INTERNATIONAL HELLENIC UNIVERSITY

On the non-linear nexus between natural gas and crude oil prices

Dimitrios Ampatzis

SID: 3302140023

SCHOOL OF SCIENCE & TECHNOLOGY

A thesis submitted for the degree of

Master of Science (MSc) in Energy Systems

DECEMBER 2015

THESSALONIKI – GREECE

INTERNATIONAL
HELLENIC
UNIVERSITY

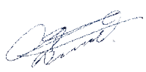
On the non-linear nexus between natural gas and crude oil prices

Dimitrios Ampatzis

SID: 3302140023

Supervisor: Dr Theologos Dergiades

Recoverable Signature

X 

Theologos Dergiades
Dr
Signed by: Theologos Dergiades

Supervising Committee Members: Dr. Dimitrios Psychoyios
Dr. Panos Pouliasis

Abstract

Energy commodities are deeply linked with the global economy and in the last decade its financial derivatives are vastly traded. Thus the examination of the relationship among them is an important task with many implications. This dissertation aims to examine the crude oil and natural gas price relationship while also taking into consideration the non-linear structure of our series. The data used consist of Western Texas Intermediate and Henry Hub spot prices as well as weather and storage shocks for the U.S. The existence of cointegration is examined by implementing the Rahbek & Mosconi (1999) fix to our the underlying VAR model as our system of variables, is a mix of endogenous I(1) and exogenous I(0) variables. After the establishment of cointegration we examine linear causality between crude oil and natural gas prices by applying the Granger causality test. However we don't want to neglect the nonlinearities of the series. The obtained residuals will be tested for i.d.d with the BDS test (1996). We implement the Breitung & Candelon (2006) frequency domain causality test to the delinearized residuals by conditioning on weather and storage shocks. The purpose of this test is to reveal the true causal relationship between natural gas and crude oil prices and surpass the limitations of the standard Granger causality test. Finally, by performing this test we are able to establish whether causality stands in the short-run or in the long-run.

Dimitrios Ampatzis

11/12/2015

Contents

Figures

Tables

1 Introduction

Natural gas and crude oil are two of the most important energy commodities. As the two commodities are linked through demand and supply side, their pricing is a very important task. In the early 90's pricing of natural gas was done with the simple rules of thumb where one million Btu of natural gas was 10 times cheaper than one barrel of crude oil (10-to-1). Later on the burner-tip-parity rule (6-to-1) prevailed where pricing of natural gas represented the difference between crude oil and natural gas energy content. However those rules can fit in historical data but cannot be used in forecasting methods. The natural gas and crude oil price relationship has been examined multiple times in the past and researchers try to find the actual nature of this relationship through the years.

Many papers such as Villar & Joutz (2006), Panagiotidis & Rutledge (2007) and Brown & Yücel (2008) suggest that natural gas and crude oil prices have a long run equilibrium relationship. In the literature this is referred as cointegration. On the contrary Siliverstovs *et al.* (2003), Erdős (2012) and Lin & Li (2015) provide evidence that the two prices have been decoupled, meaning that they are not cointegrated. This disagreement between the results can be attributed to several reasons. One reason seems to be the liberalization of the natural gas markets. US and UK have fully liberalized natural gas markets and in rest of Europe many countries have taken action to deregulate their market. Another reason appears to be the increase in shale natural gas and Liquefied Natural Gas (LNG) production in the US. This led to an oversupply and consequently lower prices in the US. The natural outcome would be the gas-to-gas arbitrage with Europe. However export limitations in the US didn't allow the Atlantic arbitrage to work out. Ji and Fan (2015) state that WTI has no longer the leading role as a benchmark price in the international world market. Adding to that in the past years WTI begun to reflect more the American demand and supply dynamics. On the other hand Henry Hub prices are determined regionally as evidence show that US natural gas prices have been decoupled from the world's natural gas prices following Siliverstovs *et al.* (2003) and Erdos (2012) examination. In this study we will examine the possibility of cointegration between natural gas and crude oil in the US. We will attempt to study the relationship covering daily quotations for a time length as large as possible. In Europe the linkage between crude oil and natural gas is clearer. Although UK was the first to deregulate the natural gas market and is one of the biggest markets, the rest of European countries must further proceed with deregulation. As Asche *et al.* (2013) and Ji *et al.* (2014) state the oil indexation of natural gas price is a fact. A large share of natural gas trade is done with long-term contracts whose price is determined by Brent crude oil price. Thus the presence of cointegration in Europe is valid.

