rapsomanikis used in their paper (2012). Since the null hypothesis under this test is nonstationarity, we should further transform the variables if we do not have sufficient evidence to reject the null hypothesis at a significance level of 0.05. Based on previous literatures and economic intuition on the price levels' random-walk behaviors, oil and corn prices are very likely to be the time series of order one. Therefore, we take the log of each of the prices and then take the first difference. Let  $dlogOil$  and  $dlogCorn$  denote the first differenced log prices of oil and corn respectively. Then we redo the ADF unit root test and stationarity can be reasonably assumed this time with p-value less than 0.05. Further following Merkusheva and Rapsomanikis's paper on the selection of optimal lag in VAR model, we apply Bayesian Information Criterion (BIC) to select the optimal lag length and ensure the residuals to be white noise. Let  $p$  denote the optimal lag. Then the VAR ( $p$ ) model is as following:

$$y_t = \Phi_0 + \sum_{i=1}^p \Phi_i y_{t-i} + \varepsilon_t \quad (13)$$

where  $y_t$  is a  $2 \times 1$  vector of approximate percentage change of oil and corn prices at time  $t$  from time  $t-1$ ,  $\Phi_0$  a  $2 \times 1$  vector of constants,  $\Phi_i$  a  $2 \times 2$  matrix relating the percentage price changes at lagged  $i$  period to current percentage changes, and  $\varepsilon_t$  a  $2 \times 1$  vector of i. i. d. vector of constants error terms. More specifically, based on the property of VAR model, each of the two variables is a function of  $p$  lags of both variables, including itself, a constant term and a contemporaneous error term. Since we use Cholesky decomposition for the orthogonalized impulse response function that is most commonly used in the previous literatures involving impulse response function analysis, the above equation can be rewritten into a vector moving average (VMA) representation as following:

$$y_t = \psi_0 + \sum_{i=1}^{\infty} \psi_i + \varepsilon_{t-1} \quad (14)$$

where  $\psi_i$  is a  $2 \times 2$  matrix with each element to be the impulse response with respect to the shock in each variable's error term.

### 4.3.3 Causality Test for Volatility Based on Bivariate Realized Beta GARCH Model

Different from Granger Causality Test for price, we extend the realized Beta GARCH model established before to test for causality in the volatility. There are two reasons for the change of model. First, it allows us to take other factors that may potentially impact the causality between oil and corn into consideration without bothering complicated multivariate test. Second, the realized Beta GARCH model provide us with the most accurate model in estimating and forecasting the dynamics between oil and corn by continuously incorporating leverage functions that can model the dependence between returns and volatility, and realized measures available in the market. Basically, we reuse the following two equations and add term  $x_{corn,t-1}$  in the first equation and  $x_{oil,t-1}$  in the second equation:

$$\log h_{oil,t} = a_{oil} + b_{oil} \log h_{oil,t-1} + c_{oil} \log x_{oil,t-1} + d_{oil} \log x_{corn,t-1} + \tau_{oil}(z_{oil,t-1}) \quad (15)$$

$$\log h_{corn,t} = a_{corn} + b_{corn} \log h_{corn,t-1} + c_{corn} \log x_{corn,t-1} + d_{corn} \log h_{oil,t} + e_{corn} \log x_{oil,t-1} + \tau_{corn}(z_{corn,t-1}) \quad (16)$$

Lag one is chosen as given according to Hansen, Lunde, and Voev's (2014). If  $e_{corn}$  can be shown against zero, then we can conclude that the returns of oil can help predict the volatility of corn one time period ahead, specifically one day ahead in this research. Namely, the size of changes associated with corn prices can be explained by the price changes of oil in the previous day, which describes the second moment causality. We can also test  $d_{oil} \log x_{corn,t-1}$  similarly in the first equation to see if it is possible that the returns of corn today can contribute to the oil volatility forecast tomorrow. However, if we do not detect any meaningful causality from the above model, it does not necessarily mean that there does not exist one. Actually, without including other variables relevant to the causality, the test may not be very comprehensive. Therefore, we also include the returns of ethanol denoted by

$x_{ethanol,t}$  that is the most typical biofuel and the returns of wheat denoted by  $x_{wheat,t}$  that is substitutable for corn in both energy production and food consumption. The reason that we use returns instead of prices is that the stationarity of the model requires stationary process. Prices are not qualified but returns can be better assumed to be stationary. Finally, we use log likelihood maximization again to estimate the coefficient vector and see if any of returns as well as other relevant variables can contribute to the prediction of the volatility. The significance of the estimated coefficients  $e_{corn}$  and  $d_{oil}$  can be obtained through the comparison between the change of log likelihood and the critical value  $\chi_{n,0.05}^2$  where  $n$  is the number of constraints in the equation. For example, if we want to test whether  $e_{corn}$ , the coefficient of  $\log x_{oil,t-1}$  in the GARCH equation of corn, is statistically significant, we first restrict  $d_{oil}$  to be 0. Then we conduct log likelihood maximization and take the difference between it and the previous log likelihood. Then we compare the change in log likelihood with the value of  $\chi_{1,0.05}^2$ . If it is larger than  $\chi_{1,0.05}^2$ , then we can reject the null hypothesis that have sufficient evidence to conclude that the returns of oil at time  $t-1$  can help predict the volatility of corn at time  $t$  at significance level 0.05.

## V. Data

My first dataset consists of daily data on the realized return, variance, and correlation of oil and corn between 2004 and 2014 from Tick Data Inc, and has been previously used in paper “A Markov Chain Estimator of Multivariate Volatility from High Frequency Data.” (Hansen, Horel, Lunde, and Archakov , 2015). It inclusively contains all the necessary information to calculate conditional variance and to better estimate the correlation between corn and oil during this period. However, because of annually short trading days around Thanks Giving and Christmas, I eliminate those days if the daily sum of trading seconds is abnormally small. The other two daily time series data used in this paper are oil price obtained from U.S.

Energy Information Administration and corn price obtained from Farm Journal. They are also adjusted according to the elimination of trading days in the previous dataset. Other relevant variables, including inflation rate represented by CPI (Consumer Price Index), and ethanol share in total use of corn, are monthly data. Without the use of mixed frequency, I can either aggregate daily data into lower monthly frequency or disaggregate monthly data into higher daily frequency. The former one is quite simple because we only need to average the daily data values over a month to convert the data from daily frequency to monthly frequency. The latter one is much more complicated because new method is required. Since we want to keep the daily price variation in the corn and oil price as what they originally have, we use Chow-Lin procedure to disaggregate monthly data, which is proposed and used in previous paper (Chow, Lin, 1971). It basically disaggregate and interpolate a relatively low frequency (monthly) time series to a relatively higher (daily) frequency time series, where either the sum, the average, the first or the last value of the resulting high frequency time series is consistent with the original low frequency data. The cross-price elasticity of demand between oil and corn is also monthly data and need to be converted to daily frequency series. More importantly, we do not have direct data on the elasticity itself. However, based on the definition of cross-price elasticity of demand and available data on ethanol price and oil sales, we can derive it from the ratio of percent change in quantity demanded of oil to the percent change of ethanol price. However, since the elasticity is derived under the “ceteris paribus” assumption and we do think that in the realistic market, supply side changes do exist, we should be very critical and skeptical toward this derived measure.

## **VII. Results and Findings**

From Figure 2, 3, 4, conditional variances of oil and corn and their conditional correlation are much more smooth and stable than the corresponding realized measures. The trend of

correlation maintains at around a level of 0.2 before 2008 and after 2012. During the middle period, the correlation doubles and fluctuates more frequently, which can be contributed to many events and changes such as financial crisis, surge of commodity prices and implementation of EISA (Energy Independence and Security Act) of 2007. From Bai and Perron procedure, we obtain that the optimal number of break dates that indicate significant structural changes is three since it has both the largest scaled F-statistic and the largest weighted F-statistic. The first break date is estimated to be April 01, 2008 that corresponds to the start of the Great Recession from 2008 to 2012. The second one is Oct 08, 2010 that marks the ever faster development of biofuel production and it is during 2010 that U.S. became the largest net exporter of ethanol in the world. At the same time, the price of corn started to increase and reach its peak. The third one is May 13, 2013 that coincides with the U.S. largest annual increase in oil production. During the same period, transportation efficiency has been fairly improved, relatively decreasing the demand of oil per vehicle per mile. These two phenomena combined can partially explain the decreasing trend of the oil price in the first half of 2013 and the sharp drop throughout the entire year of 2014. Based upon the fact that the past decade has witnessed a large amount of economic fluctuations and market changes, it is not surprising that our first Granger Causality test under full dataset is not very significant. Since the supply of oil is more exogenous, the causality is supposed to go from oil price to corn price, if there exists one. However, from the results, we can only see that when lag is limited to 1 or 2, the causality from oil price to corn price is weakly significant at the 0.05 significance level. Thus, it is necessary for us to do separated tests for different regimes endogenously determined by BP procedure.

In the first regime, as we can see from Figure 1, oil price and corn price are both relatively low and no obvious co-movement between them can be detected. More interestingly, after around Jun. 2004, these prices start to move generally in opposite directions. From the



causality test results, we can see that during this period, the Granger causality from oil price to corn price is significant for all lags from one to five at the 0.05 significance level and it is strongly significant for the first four lags at the 0.01 significance level. It makes intuitive sense that oil price does gain some predictive power in corn price because the financial market is relatively stable at that time and no exogenous shocks blow up the volatilities in both commodities. However, during the second and third period when the economy fluctuates and financial market crashes, we observe strong co-movement between oil price and corn price, which coincides with the financialization of commodities and index investment illustrated in Tang and Xiong's paper (2012). During these two periods, there does not exist any predictability either in oil price or corn price because otherwise, arbitrage would be observed in the financial market and speculation on the causal relationship can help people make a large amount of money. In the last period, it may be reasonable to assume that the causal relationship between oil and corn prices observed in the first period can be resumed to a certain degree after the financial turbulence. However, the Granger causality test in this period shows that there is some predictability in corn price, which is the opposite direction to what we expect. From Figure 1, we can observe that the corn prices decline first and then the drop in the oil prices follows. The Global Crop Production 2013 states that Corn plunged 40 percent in 2013 as the U.S. is recovering from the prior season when crops were hurt by the worst drought since the 1930s. The unexpected poor weather condition that prohibits adequate harvest of corn is obviously uncorrelated with oil market. For the subsequent fall in the oil prices, despite the low demand for oil that can be partially contributed to the growing switch away from oil to other fuels like ethanol, other political and economic reasons behind this price drop such as the turmoil in Iraq and Libya and less need to import oil are not associated with corn production and crop market. Therefore, we should be really skeptical and critical about the weak significance of the causality from corn price to oil price during

this period because it may be merely due to the fact that huge decreases in both prices make oil and corn much more volatile and the model is deficient in characterizing the whole dynamics.

