An Exploration of the Relationship Between Corn and Crude Oil Prices

FRE 501

Wicked Case

Elton Cheung
(47306162)

Background

The linkage between energy and food prices has become a prime factor in determining the affordability and accessibility of food. The strength of this food energy linkage, influenced by prevailing oil prices, has fluctuated throughout the years. However, this correlation between energy and food prices has been proven statistically significant at relatively high crude oil prices (Chen et al., 2014). As a result, periods of relatively high energy prices will affect the price of agricultural commodities, and this brings about a whole host of consequences.

When the volatility of energy markets spills over to the agricultural markets, this negatively affects consumer access to food products. In particular, the food security of low-income households is greatly impacted by significant volatility in food prices. When the price of food increases, this leads to a fall in real income and consumption. Moreover, low income consumers spend the highest proportion of their income on food; therefore, they are the most susceptible to spikes in food prices (Kalkuhl et al., 2016).

The negative impact of volatile food prices on society's welfare is not limited to consumers. Producers also face unfavorable effects from the volatility in food prices. With a greater fluctuation in prices, farmers face an increased likelihood of losses (Kalkuhl et al., 2016). Agricultural producers operate with significant capital asset expenditure. A rapid depression in prices negatively impacts the ability of farmers to meet financing obligations. Therefore, increased volatility in food prices generates uncertainty, which poses a challenge for farmers making resource allocation decisions. This leads to farmers cutting back on productive investments, which lowers potential future revenue and food availability (Kalkuhl et al., 2016).

Study Objectives

This paper aims to explore the link between energy and agricultural commodity prices. Specifically, the prices of Corn and Crude oil will be used as a proxy to investigate the presence and strength of the link between food and energy prices.

To test the long run relationship between Corn and Crude oil prices, an Ordinary Least Squares regression will be conducted on the weekly prices of corn and crude Oil from 2000-01-03 to 2019-08-19. This process firstly involves testing the price of corn and crude oil for stationarity. It is crucial to identify non-stationarity as this would lead to spurious results. In the presence of non-stationary data series we will test if the first difference of each series is stationary and both price series are integrated of the same order. Next, a test of cointegration will be carried out to determine if there is a stationary linear combination of the corn and crude oil price series.

Lastly, to explore the short run dynamics of corn and crude oil prices, an Error Correction Model will be used to estimate the speed of adjustment of corn prices back to the long run equilibrium after a temporary shock to crude oil or corn prices.

Literature Review

Firstly, Zhang et al. (2010) explored the link between agricultural commodities and crude oil from 1989 to 2008. This study concludes a statistically insignificant relationship between agricultural commodities and energy prices over a span of 19 years. However, the methodology employed in this study provides insight into the methods required to carry out a time series analysis for a more recent time period. Firstly, a Dickey Fuller test is utilized to test the unit root of the price series and determine if the data is nonstationary. Next, the first difference of each series is examined for stationarity. An Akaike Information Criterion is used to provide an accurate selection of lag periods. Next, a Johansen test of cointegration is employed to determine if the price series have a long run linear relationship. Lastly, a Vector Error Correction model is utilized to explore the short run dynamics of dependent variables returning to a long run equilibrium.

The second study by Natanelov et al. (2013) investigates the link between Crude oil and Corn prices for a more recent time period of 2006 to 2011. The authors highlighted the influence of high crude oil prices increasing the demand for corn. This positive relationship between Crude oil and Corn was established between 2008 and 2011, where crude oil prices exceeded $75 per barrel. The authors determined that at that price threshold, the conversion of corn into ethanol became a competitive substitute for fuels derived from refining crude oil. To determine the long run relationship between corn and crude oil prices, the authors utilized a Johansen test for cointegration to deal with non-stationary time series data. Additionally, the short run dynamics between corn and crude oil prices were explored using a Vector Error Correction Model.