In order to make a complete study of the crude oil-natural gas relationship, we must consider the drivers of natural gas price. These drives can be cooling and heating degree days, deviations from the normal, inventories, disruption in production from hurricanes etc. Those variables are usually stationary and are treated as exogenous in the system used. Brown & Yücel (2009), Hartley & Medlock (2014) and Nick & Thoenes (2014) indicate that these stationary exogenous variables can affect natural gas prices in the short-run. Extreme weather conditions or low inventories can be added to the deviations from the long-run equilibrium and thus extent it. We chose to include weather shocks and inventory shocks in our study because of their immediate and great effect on the natural gas prices. Following Mu (2007) speciation, degree days and working natural gas storage were transformed into shocks. It is important to incorporate those variables in our analysis for cointegration but also for causality-testing.

In the literature there is evidence of one-way causality that has a direction of crude oil to natural gas prices. Causality examination is very useful because it shows the predictive power tone that a variable may have over the other. Most of the studies implement the Granger causality method for their study and the linear relations of the two prices is revealed (see Mohammadi 2011 and Brown & Yücel 2008). However nonlinear structure in the financial markets cannot be ignored. Apart from the commonly used linear Granger causality test, Bekiros & Diks (2008) applied a nonparametric test for nonlinear causality between WTI crude oil spot and future prices. Through Vector Error Correction Model (VECM) filter residuals they confirmed strictly non-linear bidirectional causality. In another study Dergiades *et al.* (2013) establish causality from Henry Hub natural gas futures towards natural gas spot market while conditioning weather shocks. Furthermore they were able to identify that causality is in the long-run components of the NG futures. For the linear causality the standard Granger method will be applied while the Breitung & Canelon (2006) frequency domain approach is used to capture nonlinearities in the data. In order to investigate causality in a strictly nonlinear framework we use the delinearized series from a VAR model. Finally this frequency domain approach provides evidence if causality stands in low or high-frequencies.

The rest of the paper is structured as follow: Section 2 provides an overall review of the recent literature which examines the relationship between energy commodities prices (coal, crude oil, natural gas, electricity) as well as other aspects of the energy commodities cointegration and causality relations. Section 3 displays analytically the methodological framework implemented and Section 4 the data sources. Finally Section 5 provides results from the econometric tests while in Section 6 the conclusions and some possible implications are discussed.

2 Literature Review

Through the years many researchers have focused their attention on identifying the relationship between natural gas and crude oil prices. As Villar & Joutz (2006) point out, there are several economic factors that connect natural gas and crude prices from both demand and the supply side. From the demand side the two commodities are substitutes, so if one price increases then industries can switch their input to the alternative one. From the supply perspective, an increase in oil prices may affect natural gas prices both negatively and positively. The two markets have been subjected to several changes and in different regions so studies cover this relationship from different perspectives. The literature examining the relationship among energy prices is vast but most of the studies are based on cointegration and causality to reveal the long-run as well as the short-run dynamics of the relationship. Energy commodities are an important aspect of the global and regional economy. Thus studies also examine the relationship between energy prices and economic factors where cointegration and causality are again the main concern of the investigation.

Villar & Joutz (2006) studied the relationship between WTI spot oil prices and Henry Hub spot natural gas prices using seventeen years of monthly data from 1989 to 2005. They give a great emphasis on the properties of the non-stationarity of the series so as to avoid spurious results. All series were found to be non-stationary I(1) by implementing the Augmented Dickey-Fuller (ADF) test. A bivariate Vector Autoregressive (VAR) is estimate in order to begin their empirical analysis. Empirical results show that natural gas and crude oil prices have a long-run relationship as the Johansen (1988) test is applied. The short-run deviations from the long-run relationship between the two prices was examined using an Error Correction Mechanism (ECM) so as to take into account other variables such as seasonality, weather, working inventories, shifts, outliers, heating degree and cooling degree days. The estimation of the coefficients of the ECM model was done and Villar & Joutz (2006) conclude that the effect of oil prices to natural gas demand is great but supply is not clear enough. Another expected finding is that oil prices can influence natural gas prices but not the other way around.