Then we proceed with the results of impulse response functions that may provide a more visual assessment of how oil and corn adjust to one-time shocks both to themselves and to the other good. In general, Figure 5, 6, 7, 8, 9 all have the responses declining back to the steady state level. It makes intuitive sense because the VAR model includes only stationary time series. Therefore, the entire system of equations is stable. Under the full dataset, we can see that both oil and corn respond to their own exogenous impulses significantly and positively. These impacts vanish quickly by the second day without any statistically significant oscillations in later horizons. However, there are no or minor significant adjustments across commodities, corresponding to the fact that we do not detect any significant unidirectional causality in the price levels from Granger causality test. Additionally, as we have explained earlier, the first sub-period has the most stable economic environment while the second and third ones are largely disturbed. Thus, all the four responses in the first sub-period have the smallest variance to the exogenous shocks. The fourth sub-period is on its way back to equilibrium while we cannot find the expected causality from corn to oil that should mimic what found in the first sub-period. Thus, the results are relatively different among different regimes and we should analyze each regime separately.

In the first sub-period, the percentage change of corn prices responds significantly and positively to shocks in oil prices although the initial impact quickly mitigates after one day. However, it is noticeable that a shock in corn prices has an insignificantly minor impact on the oil prices as zero response is always captured in the 95% confidence interval throughout

the time. Therefore, the corn's contributions to the volatilities or the size of changes in oil prices are very limited while the oil's effect on the volatilities of corn prices is significant. In the second sub-period, we find that oil prices start to respond to its own shocks in a significantly oscillating way and the time needed to resume back to the steady state level is prolonged. Continuing to the third sub-period, we can observe that the time for oil to resume back to its steady state level is even longer due to higher and more oscillations. However, in both regimes, even if corn also has a longer response time, it still does not have any significant oscillations above and below the zero after the initial sharp response to its own shocks. Moreover, regarding the impulse responses across oil and corn, the significance lies differently through the process. In the second regime, from Figure 7 b), the percentage price change of oil does not have a significant initial response to the price shocks of corn. However, when  $t=4$ , a significant oscillation in the response appears. In the third regime, from Figure 8 c), the initial response of the percentage change of corn prices to oil price shocks is significant while no significant oscillations can be found later. Due to the highly volatile economic markets and the rapidly changing investment tastes during these two periods, any significance in the responses across commodities are supposed to be spurious and should not be considered meaningful and informative to a large degree. In the fourth post-crisis sub-period, we find inconsistency between our expectation of well-defined unidirectional causality from corn to oil and the reality of high volatilities, nonexistent arbitrage and even reversed causality. If we compare Figure 6 and Figure 9 together, we may find that the underlying mechanism or process for oil to respond to its own shocks is relatively different. In Figure 6, oil directly and quickly resumes back to its steady state level after one time horizon. However, in Figure 9, despite the same positively and significantly initial response, the percentage change of oil prices suffers from more oscillations and spends more time to eventually resume back to the zero line. Therefore, we may conclude that the

inconsistency is partially attributed to the higher volatilities in oil prices and the inelastic changing economic markets after financial crisis.

Due to the fact of nonexistent arbitrage in the financial market and seemingly spurious causality between oil and corn at the price levels, we further perform the second moment causality test in volatilities of oil and corn. As a result, we can see that the change in log likelihood for  $e_{\text{corn}}$  is highly significant compared to the critical value  $\chi_{1,0.05}^2$  while the change in log likelihood for  $d_{\text{oil}}$  is not. Therefore, returns of oil at period  $t - 1$  do statistically significantly help predict the volatility of corn at period  $t$  while the other way around is not. In other words, knowing the returns, or the price changes of oil in current period can help us forecast the uncertainty of oil price in the next period. Moreover, after controlling for the returns of ethanol and the returns of wheat, we conduct the causality test again and the estimated coefficients, whose significance represents the causal direction, correspond to our results in the previous uncontrolled test. In fact, the coefficient estimate of  $e_{\text{corn}}$ , which represents the second moment causality from the returns of oil to the volatilities of corn, becomes more significant.

## VII. Conclusions

This paper focuses on the better estimation of the correlation between oil and corn with Realized Beta GARCH model, the first moment Granger Causality test on the prices of oil and corn, and the second moment causality test on the volatilities and returns of oil and corn under structural changes. The U.S. biofuel policies, automobile engine technology, and underlying financial market changes give rise to a complex relationship between oil, ethanol and corn. Oil and ethanol can be substitutes, complements or can be completely unrelated and

their prices can move either together or apart. The correlation between oil and corn can be strengthened, mitigated or fluctuated in the past decade and the causalities of prices, volatilities and returns between these two commodities can go either direction or even be insignificant. Based on my results and findings, the correlation between oil and corn prices nearly doubled between 2008 and 2010, corresponding to the trend of increasing investment and financialization of agricultural commodities in the financial market described in Tang and Xiong's paper (2012). The decrease in the correlation after 2010 may be attributed to the unsynchronized changes in oil and corn prices. Even if both prices dramatically decline, the drop in corn prices precedes the drop in oil prices due to the unexpected exogenous shocks in agricultural production before changes in oil imports.

Conforming to Merkusheva and Rapsomanikis' conclusion (2012) that oil prices are shown to be the driven forces in the long-run grains prices, we also found that oil prices can Granger cause corn prices before April 2008, which means that current oil prices play a significant role in determining the future corn prices. However, the causality in later years has shown to be insignificant or even weakly reversed. The results from impulse response analysis largely agree to our causality tests and help draw a limited insight and explanation to certain confusing phenomena in the forth sub-period. However, there are a number of further problems that render the interpretation of impulse responses difficult since I only use the percentage change in the oil and corn prices in the VAR model for simplicity and for stationarity. The major limitation is the incompleteness or the low-dimension of our model. In real economic systems, the intertwining relationships between markets, consumers, and commodities are too complicated to be fully captured. Therefore, we usually have a simplified yet neater model for computational and interpretation efficiency. However, if important variables are omitted from the model, we may have major distortions in the impulse responses and makes the results worthless for structural and instructive

interpretations. Further improvement on the VAR model leaves to future research and exploration on the first moment causality of oil and corn prices.

However, the bivariate Granger causality model used for the first moment causality test is limited due to its simplicity and thus cannot sufficiently capture the gradual changes in the volatilities and underlying relationship between oil and corn. Therefore, we also conducted the second moment causality test in volatilities and returns of oil and corn. It is much more sensible, comprehensive and intuitive because the price variables are more likely to follow the unit root random walks that prohibit people from observing significant predictive power and from taking advantage of any meaningful relationships among the financial assets, like oil and corn. Before and after controlling for other stationary variables such as the returns of wheat and of ethanol, we found that only the coefficient that represents the causality from the returns of oil to the volatilities of corn is significant.

In conclusion, my empirical research on the evolving relationship between oil and corn here revealed some interesting and informative features of the correlation and causality between oil and corn both on the price levels and volatilities. Since 1990s, the U.S. government has implemented several biofuel policies and a variety of subsidies as a way to reduce dependence on oil as the energy source and to increase the nation' overall sustainability.

Based on our findings, during the prior crisis period, namely from 2004 to 2008, changes in the oil prices that may result from biofuel policy implementation and new standard on the ethanol production and consumption can have a predictive power in the corn prices.

Therefore, even if ethanol production utilizes a considerable share of total corn use and diverts valuable crop land away from agricultural production, the increase in the food prices and costs of animal feed that may put more pressure on crop and livestock farmers as well as low-income consumers can be estimated and forecasted to a certain degree. However, starting

from the financial crisis, any changes in the oil prices are shown to only significantly affect the volatilities rather than the mere values of corn prices. In other words, the range of the potential changes of corn prices is enlarged, which entails more uncertainties both in the energy market and in the food market. Therefore, at present, new policies regarding biofuel promotion should take these increased volatilities and volatility spillovers between oil and corn into serious consideration. In fact, based on the increased financialization in the commodity markets, as indicated by Tang and Xiong in 2012, more non-energy agricultural commodities, at least those frequently traded in the financial market and used in the energy production, are more closely associated with the energy market. Therefore, the overall methodologies used in this paper can be generalized to a larger choice of commodities and even to a multivariate analysis. The special hierarchical structure of the Realized Beta GARCH model that combines the flexibility of the GARCH modeling framework with precise incorporation of volatility measures enables it to be applied to a large variety of assets, commodities and thus multivariate analysis.

However, despite the advances brought by the relative comprehensive correlation and causation analysis of oil and corn in this paper, it still does not capture all of the dynamics between these two goods and further in the larger economic environment. On the one hand, regarding the first moment causality, I only conducted bivariate Granger causality test out of simplicity consideration, which can be further improved by taking other relevant factors into account. Additionally, the specific choice of Cholesky decomposition used in the impulse response functions is used in the previous papers but sensitive to ordering. Other variance decomposition method, such as Generalized Variance Decomposition (GVD), may be applied as well to if the results may be of any difference. On the other hand, however, due to the nonexistent arbitrage in the financial market, we may rather explore further and deeper on the second moment causality track. For example, it may be interesting to explore how news

release and public speeches from the government can affect the causalities of returns and volatilities between oil and corn. Any significant results from these modifications and improvement, in combination with this research, can provide policy makers with more practical insight and feasible guidance to mitigate the potential negative or unpredictable impact on the consumers, investors and suppliers in the related sectors.



## VII. Bibliography

- Bai, J. and Perron, P. (2003). "Computation and analysis of multiple structural change models." *Journal of Applied Econometrics* 18(1): 1-22.
- Balcombe, K. and G. Rapsomanikis, 2008. Bayesian Estimation and Selection of Nonlinear Vector Error Correction Models: The Case of the Sugar-Ethanol-Oil Nexus in Brazil. *American Journal of Agricultural Economics*, 90(3): 658-668.
- Choi, I. and Saikkonen, P. (2004). "Testing linearity in cointegrating smooth transition regressions." *Econometrics Journal* 7: 341-365.
- Chow, G.C., Lin, A.L. (1971), "Best linear unbiased interpolation, distribution and extrapolation of time series by related series," *The review of Economics and Statistics* 53, pp. 372-375.
- Kwiatkowski, D., Phillips, P.C.B., Schmidt, P., Shin, Y. (1992). "Testing the null hypothesis of stationarity against the alternative of a unit root." *Journal of Econometrics* 54: 159-178.
- Hansen, Peter R, Asger Lunde, and Valeri Voev. 2014. "Realized Beta Garch: A Multivariate Garch Model with Realized Measures of Volatility." *Journal of Applied Econometrics* 29 (April): 774–779.
- Hansen, Peter R, and Zhuo Huang. 2016. "Exponential GARCH Modeling With Realized Measures of Volatility." *Journal of Business & Economic Statistics* 34 (2) (March 17): 269–287. <http://dx.doi.org/10.1080/07350015.2015.1038543>.
- Granger, C.W.J., 1969. "Investigating causal relations by econometric models and cross spectral methods." *Econometrica* 37, 424–438.