Lastly, Chen et al. (2010) approximately combines the time frame of analysis from the earlier two studies. Looking at data from 1983 to 2010 and dividing this timeline into four distinct periods, the authors arrived at a similar conclusion. The relationship between corn and crude oil prices is only significant in the periods of high oil prices. Notably, the period from 2005 to 2008 saw a 1% rise in crude oil prices leading to a statistically significant 3.33% increase in corn prices. The authors attribute this price linkage to supply and demand factors. On the demand side, the explanation is similar to Natanelov et al. (2013). Corn faces an increase in derived demand from biofuel production, and this leads to a rise in corn prices. On the supply side, rising oil prices increase the cost of production for farmers who utilize oil as an input in farming operations. Regarding the methodology of this study, a Global Cropland Allocation Model was used to examine the long-term link between agricultural commodities and oil prices. While this model differs from the planned approach of our paper, useful methodological insights can still be gained from Chen et al. (2010). In particular, the Augmented dickey fuller test is proven to be a stringent test for stationary in time series data.

The literature reviewed informs our methodology and expected results in our investigation of the relationship between Corn and Crude oil prices.

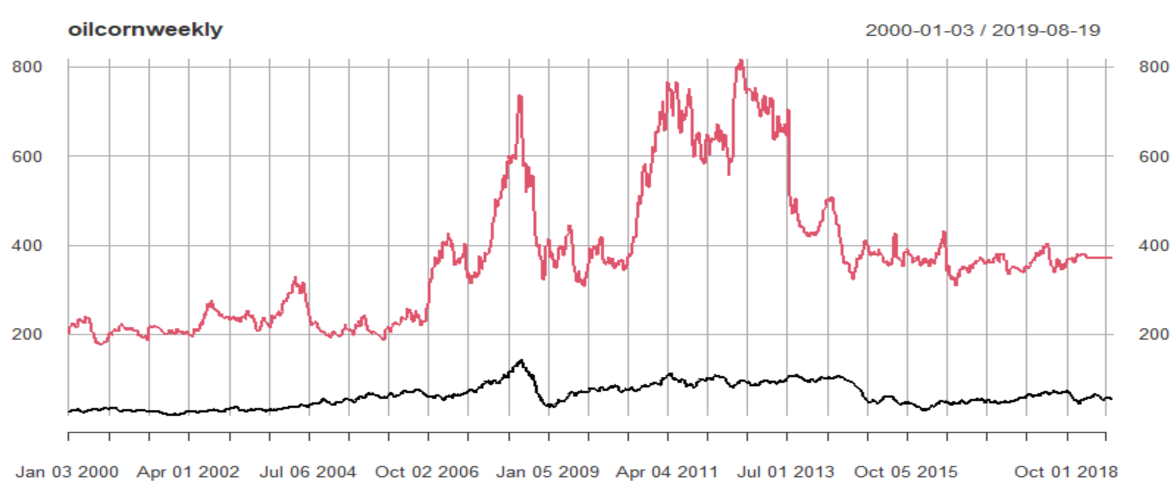
Our approach in determining the long run relationship between corn and crude oil prices will be built on the methods utilized by the studies above. This would involve a test for stationarity to identify non-stationary data. This would be followed by a test of cointegration to determine the validity of any long run relationship findings. Lastly, an error correction model will be applied to assess the short run dynamics between agricultural commodities and energy prices. However, our study will involve a more recent time span from 2000 to 2019. Additionally, the Engle-Granger test will be used to determine the statistical cointegration of crude oil and corn.

Based on the literature reviewed, we would expect a positive long run relationship between corn and crude oil prices. This can be attributed to the price of oil affecting the cost of production for farmers and demand factors influenced by the substitutability between corn-based biofuels and refined crude oil fuels. However, we would only expect to see a statistically significant long run relationship between corn and crude oil if prevailing energy prices remain relatively high.

Data and Econometric Methods

Data Overview

The data for oil and corn prices were downloaded from Investing.com. The oil data can be found in the website under Crude Oil WTI Futures and the corn data can be found under US Corn Futures. The data collected ranges from Jan 3, 2000 to Aug 19, 2019 and was collected from the website as daily data. For the purpose of this paper, the data were converted into a weekly format where the prices in a week were averaged. To do these three new variables were created that identified the year, week and day of the week using the included date format. Then new datasets (one for each commodity) were created that only included prices from a Monday or Tuesday. Then two more datasets (one for each commodity) were created that averaged the prices by week and year. Using these created datasets they were combined together by commodity type, and by matching the year and week. This created the weekly oil and corn datasets which will then be converted into time series data and merged together into one big dataset called “oilcornweekly” by matching the year and week. Furthermore, when merging, the missing data in the prices, due to mismatching dates, were filled in using the previous week’s prices.