It is well know that natural gas and crude oil have been substitutes, however the past years the number of facilities that are able to switch between those inputs has been decreasing. Brown & Yücel (2008) using an ECM, study the factors that are considered as drivers of natural gas such as weather, shut-in production, seasonality, natural gas storage, so as explain the movements of its price. The past years the energy industry used simple rules of thumb in order to price natural gas in relation to the prices of oil. For example the 10 to 1 and the burner-tip parity rules were pricing methods but neither was efficient and could describe the observed relationship with oil prices. Brown & Yücel (2008) use weekly data for WTI and HH prices, 4 weather variables in order to capture the effect of weather and seasonality, a shortage differential and a shut in production variable. Firstly they start with the ADF test for the stationarity of the series. As expected oil and natural gas prices are non-stationary were the other series are stationary. The Johansen (1995) approach to cointegration is used with and without the stationary exogenous variables to provide evidence that the two prices have a long run relationship. So as to capture the short run dynamics due to deviations from the long run relationship this paper estimates two error correction models one with the stationary exogenous variable and one without. Both models provide evidence of causality from oil prices to natural gas prices. An interesting observation is that in the model without the stationary variables gas prices adjust to close the gap at a rate of 5.77% a week while the second model the price adjustment is 12% a week. The additional stationary exogenous variables that the second model includes have increased the rate of adjustment. Finally those exogenous variables excluding cooling degree days are significant at the 0.01 significance level. These findings indicate that the price of natural gas will be higher if there is great deviation in the heating days and if storage is above seasonal norm then this will result to lower natural gas price. This study exhibits findings of a long run cointegration relationship between the two prices but also in the short run natural gas prices are driven by some other factors like seasonality, storage, weather and disruptions in production.

The drivers of natural gas prices and its implications with crude oil prices are also examined by Nick & Thoenes (2014) by estimating a VAR model for the German market. This study focuses on the natural gas price determinates and the interactions with the prices of other energy commodities. In order to make a complete study other variables must be examined so as to observe their effects during a sample period 2008 to June 2012. This time period was chosen because three major supply disruptions take place: the Russian-Ukrainian gas dispute of January 2009, the production shut-ins of 2011 in Libya and the cut in Russian gas supply in beginning of 2012. Variables included in the analysis are: supply shortfall, price of Brent crude oil, price of coal, storage, LNG imports, Heating degree days deviation and natural gas price. Nick & Thoenes (2014) generate impulse response functions to observe the effects of the variables to natural gas prices. It is clear, from the results obtained, that deviation from the normal heating degree will have a strong and direct increase in natural gas price. Disruption in supply will also have an increasing effect on the price of natural gas. Interaction with the other two commodities prices is consistent with previous literature. Coal prices impact takes place instantly in stable manner but the impact of natural gas prices appears after a delay. On the other hand the LNG import shocks don't provide clear information but a positive storage shock will increase gas prices. Variance decomposition is also performed to estimate the percentage of variance of the natural gas that is attributed to the different variables. In the short-run 37% of the natural price fluctuation is because of the supply disruptions and temperature deviations from the normal. Also storage variations have an important short-run effect while the LNG variations seem to be week. Where in the medium-run major role have coal price variations where in the long run oil price variations. If the forecast is estimated to a year then 67% of the variation of NG price is due to coal and oil prices. At the last section of the paper the three major supply disruptions are examined and the main fact is that there was an overestimation of the price of NG during those years. This happened because simultaneous demand conditions occurred such as reduced demand due to economic crisis, increased demand for storage due to spread of war in Arabic countries and increased demand due to extreme cold temperatures.

Hartley *et al.* (2008) examine the long run relationship between natural gas prices, residual fuel oil price and changing in electricity generation technology. The technology variable is included because they state that electricity generation plays an important role in effecting the relative prices of the energy commodities. They further state that electricity producers, so as to minimize their cost as the fuel price times the heat rate; they choose to switch to alternative fuels. The data consist of monthly prices of Henry Hub natural gas, wholesale residual fuel oil, WTI crude oil expressed in real $200/MMBtu and a heat rate variable. A VECM is implemented to include exogenous variables, such as inventory, weather and production disruption while allowing capturing the short-run deviations. The authors of the paper also constructed a VAR model of natural gas, WTI, residual fuel and relative heat rate so as to provide better information of the link between prices and the heat rate variable. With the implementation of the Johansen test for cointegration, a VECM is estimated and the findings indicate that WTI prices and heat rate variables do not respond to deviations in the two cointegrating relationships. Thus, it is implied that they are weekly exogenous. To have a better understanding of the linkage of the prices and to examine the short-run deviations, the Engle-Granger method is applied to the two specifications estimated by