Granger, C.W.J., 1980. "Testing for causality— a personal viewpoint." *Journal of Economic Dynamics and Control* 2, 329–352.

Granger, 1988. "Some recent developments in a concept of causality." *Journal of Econometrics* 39, 199–211.

Tang, Ke, and Wei Xiong. 2012. "Index Investment and the Financialization of Commodities." *Financial Analysts Journal* 68 (6).

Thorp, Susan. n.d. "Crude Oil and Agricultural Futures: An Analysis of Correlation Dynamics." *Discipline of Finance, The University of Sydney Business School, University of Sydney, NSW, 2006, Australia.*

Tyner, W. E. and Taheripour, F. (2008). "Policy options for integrated energy and agricultural markets." *Review of Agricultural Economics* 30: 387-396.

USDA's Joint Agricultural Weather Facility, "Global Crop Production Review, 2013"

## Appendix A. Definitions

Table 1: Definitions of Variables in the Dataset

| Variables         | Definitions  |
|-------------------|--|
| $x_{oil,t}$       | Realized returns of oil at time t calculated from the difference between log prices in a trading day     |
| $x_{corn,t}$      | Realized returns of corn at time t calculated from the difference between log prices in a trading day    |
| $y_{oil,corn,t}$  | Realized correlation between oil and corn at time t ranging from -1 to 1.                                |
| $p_{oil,t}$       | Price of oil at time (trading day) t   |
| $p_{corn,t}$      | Price of corn at time (trading day) t  |
| $p_{ethanol,t}$   | Price of ethanol at time (trading day) t   |
| $p_{wheat,t}$     | Price of wheat at time (trading day) t   |
| $r_{inflation,t}$ | Monthly Consumer Price Index (CPI) as a proxy  |
| $\eta_t$          | Cross-price elasticity between oil and corn  |
| $s_t$             | Quarterly corn share used in the energy production   |
| $x_{ethanol,t}$   | Realized returns of ethanol at time t calculated from the difference between log prices in a trading day |
| $x_{wheat,t}$     | Realized returns of wheat at time t calculated from the difference between log prices in a trading day   |

Note: Every variable presented above can be directly obtained from the dataset except for the cross-price elasticity between oil and corn. It is indirectly derived from the ratio of percent change in quantity demanded of oil to the percent change of ethanol price.

## Appendix B. Summary Statistics

Table 2. Summary Statistics of Variables in the Dataset

| <b>Variable</b>   | <b>Obs.</b> | <b>Mean</b> | <b>Std.Dev.</b> | <b>Min</b> | <b>Max</b> |
|-------------------|-------------|-------------|-----------------|------------|------------|
| $x_{oil,t}$       | 2765        | -0.0518     | 1.935444        | -10.97     | 10.23      |
| $x_{corn,t}$      | 2765        | 0.03204     | 1.819440        | -7.79      | 8.97       |
| $y_{oil,corn,t}$  | 2765        | 0.10306     | 0.155971        | -0.55      | 0.79       |
| $p_{oil,t}$       | 2765        | 80.1263     | 4.63852         | 46.86      | 151.72     |
| $p_{corn,t}$      | 2765        | 4.12045     | 1.22309         | 1.72       | 8.13       |
| $p_{ethanol,t}$   | 2765        | 2.22586     | 0.453011        | 1.20       | 3.58       |
| $p_{wheat,t}$     | 2765        | 2.46446     | 0.70961         | 1.14290    | 4.39720    |
| $r_{inflation,t}$ | 2765        | 2.32812     | 1.42133         | -2.1       | 5.6        |
| $\eta_t$          | 2765        | 0.28652     | 1.40390         | -1.90264   | 11.69656   |
| $s_t$             | 2765        | 29.57385    | 13.15174        | 9.40701    | 61.49924   |
| $x_{ethanol,t}$   | 2765        | 0.12578     | 0.89921         | 0.00000    | 0.80912    |
| $x_{wheat,t}$     | 2765        | 0.09325     | 1.34576         | 0.00000    | 0.54638    |

Note: This table presents the number of observations (Obs.), average (Mean), standard deviation (Std. Dev.), minimum value (Min) and maximum value (Max) for each variable defined in Table 1. The number of observations is consistently 2765, which is the number of remaining trading days after I applied the abnormally short trading days criterion.

## Appendix C. Preliminary Estimates

Table 3. Derived Variables and Estimated Coefficients in Realized Beta GARCH Model

| Variable            | Obs. | Mean    | Std.Dev. | Min    | Max   |
|---------------------|------|---------|----------|--------|-------|
| $r_{oil,t}$         | 2765 | -0.0408 | 2.05342  | -8.234 | 8.357 |
| $h_{oil,t}$         | 2765 | 3.80042 | 3.837562 | 0.06   | 33.84 |
| $r_{corn,t}$        | 2765 | 0.03521 | 1.70349  | -6.23  | 5.85  |
| $h_{corn,t}$        | 2765 | 3.35326 | 1.675413 | 0.47   | 12.30 |
| $\rho_{oil,corn,t}$ | 2765 | 0.32025 | 0.10349  | -0.02  | 0.53  |
| $u_{oil,t}$         | 2765 | 0.01341 | 0.439861 | -1.94  | 5.25  |
| $u_{corn,t}$        | 2765 | 0.00032 | 0.479962 | -2.04  | 2.57  |
| $v_{oil,corn,t}$    | 2765 | 0.00024 | 0.134454 | -0.65  | 0.96  |

### Coefficient Estimates (including $p_{ethanol,t}$ , $p_{wheat,t}$ , $r_{inflation,t}$ , $\eta_t$ , $S_t$ )

|                  |  |                      |   |
|------------------|--|----------------------|---|
| $a_{oil}$        | 0.03905  | $\sigma_{corn}^2$    | 0.22983   |
| $b_{oil}$        | 0.84277  | $\xi_{corn}$         | -0.4810   |
| $c_{oil}$        | 0.13508  | $\varphi_{corn}$     | 1.21453   |
| $\xi_{oil}$      | -0.21268   | $\delta_{corn}$      | $0.01997z_{corn,t-1} + 0.147(z_{corn,t-1}^2 - 1)$ |
| $\varphi_{oil}$  | 1.06920  | $\xi_{oil,corn}$     | 0.00515   |
| $\delta_{oil}$   | $-0.07177z_{oil,t-1} + 0.10497(z_{oil,t-1}^2 - 1)$ | $\varphi_{oil,corn}$ | 0.78621   |
| $a_{corn}$       | 0.16734  | $h_{oil,0}$          | -1.85931  |
| $b_{corn}$       | 0.643600   | $h_{corn,0}$         | 0.39018   |
| $c_{corn}$       | 0.213110   | $h_{oil,corn,0}$     | 0.05736   |
| $d_{corn}$       | 0.016750   | $\mu_{corn}$         | 0.04769   |
| $\sigma_{oil}^2$ | 0.16583  | $\sigma_{corn}^2$    | 0.22983   |
| $\mu_{oil}$      | -0.03352   |                      |   |

Note: This table shows the better estimated conditional variables and related coefficients that are jointly defined by the Realized Beta GARCH model, or namely equations (1), (2), (3), (4) and (5).

Table 4. OLS Estimates and Structural Breaks from Bai and Perron Procedure

(Dependent Variable:  $\rho_{oil, corn, t}$ , N:2765)

| Variable                               | Coefficient | Std. Error | t-Statistic | Prob.  |
|--|-------------|------------|-------------|--------|
| Oil_Price**                            | 0.07657     | 0.035681   | 2.145960    | 0.0160 |
| Corn_Price                             | 0.34672     | 0.507782   | 0.682813    | 0.5002 |
| Ethanol_Price                          | -0.202436   | 0.445601   | -0.454299   | 0.6496 |
| Inflation_rate                         | 0.085486    | 0.126690   | 0.674761    | 0.4999 |
| Cross-Elasticity**                     | 0.091926    | 0.0445601  | 2.458399    | 0.0140 |
| Wheat Price                            | 1.458716    | 0.502174   | 2.904803    | 0.0037 |
| Ethanol_Share_in_<br>Total_Corn_Use*** | 0.055116    | 0.016327   | 3.375690    | 0.0007 |
| R-squared                              | 0.789075    |            |             |        |
| Adjusted R-squared                     | 0.788693    |            |             |        |
| S.E. of regression                     | 0.796453    |            |             |        |
| Sum squared resid                      | 1750.137    |            |             |        |

Note: 1. \* Significant at 10%; \*\* Significant at 5%; \*\*\* Significant at 1%.  
 2. We use HAC (Heteroskedasticity and Autocorrelation Consistent) covariance estimation in order to allow for serial correlation in the errors.

| Breaks | F-statistic | Scaled F-statistic | Weighted F-statistic | Critical Value |
|--------|-------------|--------------------|----------------------|----------------|
| 1*     | 15.53604    | 93.21623           | 93.21623             | 20.08          |
| 2*     | 28.28419    | 169.7051           | 196.1819             | 17.37          |
| 3*     | 48.56434    | 291.3860           | 375.5476             | 15.58          |
| 4*     | 27.90480    | 167.4288           | 241.8683             | 13.90          |
| 5*     | 32.48390    | 194.9034           | 327.7773             | 11.94          |

Note: \* Significant at 5%

Estimated Break Dates: 1: Mar 04, 2009

2: Sept 15, 2010, May 13, 2013

3: Apr 01, 2008, Oct 08, 2010, May 13, 2013

4: Nov 1, 2006, Jul 14, 2008, Oct 08, 2010, May 13, 2013

5: Oct 03, 2005, Jun 01, 2007, Jan 22, 2009, Oct 08, 2010  
 May 13, 2013

**Number of Breaks Chosen: 3**

---

Note: This table first details the OLS regression according to equation (7) with dependent variable to be the conditional correlation  $\rho_{oil,corn,t}$  we obtained earlier in the Realized Beta GARCH model and independent variables to be the other seven relevant daily and dis-aggregated from monthly to daily variables. The optimal number of structural breaks and the estimation of the dates are provided in the second part of the table based on the criterion of minimum sum of squared residuals from equation (8) and (9).

Table 5. Bivariate Granger Causality Test Results between Oil and Corn Prices  
(Lag=1, 2, 3, 4, 5 without Structural Breaks)

| Null Hypothesis                              | Num. of Lags | F-statistic | Prob.  |
|--|--------------|-------------|--------|
| Oil_Price does not Granger Cause Corn_Price  | 1            | 0.51115     | 0.4747 |
| Corn_Price does not Granger Cause Oil_Price* | 1            | 6.04841     | 0.0140 |
| Oil_Price does not Granger Cause Corn_Price  | 2            | 0.60263     | 0.5474 |
| Corn_Price does not Granger Cause Oil_Price* | 2            | 3.20563     | 0.0407 |
| Oil_Price does not Granger Cause Corn_Price  | 3            | 0.47950     | 0.6966 |
| Corn_Price does not Granger Cause Oil_Price  | 3            | 2.32010     | 0.0734 |
| Oil_Price does not Granger Cause Corn_Price  | 4            | 1.64446     | 0.1603 |
| Corn_Price does not Granger Cause Oil_Price  | 4            | 1.85162     | 0.1162 |
| Oil_Price does not Granger Cause Corn_Price  | 5            | 1.41912     | 0.2140 |
| Corn_Price does not Granger Cause Oil_Price  | 5            | 1.44856     | 0.2036 |

Note: 1. \* represents we can reject the null hypothesis at 5% significance level.