Here we see the red line are the corn prices and the black line are the oil prices over time. From this plot we see that there are two distinct humps for corn, one from July 2007 to Jan 2009 and the other from July 2010 to July 2013. For oil prices there is only one distinct hump from July 2007 to Jan 2009. This visual depiction is in line with the information described in the literature review. Prices before 2005 appear to be not correlated, where corn price fluctuations appear independent compared to the relatively gradual increase in oil prices. However, the sharp peak in oil prices in 2007 coincides with a similar spike in corn prices.

Economic Model Overview

In this paper we will be estimating the relationship between the price of corn and the price of oil. The underlying theory for this relationship is that corn is used in the production of ethanol fuel, and ethanol is a substitute for oil. Therefore, we will be using oil prices as the explanatory variable in order to estimate corn prices. This allows us to measure how well oil prices are able to forecast the prices of corn and estimate the speed at which corn prices react to changes in oil prices.

Statistical Methods

The first statistical method used is to check if the data is stationary. This is important because if the data was determined to be non-stationarity then it is considered unpredictable and cannot be forecasted. This means the model would get spurious results with no real meaning (Erica, 2020 Sept 1). To test for stationarity the Augmented Dickey-Fuller (ADF) test will be used over the Dickey-Fuller (DF) test. The ADF test builds on the DF test by addressing autocorrelation in the $\Delta y_t = \mu + \beta t + \delta y_{t-1} + \varepsilon_t$ regression.

This is because unlike the DF test, the ADF test allows us to include lags in the model to deal with autocorrelation (Erica, 2020 Sept 1). To select the amount of lags in the model the Acf() and the Pacf() functions will be used. These functions create plots that show the autocorrelation between current prices with lagged prices for a commodity. In the ADF test, when vertical black lines exceed the horizontal significance lines, it highlights the presence of autocorrelation. Since the Pacf() function isolates which lags are more relevant in today's prices, it will be used to determine the amount of lags to include by identifying where the vertical black lines exceed the horizontal significance lines. Once the ADF tests are conducted for each of the commodity prices the generated t statistics of the lagged prices are compared with the tau3 critical values. Where if the t statistic is greater (less) than tau3 we fail to reject (reject) the null hypothesis of a unit root, and therefore conclude that the data is non-stationary (stationary).

The second method is testing for integration using an ADF test on the first difference of corn and oil prices. This allows us to test if the two non-stationary time series data are integrated of the same order (Erica, 2020 July 21). When the two individual non-stationary data are determined to be integrated on the same order then a linear combination of them is stationary (Erica, 2020 July 21). However, before running the ADF test we will again use the Pacf() function on the first difference of corn and oil prices to determine the amount of lags to use. When running the ADF test on the first difference for each commodity this will produce t statistics and tau3 critical values. Where if we fail to reject (reject) the null hypothesis of a unit root and conclude that both series are integrated (not integrated) of the same order.

The third method uses the Engle-Granger test. The first step of the Engle-Granger test is to run the following regression $p^{corn} = \alpha + \beta_1 p^{oil} + \varepsilon$, in order to estimate the cointegrating vector $\alpha$and β. Furthermore, determine if the coefficients are significant or not. The second step of the Engles-Granger test is to test the residuals of the regression using the ADF test to see if the linear combination of corn and oil prices are stationary or not (Erica, 2020 July 21). Here will be also using the Pacf() function to determine the amount of lags to include. If we reject the null hypothesis of a unit root of the residuals there would be evidence for cointegration and thus a long-run relationship.

The fourth and final method is the Error Correction Mode (ECM). Once we determine that the linear combination is cointegrated, meaning we would not get spurious results, we can run the ECM to estimate the speed of adjustment of Corn prices back to the long run equilibrium after a temporary shock to crude oil or corn prices.