Ordinary Least Squares (OLS). One interesting finding is that an improvement in the heat rate of natural gas relative to oil fuel has increased the price of natural gas relative to the residual fuel oil. Finally the ECM model is estimated including stationary variables such as weather, inventories, disruption in production etc. Finding support the argument that the relationship between natural gas and crude oil is indirect, through competition of residual fuel with natural gas. Changes in crude oil prices effect natural gas and residual fuel prices, indicating that the former is weakly exogenous in a system including the latter. This paper finds that inventories, weather, production shut-down can have an effect on the short-run dynamics of the prices and thus should not be neglected. Finally the novelty of this study, to include the technological change permits to explain the drift in the long-run linkage between natural gas and residual fuel oil prices.

Hartley & Medlock (2014) did another study about natural gas, crude oil and technology but this time they also examine the role of the exchange rate in this relationship. They construct a simple model were oil is traded between the home country and a foreign one but natural gas doesn't. The authors point out that because of the limited switching between natural gas and crude oil in industries the exchange rate can have an important role in determining the price of the two commodities. Monthly data of Henry Hub and Brent crude prices are used from January 1995 to December 2011. In the cointegration relationship the heat rate variable is included to account for the technological change and also the foreign exchange value of U.S. dollar (Broad Trade Weighted Exchange Rate Index). The first step of the empirical analysis is to test if the series are integrated of order one. The ADF test, the Phillips-Perron (PP) and KPSS test indicate that all the variables are I(1). The Johansen test is applied and it suggests that there is one cointegration vector at the 0.05 significance level. The heat rate and exchange rate variable is taken into consideration to find out that the coefficient of the heat rate is negative. This is an indication that the better the relative thermal efficiency of the natural gas compared to that of oil the higher the relative price of natural gas to the oil price. Those results are indeed consistent with their previous study. As far as the exchange rate is concerned, they prove that the changes in the actual exchange rate can describe some of the long-term effects in the natural gas-crude relative prices. Finally some exogenous variables are included such as weather-related events, working inventories and shut-in production that can explain the short-run deviations. An interesting finding is that natural gas prices can respond quickly to inventory and weather shocks but much slowly to deviations from the long-run relationship.

Economic theory suggests that global economic activity must be accounted for the modeling of crude oil demand, thus it is a basic determinant of crude oil prices. He *et al.* (2010) examine the possibility of a cointegration relationship between real future prices of crude oil, the Killian economic index and a weighted US dollar exchange rate. The authors begin with the construction of a demand and supply structural model where demand for crude oil is an equation including its own price, global economic activity and exchange rate variable and the supply equation consists of crude's price and a constant variable. The data used in the VECM are monthly quotations of WTI crude oil futures deflated by US CPI, global oil production, petroleum inventories, weighted US dollar index and the Killian economic index from 1998 to 2007. The empirical analysis starts with implementing the ADF test to ensure that all of the variables are non-stationary with the exception of the production variable that was found to be stationary. Afterwards the Johansen approach is applied to find out that crude real price, the Killian economic index and weighted exchange rate are cointegrated. Additionally, tests are performed so as to identify the direction of causality between the variables. The results suggest that the Killian global economic index Granger causes the crude oil prices in the long-run as the former was found to be weakly exogenous. As a consequence oil prices are affected by deviations of the Killian economic index through long-run disequilibrium and short-run effects. Finally they constructed a general ECM model that is extracted from an Autoregressive Distributed Lag Model (ADL) to show that a permanent change in the global economic activity takes longer time for the crude price to adjust than with a permanent change in the US dollar index.