2. The F-statistic is defined as  $\frac{(SSR_r - SSR_u)/n}{SSR_u/(T - (m+n+1))}$  where  $SSR_r$  and  $SSR_u$  are the two sums of squared residuals related to the restricted and unrestricted form of the VAR model in section 4.3.1. The degrees of freedom are  $T=2765$ , which is the number of observations in the full dataset while  $n$  and  $m$  are the number of lags as it can be seen from equation (11).



Table 6. Bivariate Granger Causality Test Results between Oil and Corn Prices

(Lag=1, 2, 3, 4, 5 with Structural Breaks: Jan 05, 2004 to Apr 01, 2008)

| Null Hypothesis                               | Num. of Lags | F-statistic | Prob.  |
|---|--------------|-------------|--------|
| Oil_Price does not Granger Cause Corn_Price** | 1            | 0.9241      | 0.0091 |
| Corn_Price does not Granger Cause Oil_Price   | 1            | 3.27451     | 0.0708 |
| Oil_Price does not Granger Cause Corn_Price** | 2            | 8.70230     | 0.0002 |
| Corn_Price does not Granger Cause Oil_Price   | 2            | 0.93887     | 0.0407 |
| Oil_Price does not Granger Cause Corn_Price** | 3            | 6.97756     | 0.0001 |
| Corn_Price does not Granger Cause Oil_Price   | 3            | 0.86169     | 0.4605 |
| Oil_Price does not Granger Cause Corn_Price** | 4            | 5.20136     | 0.0004 |
| Corn_Price does not Granger Cause Oil_Price   | 4            | 1.09415     | 0.3580 |
| Oil_Price does not Granger Cause Corn_Price*  | 5            | 4.09392     | 0.0011 |
| Corn_Price does not Granger Cause Oil_Price   | 5            | 0.86604     | 0.5033 |

Note: 1. \* represents we can reject the null hypothesis at 5% significance level. \*\* represents we can reject the null hypothesis at 1% significance level.

2. The F-statistic is defined as  $\frac{(SSR_r - SSR_u)/n}{SSR_u/(T - (m + n + 1))}$  where  $SSR_r$  and  $SSR_u$  are the two sums of squared residuals related to the restricted and unrestricted form of the VAR model in section 4.3.1. The degrees of freedom are  $T=1062$ , which is the number of observations in the first regime while  $n$  and  $m$  are the number of lags as it can be seen from equation (11).

Table 7. Bivariate Granger Causality Test Results between Oil and Corn Prices

(Lag=1, 2, 3, 4, 5 with Structural Breaks: Apr 02, 2008 to Oct 08, 2010)

| Null Hypothesis                             | Num. of Lags | F-statistic | Prob.  |
|---|--------------|-------------|--------|
| Oil_Price does not Granger Cause Corn_Price | 1            | 2.03293     | 0.1542 |
| Corn_Price does not Granger Cause Oil_Price | 1            | 1.69042     | 0.1938 |
| Oil_Price does not Granger Cause Corn_Price | 2            | 0.83878     | 0.4327 |
| Corn_Price does not Granger Cause Oil_Price | 2            | 1.70748     | 0.1822 |
| Oil_Price does not Granger Cause Corn_Price | 3            | 0.58229     | 0.6268 |
| Corn_Price does not Granger Cause Oil_Price | 3            | 1.16570     | 0.3220 |
| Oil_Price does not Granger Cause Corn_Price | 4            | 1.60095     | 0.1724 |
| Corn_Price does not Granger Cause Oil_Price | 4            | 1.13236     | 0.3401 |
| Oil_Price does not Granger Cause Corn_Price | 5            | 1.54497     | 0.1738 |
| Corn_Price does not Granger Cause Oil_Price | 5            | 0.93360     | 0.4585 |

Note: 1. \* represents we can reject the null hypothesis at 5% significance level.

2. The F-statistic is defined as  $\frac{(SSR_r - SSR_u)/n}{SSR_u/(T - (m+n+1))}$  where  $SSR_r$  and  $SSR_u$  are the two sums of squared residuals related to the restricted and unrestricted form of the VAR model in section 4.3.1. The degrees of freedom are  $T=637$ , which is the number of observations in the second regime while  $n$  and  $m$  are the number of lags as it can be seen from equation (11).

Table 8. Bivariate Granger Causality Test Results between Oil and Corn Prices

(Lag=1, 2, 3, 4, 5 with Structural Breaks: Oct 11, 2010 to May 13, 2013)

| Null Hypothesis                              | Num. of Lags | F-statistic | Prob.  |
|--|--------------|-------------|--------|
| Oil_Price does not Granger Cause Corn_Price  | 1            | 1.10268     | 0.2941 |
| Corn_Price does not Granger Cause Oil_Price  | 1            | 0.02249     | 0.8808 |
| Oil_Price does not Granger Cause Corn_Price  | 2            | 1.17854     | 0.3084 |
| Corn_Price does not Granger Cause Oil_Price  | 2            | 0.15564     | 0.8559 |
| Oil_Price does not Granger Cause Corn_Price  | 3            | 0.75740     | 0.5183 |
| Corn_Price does not Granger Cause Oil_Price  | 3            | 0.12736     | 0.9439 |
| Oil_Price does not Granger Cause Corn_Price* | 4            | 3.53081     | 0.0415 |
| Corn_Price does not Granger Cause Oil_Price  | 4            | 0.28197     | 0.8897 |
| Oil_Price does not Granger Cause Corn_Price  | 5            | 1.32127     | 0.2531 |
| Corn_Price does not Granger Cause Oil_Price  | 5            | 0.27437     | 0.9272 |

Note: 1. \* represents we can reject the null hypothesis at 5% significance level.

2. The F-statistic is defined as  $\frac{(SSR_r - SSR_u)/n}{SSR_u/(T - (m+n+1))}$  where  $SSR_r$  and  $SSR_u$  are the two sums of squared residuals related to the restricted and unrestricted form of the VAR model in section 4.3.1. The degrees of freedom are  $T=653$ , which is the number of observations in the third regime while  $n$  and  $m$  are the number of lags as it can be seen from equation (11).

Table 9. Bivariate Granger Causality Test Results between Oil and Corn Prices

(Lag=1, 2, 3, 4, 5 with Structural Breaks: May 14, 2013 to Dec 31, 2014)

| Null Hypothesis                               | Num. of Lags | F-statistic | Prob.  |
|---|--------------|-------------|--------|
| Oil_Price does not Granger Cause Corn_Price   | 1            | 0.64520     | 0.4223 |
| Corn_Price does not Granger Cause Oil_Price** | 1            | 7.24677     | 0.0074 |
| Oil_Price does not Granger Cause Corn_Price   | 2            | 0.79291     | 0.4532 |
| Corn_Price does not Granger Cause Oil_Price*  | 2            | 4.51118     | 0.0115 |
| Oil_Price does not Granger Cause Corn_Price   | 3            | 1.40360     | 0.2412 |
| Corn_Price does not Granger Cause Oil_Price*  | 3            | 2.78277     | 0.0407 |
| Oil_Price does not Granger Cause Corn_Price   | 4            | 1.141201    | 0.2292 |
| Corn_Price does not Granger Cause Oil_Price*  | 4            | 2.73462     | 0.0287 |
| Oil_Price does not Granger Cause Corn_Price   | 5            | 1.29559     | 0.2649 |
| Corn_Price does not Granger Cause Oil_Price   | 5            | 2.15628     | 0.0582 |

Note: 1. \* represents we can reject the null hypothesis at 5% significance level. \*\* represents we can reject the null hypothesis at 1% significance level.

2. The F-statistic is defined as  $\frac{(SSR_r - SSR_u)/n}{SSR_u/(T - (m+n+1))}$  where  $SSR_r$  and  $SSR_u$  are the two sums of squared residuals related to the restricted and unrestricted form of the VAR model in section 4.3.1. The degrees of freedom are  $T=413$ , which is the number of observations in the fourth regime while  $n$  and  $m$  are the number of lags as it can be seen from equation (11).

Table 10. Augmented Dickey and Fuller's Unit Root Test for Full Dataset

(Null Hypothesis: Oil\_Price/ Corn\_Pirce has a unit root)

| Variable                     | t-Statistic | Prob.          |
|------------------------------|-------------|----------------|
| Oil                          | 2.313645    | 0.1677         |
| Corn                         | -1.697112   | 0.4327         |
| Augmented Dickey-Fuller Test | Level       | Test Statistic |
| Critical Values:             | 1% level    | -3.432524      |
|                              | 5% level    | -2.862386      |
|                              | 10% level   | -2.567265      |

Note: This table shows the results of Augmented Dickey and Fuller's Unit Root Test for the full dataset and provides three threshold statistics at the 1%, 5%, and 10% levels. The Augmented Dickey and Fuller's Unit Root Test is based on the model  $\Delta y_t = \alpha + \beta t + \gamma y_{t-1} + \delta_1 \Delta y_{t-1} + \delta_2 \Delta y_{t-2} + \dots + \delta_{p-1} \Delta y_{t-p+1} + \varepsilon_t$  where  $y_t$  is time series Oil or Corn respectively and  $p$ , the optimal lag length, has been selected through Akaike information criterion (AIC). The test statistic can be calculated by the ratio  $\frac{\hat{\gamma}}{SE(\hat{\gamma})}$ .

Table 11. Augmented Dickey and Fuller's Unit Root Test for 4 Sub-periods

(Null Hypothesis: Oil\_Price/ Corn\_Pirce has a unit root)

| Sub-period | Variable | t-Statistic | Prob.  |
|------------|----------|-------------|--------|
| 1          | Oil      | -0.650790   | 0.8565 |
|            | Corn     | 1.506764    | 0.9994 |
| 2          | Oil      | -1.405033   | 0.5808 |
|            | Corn     | -1.686196   | 0.4378 |
| 3          | Oil      | -2.03465    | 0.3284 |
|            | Corn     | -2.788445   | 0.0605 |
| 4          | Oil      | 2.170562    | 0.9999 |
|            | Corn     | -2.838834   | 0.0538 |

| Augmented Dickey-Fuller Test Critical Values: | Level     | Test Statistic |
|---|-----------|----------------|
|   | 1% level  | -3.432524      |
|   | 5% level  | -2.862386      |
|   | 10% level | -2.567265      |

Note: This table shows the results of the same test. However, it is applied to the four regimes separately instead. Similarly, three threshold statistics at the 1%, 5%, and 10% levels are provided at the bottom of the table.