The initial step in the ECM is to generate the residuals from the first stage regression: $p^{corn} = \alpha + \beta_1 p^{oil} + \varepsilon$. The residuals $\varepsilon = (p_t^{corn} - \alpha - \beta_1 p_t^{oil}\ )$ are then collected to run the following regression and estimate $\Delta P_t^{corn}\ =\ \beta_1 + \beta_2 \Delta P_t^{oil} - \lambda(p_t^{corn} - \alpha - \beta_1 p_t^{oil}) + \varepsilon$. Where $\beta_2$determines the short-run effect of a change in P_t^{oil}on a change in P_t^{corn} and λ represents the speed of adjustment in P_t^{corn}(Krisha's Error Correction Model Notes).

The simple ECM model is then expanded to include additional lags on the price of corn and crude oil (Krisha's Error Correction Model Notes).

$$\Delta P_t^{corn}\ =\ \beta_1 + \beta_2 \Delta P_t^{oil} + \beta_3 \Delta P_{t-1}^{corn} + \beta_4 \Delta P_{t-1}^{oil} - \lambda(p_t^{corn} - \alpha - \beta_1 p_t^{oil}) + \varepsilon$$

More specifications and variables are added to this model to capture more dynamics that might be going on. For example, we wanted to check if price shocks in past periods might cause an impact at a later point. Additionally, the inclusion of the lagged difference in the price of corn and the price of crude oil will work to address any autocorrelation that might be present in our data.

From the expanded ECM model, $\beta_4$tells us the effect of a change in oil prices 1 period ago on the change in corn prices in time t. Similarly, λ represents the speed of adjustment in P_t^{corn}. Lastly, ΔP_{t-1}^{corn} was added to address autocorrelation and β_3 tells us how the change in corn prices in a previous time period affects the change in corn prices today (Krisha's Error Correction Model Notes).

Results

Stationarity Tests for Crude Oil and Corn ***Price Series***

Table 1: ADF test for Unit Root

Lagged price, y_{t-1}	Oil	Corn
Estimated coefficient and SE	-5.886e-03* (2.962e-03)	-0.0066905* (0.0032958)
t statistic	-1.9872	-2.03
Critical t, DF Table	-3.41	-3.41
Ho: Unit Root	Fail to Reject	Fail to Reject

An Augmented Dickey Fuller test was carried out to test each price series for stationarity. The input lags for Crude Oil were obtained from Appendix 1 and the lags from Corn were obtained from Appendix 2. Using the Pacf() function it was determined that 2 lags be included for each commodity when running the ADF test in order to correct for autocorrelation. From Table 1, the t statistic of the lagged price for oil and corn are -1.99 and -2.03 respectively. Which is bigger than the critical value of tau3 at -3.41. This means we fail to reject the null hypothesis of a unit root at a 5% significance level. We conclude that corn and oil prices are non-stationary.

Stationarity Tests for Each Commodity's First Difference

The results of and ADF stationarity test on the first difference of Corn and Crude oil is listed in Appendix 3. Using the Pacf() function on each commodities first difference, it indicated that 2 lags were needed in order to correct for autocorrelation (Appendix 3). When we do a ADF test on the first difference of oil and corn prices, the test statistics are -25.9697 and -27.0552, respectively (Appendix 4). We reject the null hypothesis of a unit root for both series.
and therefore, conclude that both series are integrated of order 1. We can now proceed with the test for cointegration

First Step of Engle-Granger Test

Table 2: Step 1 Regression Output of Corn Prices Against Crude Oil Prices

	Dependent variable: p_corn
p_oil	4.562***
	(0.113)
Constant	93.552***
	(7.620)
Observations	1,062
R^2	0.606
Adjusted R^2	0.606
Residual Std. Error	96.952 (df = 1060)
F Statistic	1,632.455*** (df = 1; 1060)
Note:	*p<0.5; **p<0.01; ***p<0.001

When running the first step of the Engle-Granger test the results in Table 2 show that the a 1 unit increase in the price per barrel of crude oil prices will mean the price of corn will increase by 4.562 units per bushel, on average.

Second Step of the Engle-Granger Test

Based on Appendix 5, the plot of residuals can be examined in 3 distinct periods. The first third of the residual plot show the residuals distributed approximately evenly about the mean. A similar distribution is observed in the final third of the residual plot. However, the middle section of the residual plot shows a slight increasing trend that is followed by a decreasing trend in the pattern of residuals.