The majority of the past literature studies the relationship of crude oil and natural gas prices with only Henry Hub or NBP prices taken into consideration. The novelty of the Ji *et al.* (2014) is the examination of the effects of crude prices and economic activity on different natural gas import prices as they are formed in separate regions with unique features. The data source consist of monthly data of Henry Hub spot prices, US pipeline-European pipeline monthly average import prices, US-European-Japanese LNG monthly average import price and the Kilian index to account for the global economic activity. The existence of cointegration is tested with the Johansen-Fisher, Pedroni (2004) and Kao (1999) tests converging to cointegration. The cointegration tests support the argument of a long-run co-movement between global economic activity, international crude oil prices and natural gas import prices. The long-run equation is estimated using the Fully Modified Ordinary Least Squares (FMOLS) and the estimated coefficients prove that global economic activity and crude oil prices have a positive effect on natural gas import prices but the degree of the impact differs between the regions. In North America international crude oil prices have a weak impact on natural gas import prices while the impact is greater in Japan and Europe. In contrary the effect of global economic activity is greater in North America and less in Japan and Europe where natural gas prices are still based on the oil-index. Finally based on a VAR model, the impulse response functions and the variance decomposition analysis are generated to abstract more information regarding the impact of crude oil volatility to natural gas prices. International crude oil prices volatility has a negative effect on natural gas import prices which is weak in North America and Europe and in Japan has some implication in the production activities. Another interesting observation is the asymmetric response of natural gas prices in deviations of the international crude oil prices. Again the responses differ between the regions where in North America an increase or decrease of crude oil prices has a minor impact on natural gas prices and in Japan and Europe the impact is more significant.

Whether an energy commodity's price is determined regional or international can affect the existing relationship with other energy commodities prices. Siliverstovs *et al.* (2005) aims to examine the relationship of the European, North American and Japan regional natural gas markets by using monthly import prices (pipeline gas and LNG) from those three regions in the time period of 1990 to 2004. In order to take into account the market condition in North America, monthly average spot prices of Henry Hub are included as well as Brent crude monthly average spot prices to investigate the international interaction of the two prices. The authors examine the integration of these markets mainly due to three facts: firstly natural gas imports are limited geographically within those tree regions; secondly the North American natural gas market is now liberalized while in Europe is under transition and thirdly LNG trade has increased during the recent past years. The degree of integration is tested with ADF, PP and KPSS tests. Results show that all of the series are non-stationary. To test if the markets are integrated the Principal Components Analysis (PCA) is applied. Results provide evidence that two principal components attribute to almost all the variation. Two sets of variables reflect the same characteristics: US pipeline with Henry Hub as one group and the other LNG Europe, Japan European Pipeline and Brent as the other. This is an indication that there is a co-movement in the regional natural gas prices but not in the international one. To test the cointegration of the prices more formally the Johansen test is applied to the VAR model transformed to an ECM. The results from the cointegration bivariate tests provide evidence that there is strong integration in the regional natural gas prices in North America and Continental Europe. Also price co-movement is observed between European and Japanese natural gas prices. However the absence of global market integration is underlined. This shows that he natural markets across the Atlantic are divided and this separation took place 1990s. A final observation is the North American natural gas and Brent crude prices are not cointegrated and Siliverstovs *et al.* (2005) support the argument that the North American natural gas prices are decoupled from the crude oil ones due to the liberalization of the regional natural gas market.

On the same direction Ji & Fan (2015) reinvestigate the argument and examine the diversion of the crude oil prices. Through their empirical analysis try to find which crude oil variety is a price setter across the sample data. The crude oil varieties selected for this analysis are WTI, Brent, Dubai, Bonny and Tapis and the data-source consists of daily data from January 2000 to March 2014. From the data statistics of this crudes WTI seems to be the steadiest price where the others have higher standard deviation. Also correlation between the crudes is high except the cases with the WTI in which correlation is lower. However the abnormal return correlations are much smaller between the five crudes which indicates s that the may be regional factors effecting the prices. The analysis begins with a time varying correlation combined with the construction of an average distance of the crude oil regional prices with a window width of 250 days. Results from this method point out that prices diverge more in the short term. Another interesting observation is that WTI seems not to be integrated always with other crude prices probably because of the different market conditions in the US. Bai & Perron (1998) tests for structural breaks where posed and double maximum UDmax or WDmax, so as to identify breaks. A break is proven to occur at October of 2011. The average distances shown in this paper support the argument that WTI has been separated from the other crude prices in September 2010. The next step to this papers analysis is an implementation of ECM based on Directed Acyclic Graph (DAG) with the sample divided to two subintervals: January 2000 to 20 September 2010 and 21 September to March 2014. The Johansen test applied shows that there are 4 cointegrating equations in the specification for the first sample period. Thus there is a cointegrating linkage of the 5 crude oil markets. Finally the DAGs in this paper demonstrate the changes in the relationships between the crude oils. WTI, Brent and Dubai are still the main benchmarks; Tapis remains the role of price taker. However in the first sub-period WTI and Dubai have a clear effect on Brent prices where in the second sub-period Brent crude is the one that affects the other 2. To conclude it is obvious that WTI crude separates from the rest prices but it is not clear if this is a permanent phenomenon or not. For the rest of the world's crude oil market the strong integration holds and Brent crude oil has become the price setter that affects the other ones.