Table 12. Augmented Dickey and Fuller's Unit Root Test for Full Dataset

(Null Hypothesis: dlogOil/ dlogCorn has a unit root)

| Variable   | t-Statistic | Prob.          |
|--|-------------|----------------|
| dlogOil  | -2.953791   | 0.0344         |
| dlogCorn   | -3.023083   | 0.0187         |
| Augmented Dickey-Fuller Test<br>Critical Values: | Level       | Test Statistic |
|  | 1% level    | -3.432524      |
|  | 5% level    | -2.862386      |
|  | 10% level   | -2.567265      |

Note: This table shows the results of the same test. However, it is applied to variables dlogOil and dlogCorn, which are the first difference of the log of the original price variables under the full dataset. Namely, it represents the percentage change in the prices since  $\log p_{t+1} - \log p_t \cong \% \Delta p = \frac{p_{t+1} - p_t}{p_t}$ . Similarly, three threshold statistics at the 1%, 5%, and 10% levels are provided at the bottom of the table.

Table 13. Augmented Dickey and Fuller's Unit Root Test for 4 Sub-periods  
(Null Hypothesis: dlogOil/ dlogCorn has a unit root)

| Sub-period                                    | Variable | t-Statistic | Prob.          |
|---|----------|-------------|----------------|
| 1   | dlogOil  | -3.72945    | 0.0065         |
|   | dlogCorn | -3.77142    | 0.0034         |
| 2   | dlogOil  | -4.67331    | 0.0058         |
|   | dlogCorn | -2.87025    | 0.0478         |
| 3   | dlogOil  | -2.95246    | 0.0401         |
|   | dlogCorn | -3.45988    | 0.0105         |
| 4   | dlogOil  | -2.98162    | 0.0488         |
|   | dlogCorn | -2.93417    | 0.0310         |
| Augmented Dickey-Fuller Test Critical Values: |          | Level       | Test Statistic |
|   |          | 1% level    | -3.432524      |
|   |          | 5% level    | -2.862386      |
|   |          | 10% level   | -2.567265      |

Note: This table shows the results of the same test using dlogOil and dlogCorn. However, it is applied to the four regimes separately instead. Similarly, three threshold statistics at the 1%, 5%, and 10% levels are provided at the bottom of the table.



Table 14. Causality Estimates for Volatility

(Not including  $x_{ethanol,t}$  and  $x_{ethanol,t}$ )**Coefficient Estimates**

|   |   |                         |  |
|---|---|-------------------------|--|
| $a_{oil}$   | 0.06312   | $\sigma_{corn}^2$       | 0.22983  |
| $b_{oil}$   | 0.75984   | $\xi_{corn}$            | -0.4810  |
| $c_{oil}$   | 0.20326   | $\varphi_{corn}$        | 1.21453  |
| $d_{oil}$   | 0.00872   | $e_{corn}$              | 0.27632  |
| $\xi_{oil}$                                       | -0.21268  | $\delta_{corn}$         | $-0.07177z_{oil,t-1} - 0.00347(z_{oil,t-1}^2 - 1)$ |
| $\varphi_{oil}$                                   | 1.06920   | $\xi_{oil,corn}$        | -0.0809  |
| $\delta_{oil}$                                    | $-0.05973z_{oil,t-1} - 0.0017(z_{oil,t-1}^2 - 1)$ | $\varphi_{oil,corn}$    | 0.78621  |
| $a_{corn}$  | 0.14784   | $h_{oil,0}$             | -4.68776   |
| $b_{corn}$  | 0.74039   | $h_{corn,0}$            | -1.61848   |
| $c_{corn}$  | 0.15935   | $h_{oil,corn,0}$        | 0.057362   |
| $d_{corn}$  | -0.1158   | $\sigma_{corn}^2$       | 0.22983  |
| $\mu_{oil}$                                       | -0.02626  | $\sigma_{oil}^2$        | 0.16583  |
| $\mu_{corn}$                                      | 0.04268   |                         |  |
| Change in log likelihood for $e_{corn}$ : 14.8237 |   | $\chi_{1,0.05}^2=3.841$ |  |
| Change in log likelihood for $d_{oil}$ : 1.64321  |   | $\chi_{1,0.05}^2=3.841$ |  |

Note: this tables shows the results of the re-estimated coefficients from the previous Realized Beta GARCH model by including the realized returns of one commodity at time period  $t-1$  in the GARCH equation of the other, which can be seen from equations (15) and (16). The test statistic here is the change in the log likelihood and the threshold value is  $\chi_{1,0.05}^2 = 3.841$ .

Table 15. Causality Estimates for Volatility

(Including  $x_{ethanol,t}$  and  $x_{wheat,t}$ )**Coefficient Estimates**

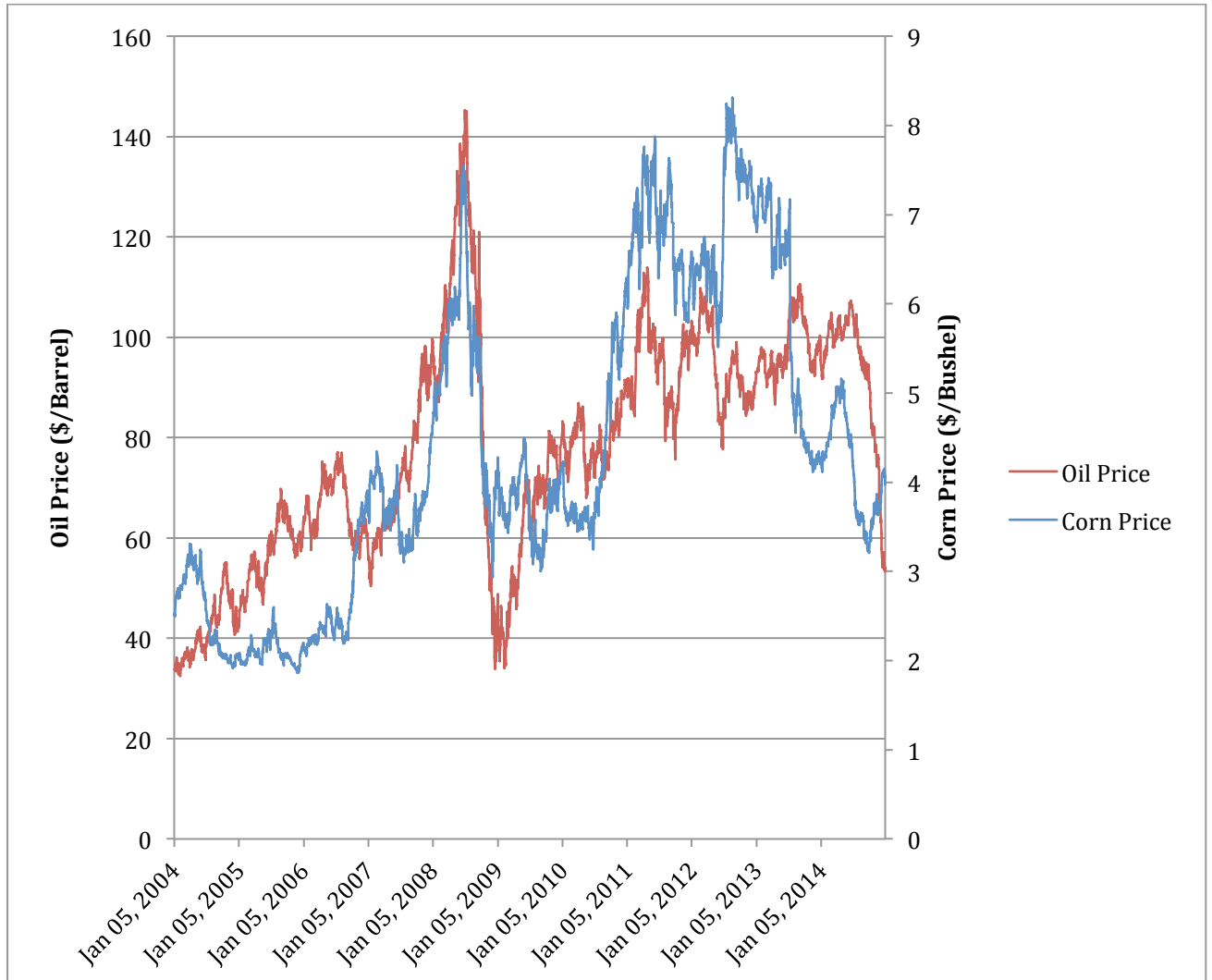
|   |  |                         |          |
|---|--|-------------------------|----------|
| $a_{oil}$   | 0.06311  | $\sigma_{corn}^2$       | 0.22983  |
| $b_{oil}$   | 0.75965  | $\xi_{corn}$            | -0.4810  |
| $c_{oil}$   | 0.20365  | $\varphi_{corn}$        | 1.21453  |
| $d_{oil}$   | 0.00219  | $e_{corn}$              | 0.32358  |
| $\varphi_{oil}$                                   | 1.06920  | $\xi_{oil,corn}$        | 0.00515  |
| $\delta_{oil}$                                    | $-0.07177z_{oil,t-1} + 0.10497(z_{oil,t-1}^2 - 1)$ | $\varphi_{oil,corn}$    | 0.78621  |
| $a_{corn}$  | 0.16734  | $x_{ethanol}$           | 0.00642  |
| $b_{corn}$  | 0.643600   | $x_{wheat}$             | -0.02173 |
| $c_{corn}$  | 0.213110   | $h_{oil,0}$             | -1.85931 |
| $d_{corn}$  | 0.016750   | $h_{corn,0}$            | 0.39018  |
| $\sigma_{oil}^2$                                  | 0.16583  | $h_{oil,corn,0}$        | 0.057362 |
| $\sigma_{corn}^2$                                 | 0.22983  | $\mu_{corn}$            | 0.04769  |
| $\mu_{oil}$                                       | -0.03352   |                         |          |
| Change in log likelihood for $e_{corn}$ : 20.6085 |  | $\chi_{1,0.05}^2=3.841$ |          |
| Change in log likelihood for $d_{oil}$ : 1.20572  |  | $\chi_{1,0.05}^2=3.841$ |          |

Note: this tables shows the results of the re-estimated coefficients from the previous Realized Beta GARCH model by including the realized returns of one commodity at time period  $t-1$  in the GARCH equation of the other, which can be seen from equations (15) and (16). Additionally, it also includes the realized returns of ethanol and wheat as two close substitutes for oil and corn respectively. The test statistic here is the change in the log likelihood and the threshold value is  $\chi_{1,0.05}^2 = 3.841$ .

## Appendix D. Figures

Figure 1.