Table 3: Stationarity Test on Residuals

Lagged residual, y_{t-1}	Residual Corn Oil
Estimated coefficient and SE	-0.016522** (0.005037)
t statistic	-3.2801
Critical t, DF Table	-1.95
Ho: Unit Root	Reject null
Note:	*p<0.5; **p<0.01; ***p<0.001

When we run the DF test on the residuals of the regression, we get a test statistic of -3.2801, as observed from Table 3. Since this is less than the t critical of -1.95, we reject the null hypothesis of a unit root of the residuals, and now have evidence of cointegration at the 1% significance level.

Error Correction Model

Table 4: Error Correction Model Results

	Corn (Y_t)	
Dep Variable	Simple Crude Oil	Expanded Crude Oil
Intercept	0.113728 *(0.439005)*	0.096109 *(0.432427)*
β_2	1.546751*** *(0.185158)*	1.517084*** *(0.187943)*
β_3		0.174196*** *(0.030499)*
β_4		-0.597188** *(0.191227)*
λ	-0.008294° *(0.004542)*	-0.010615* *(0.004488)*
R^2	0.06135	0.09067
Observations	1062	1062
Note:	$p° < 0.1$	$^{*}p<0.5$; $^{**}p<0.01$; $^{***}p<0.001$

The results in Table 4 show that for the simple error correction model that the coefficient of ΔP_t^{oil} is 1.546 and statistically significant at the 0.1% level. This suggests that the change in crude oil does change the price of corn within the same period. The estimate on the speed of adjustment λ is -0.008 and statistically significant at the 10% level. This suggests that when the equilibrium relationship experiences a shock, corn prices adjust to bring the relationship back to the long-run equilibrium, but at a relatively slow speed.

For the expanded model, the coefficient for β_2 is -1.517 and statistically significant at the 0.1% level. This coefficient is still similar to the β_2 from the simple model. The coefficient for the change in crude oil a period ago, β_4, is -0.597 and statistically significant at the 1% level. Notably, the speed of adjustment λ has increased to -0.0106 and increased its significance to the 5% level.

Conclusion

This paper attempted to establish the long run relationship and short run dynamics of Crude oil and Corn prices. A time series analysis was carried out on weekly corn and crude oil prices for 1062 periods. We began by establishing that each price series was non-stationary. This was followed by determining that both price series were integrated of the same order. An Engle-Granger Test was then used to identify the cointegrating vector β_2 was 4.562 and statistically significant at the 1% level. To confirm the stationarity of the residuals, we carried out an Augmented Dickey Fuller Test which provided evidence to conclude that prices of corn and crude oil were cointegrated. Thus, we concluded that a 1 unit increase in the price per Crude oil barrel will increase corn prices by 4.562 units per bushel.

Short run dynamics were investigated through an Error Correction Model. Based on the value of -0.008294, we concluded that the speed of adjustment in P_t^{corn} to a price shock in crude oil or corn prices would be relatively slow.

Limitations

Like many other papers, this study is subject to several limitations. One limitation being that the number of observations might not be enough to measure the true long-run relationship between crude oil and corn prices. Another limitation is not running any post regression tests for the first step of the Engle-Granger test. As doing this would help make the results more accurate. Furthermore, when cleaning the data, we replaced missing data with the previous week's prices which would also affect the accuracy and reliability of the results.

Policy Implications

Policies should be created with the understanding that agricultural commodities and energy prices have a long run positive relationship. Given the concerns surrounding volatile food prices, caution should be placed on enacting policies that will strengthen the food energy linkage. In particular, policies that promote the use of biofuels should ensure that the cost of volatile food prices are part of the discussion for the overall cost and benefit analysis of the policy.

Future Research

This paper presents opportunities for further expansion. The prices of multiple agricultural commodities can be tested against the price of crude oil. Additionally, the model can be expanded to include multiple explanatory variables. This would be done by regressing the price of multiple agricultural commodities against multiple energy products.

As noted in the literature review section, the statistical significance of the correlation between energy and food prices is dependent on high energy prices. Therefore, further research can attempt to identify the strength of the food energy linkage at different oil price levels.