The relationship between energy commodities is another subject that is studied as they are used as inputs in the power industry. The relationship between natural gas prices, crude oil and coal is investigated by Mohammadi (2009). He attempts to find the causal relation of the prices and if those are symmetric or asymmetric. The empirical analysis is divided into two parts, the one using annual data from 1970 to 2007 and the other monthly from 1976 to 2008 both in real prices of North America. In the annual data model the Johansen test (1988) fails to reject the no-cointegration hypothesis in the tri-variable and bi-variable model, provided that, an unconstrained VAR is constructed to examine the short-run dynamics of the relationship. Findings show that changes in oil prices have an effect on gas prices but not the other way around. Impulse response functions by Pesaran & Shin (1998) and variance decompositions are generated to show the price movements of the three commodities. A shock in the log of crude oil affects positively all three energy prices in two periods and fades out after that. Also variance decomposition findings show that natural gas and coal price innovations don't explain variations in crude oil prices. However 40% of the natural gas price variation is affected by oil price innovations and coal prices are not affected by any changes in the two other markets. On the other hand by using monthly data the Johansen test rejects the hypothesis of no-cointegration and with a VECM model there is evidence of a long-run causality from oil to natural gas prices. In this model impulse response functions support the argument that innovations in crude oil prices have a persistent effect on natural gas prices and its own ones and variance decompositions are similar with the annual model findings. Finally, the MTAR specification provides evidence of asymmetric adjustments in the long-run relationship.

The relationship of the three commodities prices is also examined by Bachmeier & Griffin (2006). However in this study the market integration of crude oils and regional coals separately is also examined. Daily price quotations are used since 1989 for five different crude variations; WTI, Brent, Arun and Alaskan crude to test for market integration. By constructing an ECM model proposed by Engle & Granger (1987) their findings support the argument of highly integrated world oil market. In order to extract more information about the degree of market integration Bachmeier & Griffin (2006) calculate the percentage of instantaneous long run increase denote as "instant %". Adding to that the "half-life" is capturing any disequilibrium that the instant % has not taken into account. As they state in the work, two markets are likely strongly integrated and establish an economic market if they have high instant % with a short half-life. In their ECM for four pairs of crude oils the WTI-Brent reveals the higher market integration with 96% of instant % and 6.8 days half-life. That indicates that almost 96% of the adjustment is happening instantaneous while the remaining adjustment happens after 6.8 days. All four pairs exhibit sings of high market integration even with the monopolistic Arun crude oil. In North American coal market their examination is between the prices of Western and Eastern coals. Empirical findings show a cointegration relationship between and among those regional coals. However, cointegration appears to be stronger among the western coals and weaker between the western and eastern ones. Furthermore low instant % and huge half-life show low market integration. As far as the linkages among the primary energy fuels is concerned there is a cointegration relationship among them but a weak one. Bachmeier & Griffin (2006) expected that there will be a strong relationship between natural gas and coal but empirical findings don't support this expectation. Natural gas and oil have a long-run equilibrium relationship and show signs of an integrated market but again a weak one compared to the crude oils and western coals market integration.

Many studies have incorporated in their investigation the electricity market as the energy commodities are essential inputs for the power industries. In that direction Bancivenga *et al.* (2010) study the relationship of crude oil, natural gas and electricity in the UK. While the liberalization of the natural gas and electricity markets has been implemented in the past years, the dependence on oil prices is still valid. Daily prices of Brent crude, National Balancing Point (NBP) for natural gas and European Energy Exchange for electricity are used from September 2001 to December 2009. Firstly by estimating a rolling correlation the researches attempted to capture the short-run relationship of the variables. However the results do not provide the necessary information for the nature of the relationship between the prices of the three commodities. The cointegration analysis between the prices is performed using 2 methods: the Johansen and the Engle-Granger method. In the Johansen procedure a VECM is estimated based on the reduced rank regression model. Their findings support the existence of two cointegrating vectors and one common trend which is implied to be crude oil. The Engle-Granger method is applied in order to examine the cointegration relationships individually using OLS. The results support the findings of the Johansen method and each pair has a long-run relationship. Finally, the existence of cointegration among the variables allows the researchers to use the ECM to capture the short-run dynamics. The short-run influences among the two variables may be contributed to other factors that can affect natural gas prices independently.