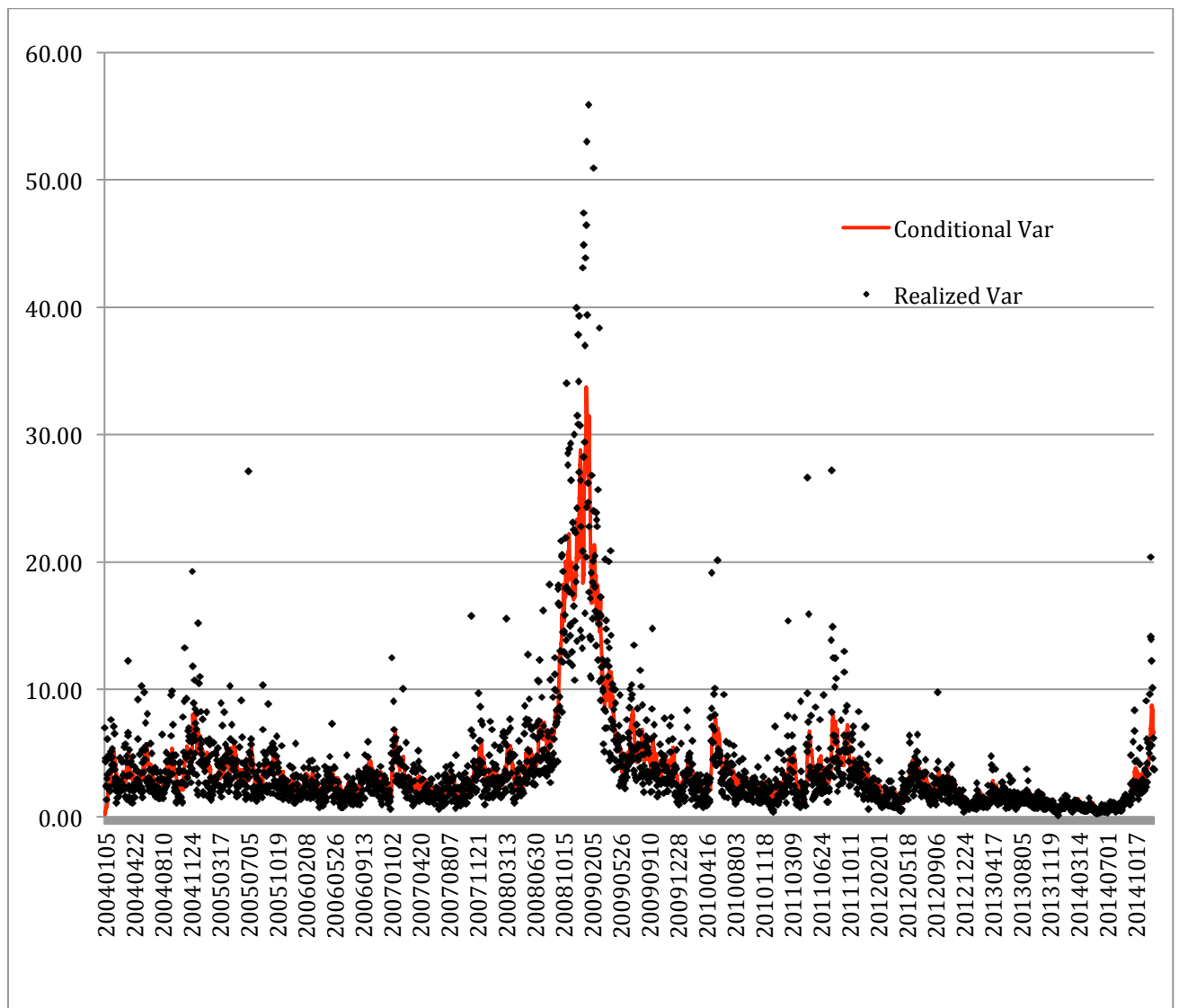
Oil and Corn prices from 2004 to 2014



This figure is the direct plot of oil and corn prices from 2004 to 2014 in the dataset.

Figure 2.

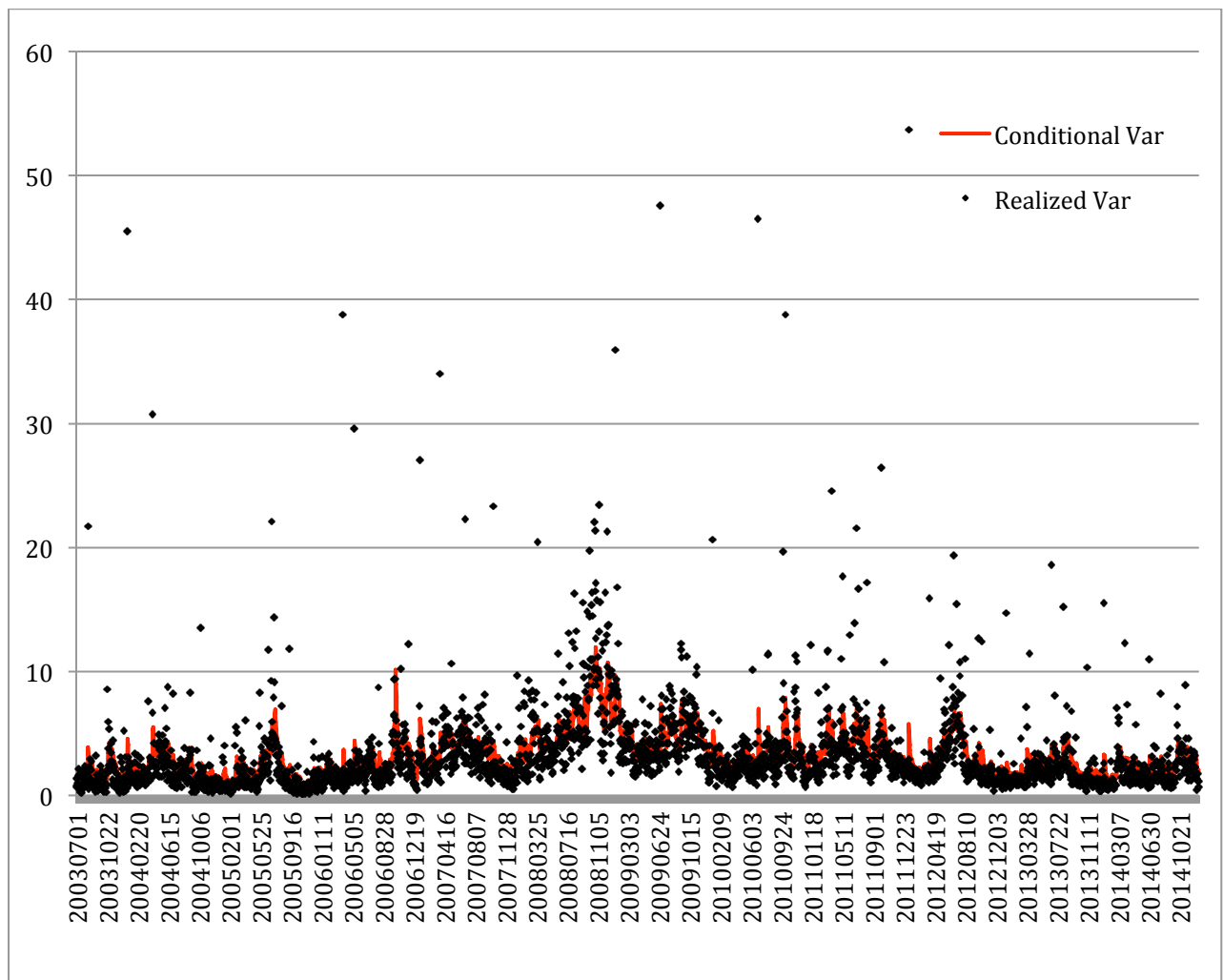
### Conditional and Realized Variance of Oil from 2004 to 2014



This figure shows the realized variances of oil that can be directly plotted based on the variable  $x_{oil,t}$  in the dataset and the conditional variances that we estimated by the Realized Beta GARCH model in section 4.1.

Figure 3.

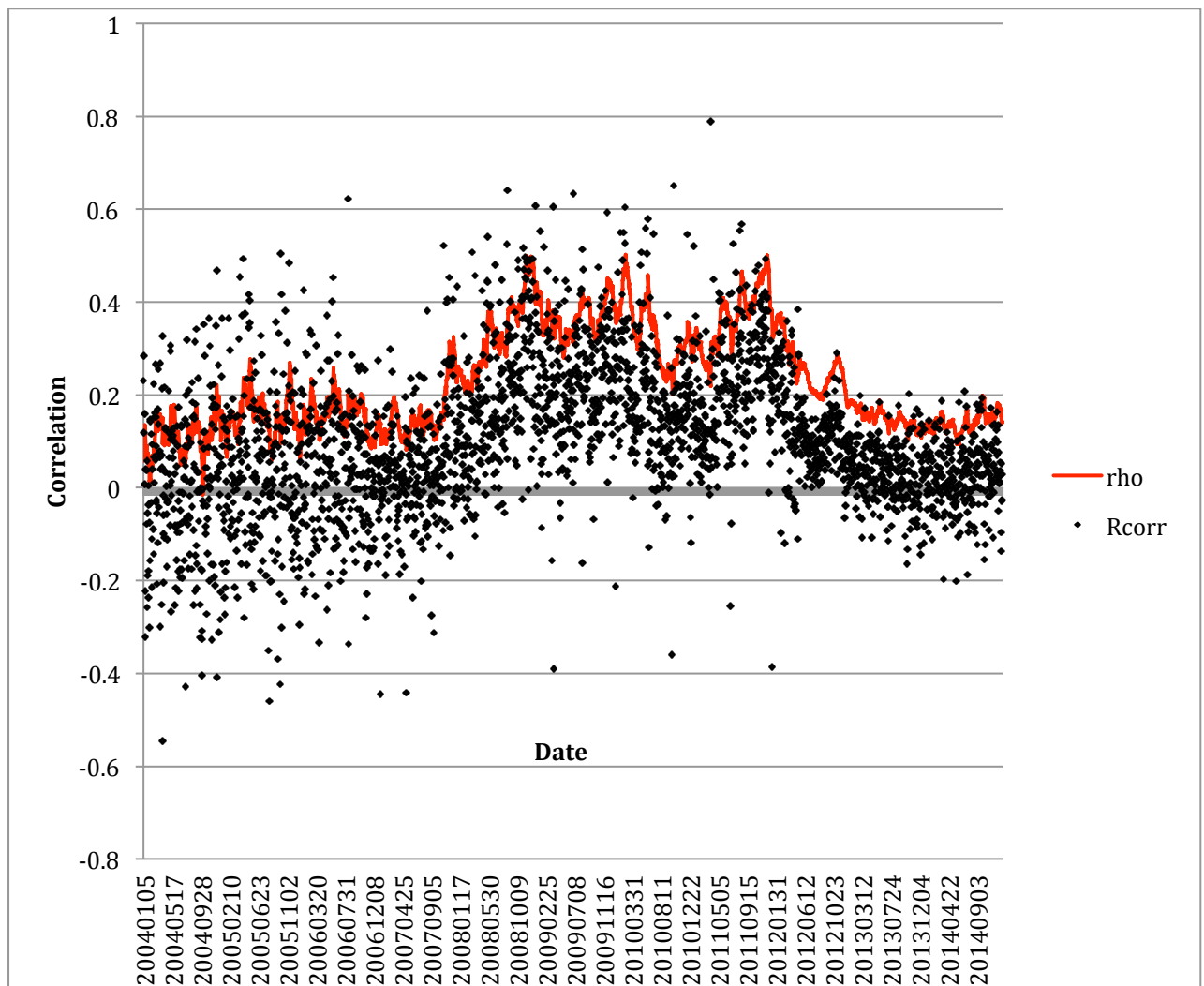
### Conditional and Realized Variance of Corn from 2004 to 2014



This figure shows the realized variances of corn that can be directly plotted based on the variable  $x_{\text{corn},t}$  in the dataset and the conditional variances that we estimated by the Realized Beta GARCH model in section 4.1.