These methods will be useful in providing a more comprehensive assessment of the varying relationships between agricultural commodity and energy prices.

References

Chen, S., Kuo, H., & Chen, C. (2010). Modeling the relationship between the oil price and global food prices. *Applied Energy, 87*(8), 2517-2525. doi:10.1016/j.apenergy.2010.02.020

Erica. (2020, July 21). A Guide to Conducting Cointegration Tests. Retrieved November 09, 2020, from https://www.aptech.com/blog/a-guide-to-conducting-cointegration-tests/

Erica. (2020, September 11). How to Conduct Unit Root Tests in GAUSS. Retrieved November 09, 2020, from https://www.aptech.com/blog/how-to-conduct-unit-root-tests-in-gauss/

Kalkuhl M., von Braun J., Torero M. (2016). Food Price Volatility and Its Implications for Food Security and Policy. Springer, Cham. https://doi.org/10.1007/978-3-319-28201-5_1

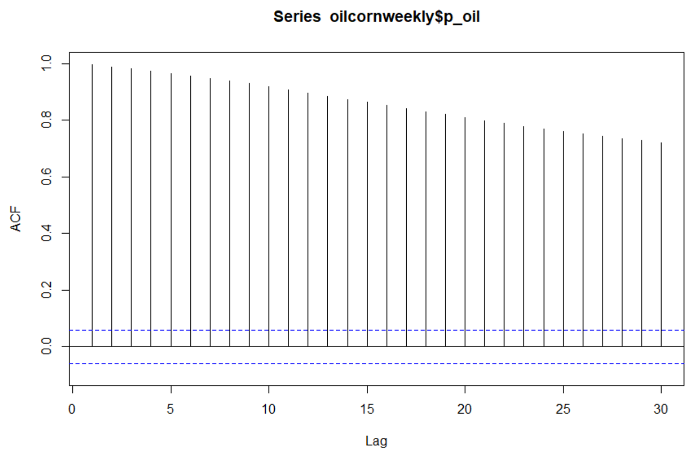
Lim, Krisha (2020). Error Correction Model Notes.

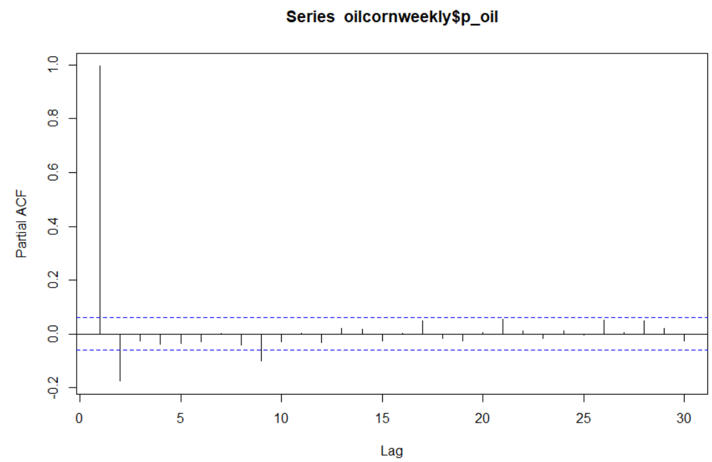
Natanelov, V., Mckenzie, A. M., & Huylenbroeck, G. V. (2013). Crude oil–corn–ethanol – nexus: A contextual approach. *Energy Policy, 63*, 504-513. doi:10.1016/j.enpol.2013.08.026

Zhang, Z., Lohr, L., Escalante, C., & Wetzstein, M. (2010). Food versus fuel: What do prices tell us? *Energy Policy, 38*(1), 445-451. doi:10.1016/j.enpol.2009.09.034