Figure 4.

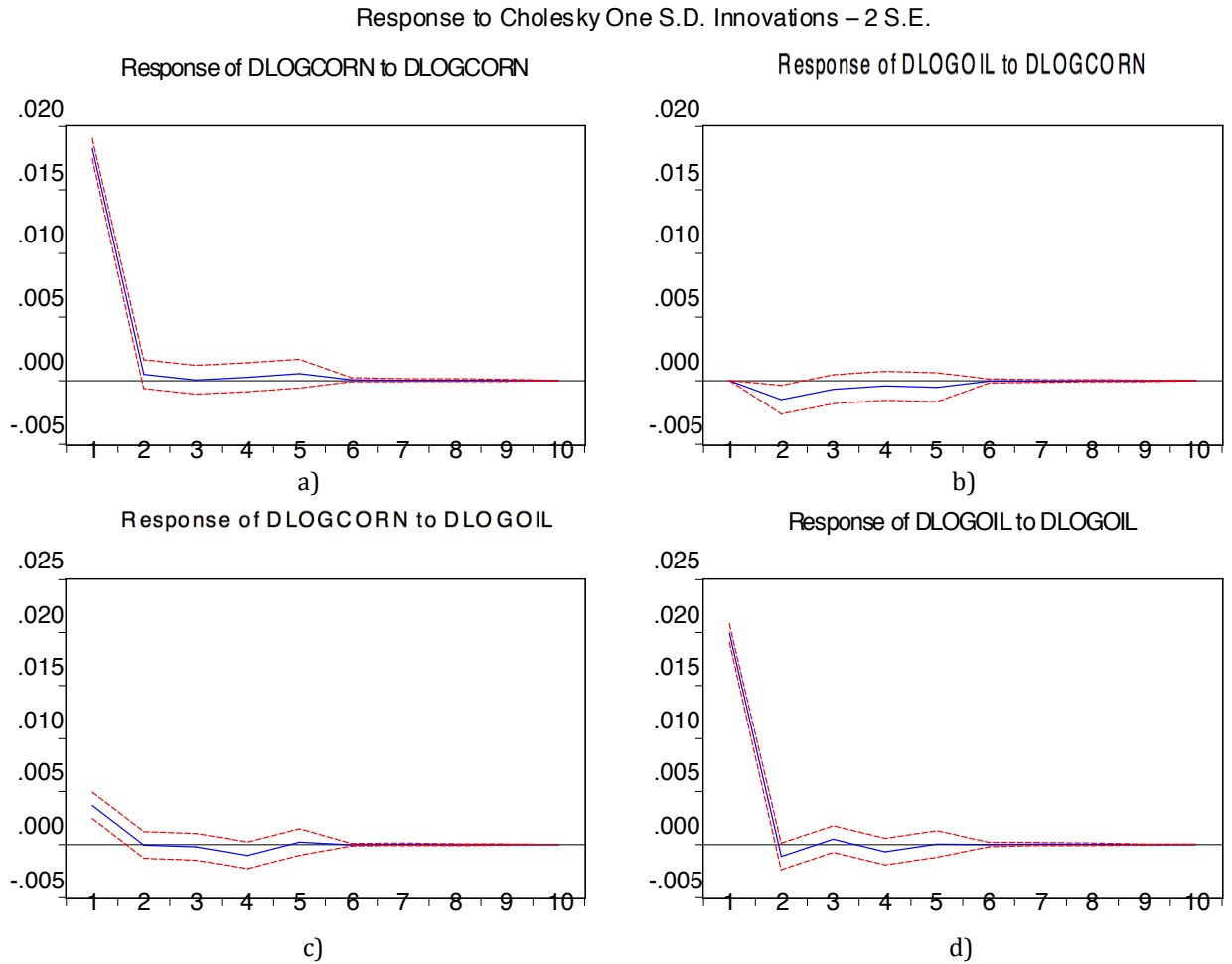
Correlation between Oil and Corn from 2004 to 2014



This figure shows the realized correlation between oil and corn that can be directly plotted based on the variable  $y_{oil,corn,t}$  in the dataset and the conditional correlation that we estimated by the Realized Beta GARCH model in section 4.1.

Figure 5.

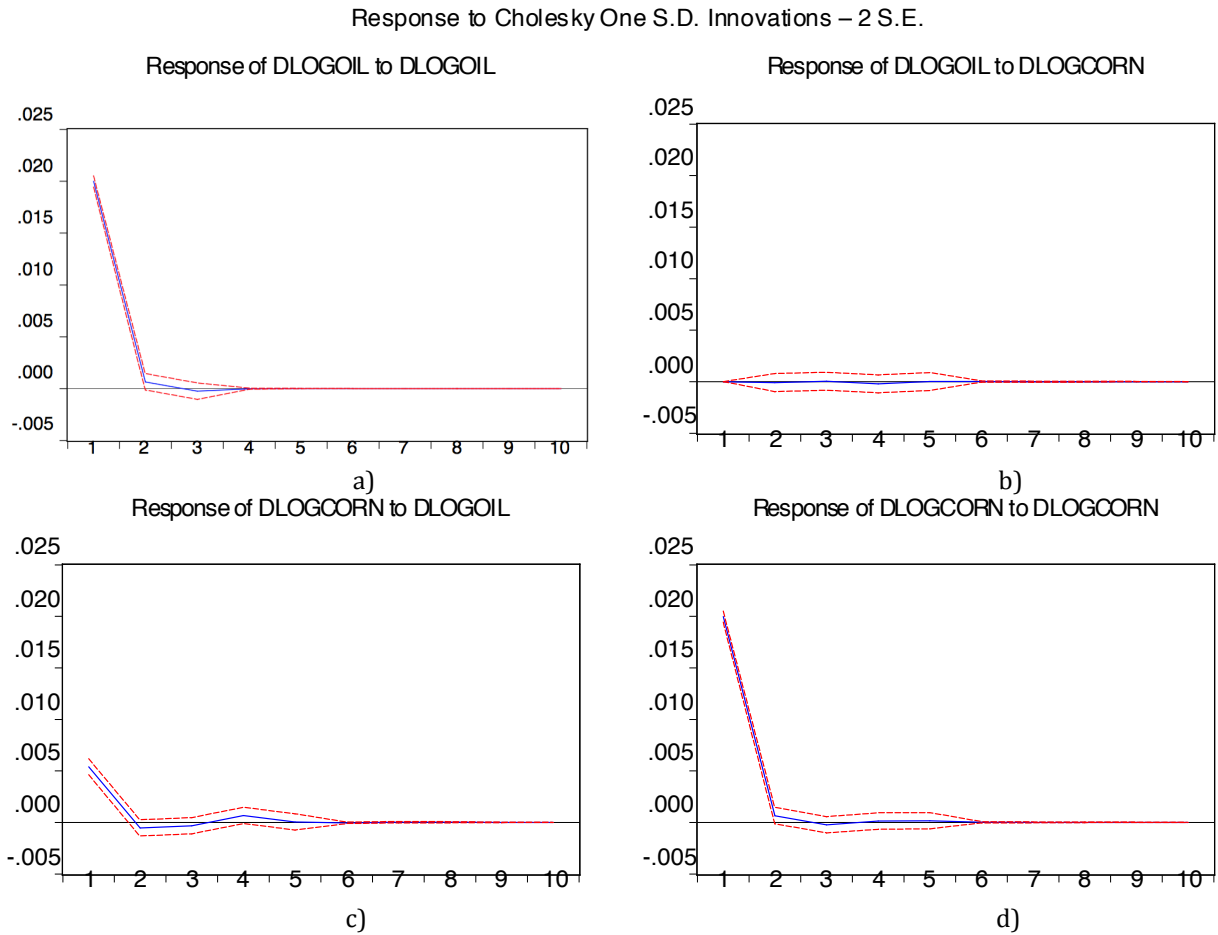
### Impulse Response Functions for the Full Data



This figure shows the plot of the responses of one commodity in terms of percentage change in its prices to the on-time price shock of the other commodity from one up to ten time periods after the shock happens under the full dataset. They are generated from impulse response functions defined in section 4.3.2 and directly referred to equations (13) and (14).

Figure 6.

Impulse response functions for the first period: Jan 05, 2004 to Apr 01, 2008

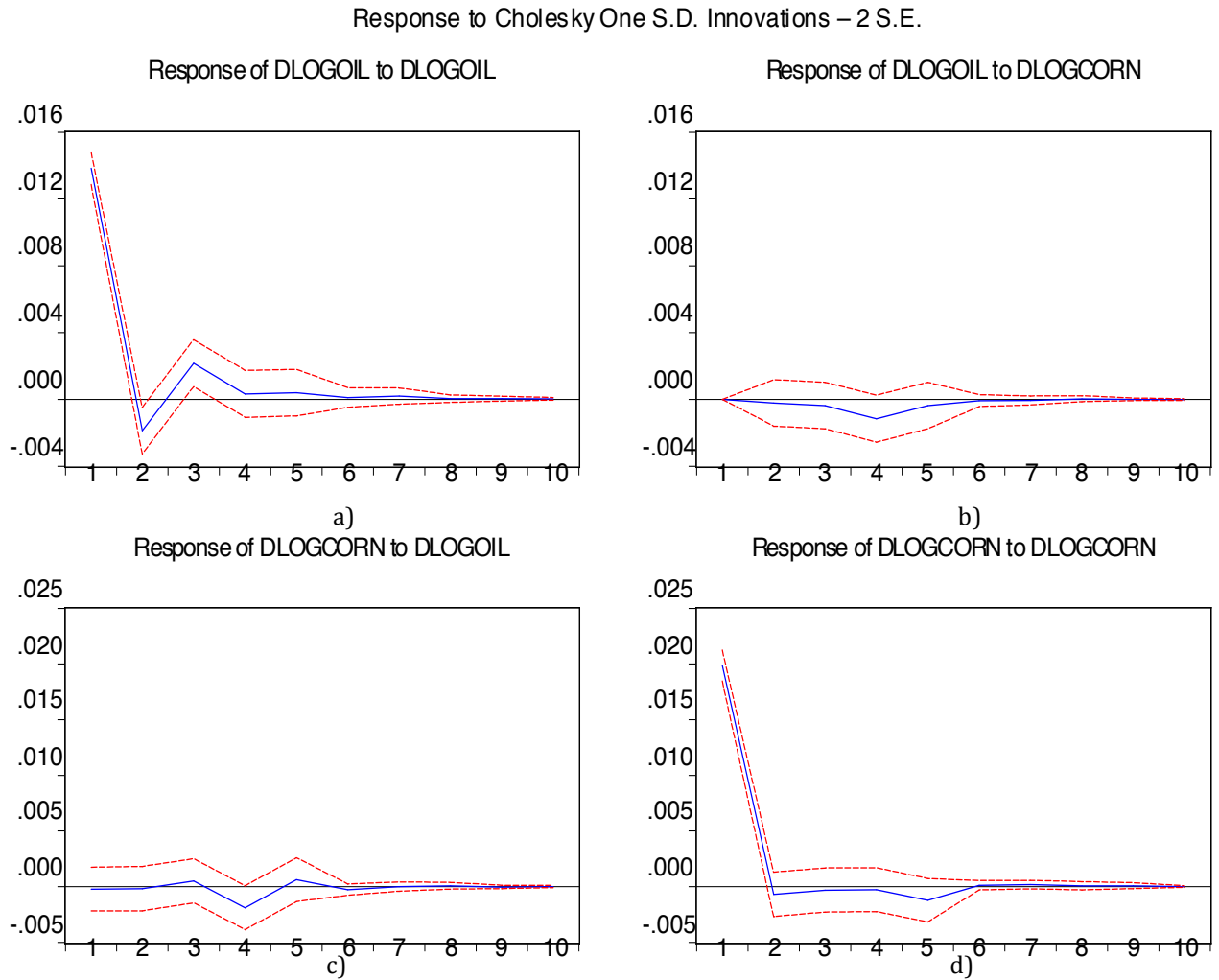


This figure shows the plot of the responses of one commodity in terms of percentage change in its prices to the on-time price shock of the other commodity from one up to ten time periods after the shock happens in the first regime. They are generated from impulse response functions defined in section 4.3.2 and directly referred to equations (13) and (14).



Figure 7.

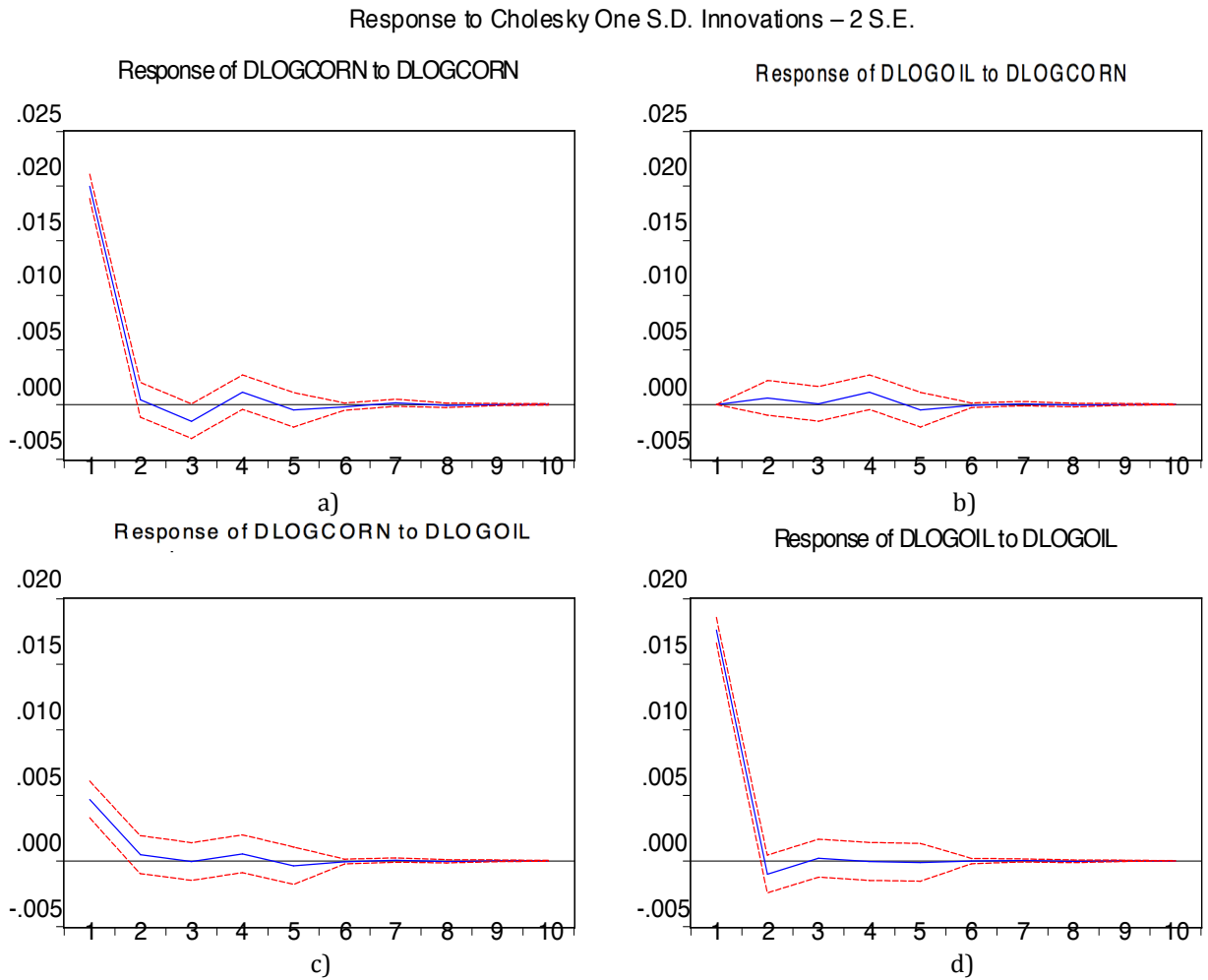
Impulse response functions for the second period: Apr 02, 2008 to Oct 08, 2010



This figure shows the plot of the responses of one commodity in terms of percentage change in its prices to the on-time price shock of the other commodity from one up to ten time periods after the shock happens in the second regime. They are generated from impulse response functions defined in section 4.3.2 and directly referred to equations (13) and (14).

Figure 8.

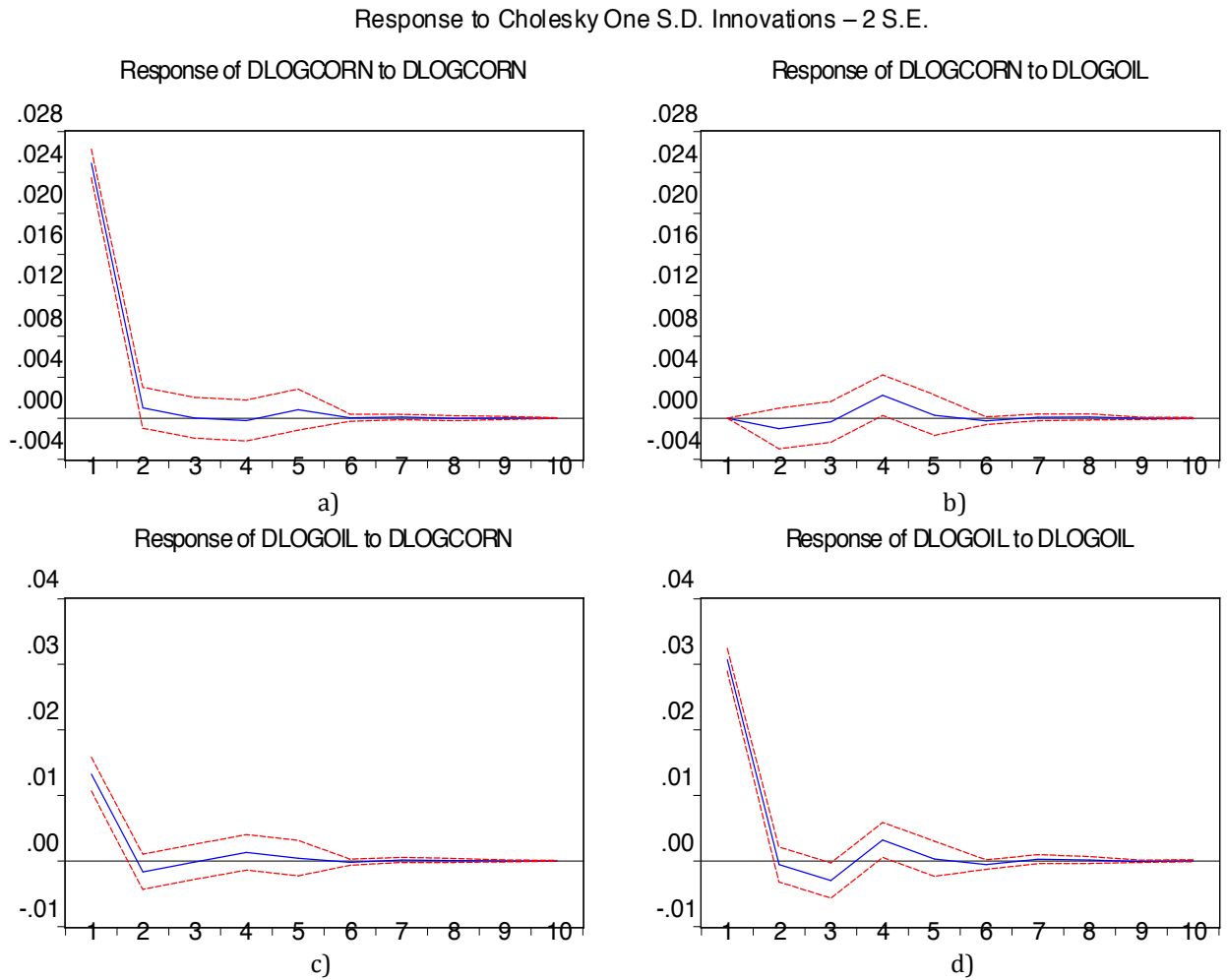
Impulse response functions for the third period: Oct 11, 2010 to May 13, 2013



This figure shows the plot of the responses of one commodity in terms of percentage change in its prices to the on-time price shock of the other commodity from one up to ten time periods after the shock happens in the third regime. They are generated from impulse response functions defined in section 4.3.2 and directly referred to equations (13) and (14).

Figure 9.

Impulse response functions for the forth period: May 14, 2013 to Dec 31, 2014



This figure shows the plot of the responses of one commodity in terms of percentage change in its prices to the on-time price shock of the other commodity from one up to ten time periods after the shock happens in the forth regime. They are generated from impulse response functions defined in section 4.3.2 and directly referred to equations (13) and (14).