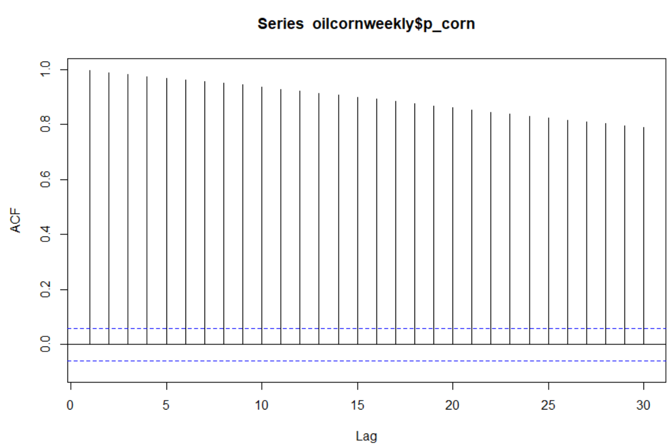
Appendix

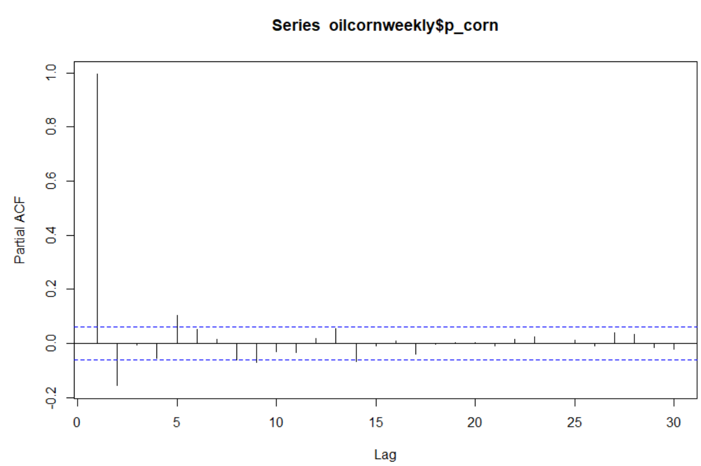
Appendix 1: ACF and PACF for Crude Oil Prices



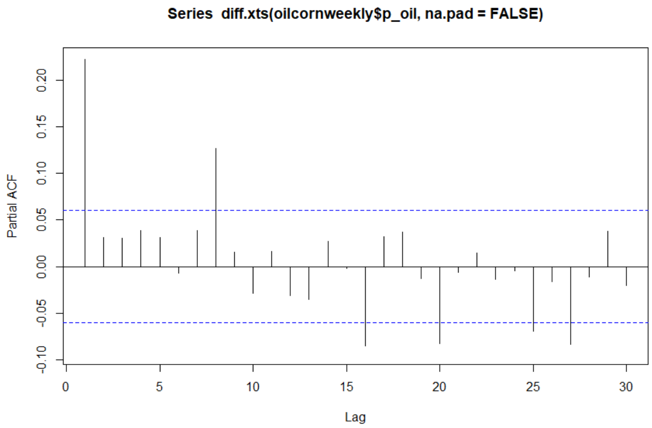


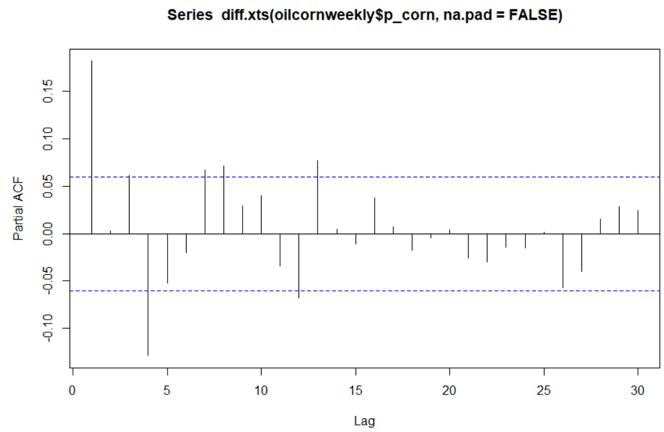
Appendix 2: ACF and PACF for Corn Prices





Appendix 3: PACF to Identify Number of Lags Used in the ADF Test for the First Difference in Crude oil and Corn





Appendix 4: ADF Stationarity Test Results for the First Differences in Crude Oil and Corn

Lagged price, y_{t-1}	Oil	Corn
Estimated coefficient and SE	-0.7322288*** (0.0449021)	-0.7663347*** (0.0466646)
t statistic	-16.3072	-16.4222
Critical t, DF Table	-3.41	-3.41
Ho: Unit Root	Reject null	Reject null

Appendix 5: Plot of Residuals