Mjelde & Bessler (2009) focus on the relationship of North America electricity prices with the major fuel sources used in the market. They use weekly spot peak and non-peak wholesale electricity prices from 2 different regional markets: Pennsylvania, New Jersey, Maryland Interconnection (PJM) and Mid-Columbia (Mid-C) along with weekly prices of Henry Hub natural gas, Penn railcar coal, WTI crude oil and uranium from June 2001 to April 2008. A VECM is constructed including the eight prices, heating degree and cooling degree day variables. Since all prices were found non-stationary using the ADF test the Johansen's trace test statistic is applied to find the number of cointegrating vectors. Results show the existence of 4 cointegrating vectors with a constant within the cointegrating space. The presence of less than n-1 cointegrating vectors in the system implies that there is no market integration. Exogeneity tests are performed to reveal that coal, natural gas and crude oil prices are weakly exogenous in the error correction model. Mjelde & Bessler (2009) apply a one-time-only shock to each energy prices so as to examine their responses by applying the two-step GES algorithm (Chickering 2003). Electricity prices respond less to oil shocks in comparison to other fuels. The peak electricity prices of the two markets respond similarly to shocks in natural gas prices as natural gas is commonly used for peak power generation. The most extreme responses are reported from shocks in the coal market. This can be attributing to the 50% share of coal in electricity production in the two markets. Another interesting observation for the coal market is that is not affected by shocks from other fuel sources as it differs in transportation cost and the long-term contracts pricing. Finally it is noted that in contemporaneous time peak electricity prices can affect natural gas prices and in turn crude oil prices. However in the long run fuel prices can influence peak and off-peak electricity prices.

A similar study of Serletis & Herbert (1999), examine the relationship between crude oil, natural gas and power prices in North America. There seems to be a connection between those prices because both are used as inputs in power generation for the peaking hours and also are substitutes for the industrial sector. Daily data of futures are used from October 1996 to November 1997 for Henry Hub and Transco Zone 6 natural gas prices, fuel oil for New York delivery and Pennsylvania, New Jersey, Maryland (PJM) power market for electricity future prices. The ADF unit root test is applied to the data series to find out that natural gas and fuel oil prices are non-stationary where the power prices are stationary. For the non-stationary data, which is a prerequisite for the cointegration analysis, the Engle-Granger test is implemented where the null hypothesis of no cointegration is rejected at the 0.01 significance level. This can be interpreted that there is a long-run equilibrium relationship between the Henry Hub, Transco Zone 6 and fuel oil prices which reveals the existence of effective arbitrage mechanisms for the prices of this markets. In order to capture the short-run dynamics of the relationship a bivariate VAR in first differences augmented with an error-correction term is estimated and Granger-causality tests are performed. It is shown that Henry Hub prices Granger-cause Transco Zone 6 prices and vice versa. To conclude this paper shows that there is a long-run stable relationship between the two prices of natural gas and the fuel oil but not with power electricity prices. Another finding of the paper is the existence of causality between the price pairs. The findings have important implications for policy makers and for energy or financial companies to design and pursue a successful trading strategy; however the authors acknowledge that the time period of data they used is short and results must be treated with caution.

In the previous years there was an argument that liberalization of the Natural Gas and Electricity markets can change the trends of the prices and thus alter the existing relationship. Asche *et al.* (2006) examine the link of the prices and market integration in the period where the UK electricity market was liberalized and the natural gas market was isolated and deregulated just before the Interconnector pipeline started to operate. The data they use are monthly wholesale prices for the period of January of 1995 to July of 1998 and for comparison reasons July 1998 to December 2002 period is also examined. The ADF test was implemented and all prices were non-stationary I(1). The Johansen procedure is applied to allow the testing for cointegrantion and exogeneity by constructing a VECM. Due to the short range of the data used, tests for autocorrelation, ARCH and heteroscedasticity were reported and no evidence of misspecification was found in both bivariate and multivariate model. The two models support the argument that there are cointegrating vectors in the series and prices are moving with a stochastic trend. High market integration is observed in both models as the test for proportionality states (the Law of one Price). The energy commodity that affects the other two and serves as the price leader is oil. Tests for week exogeneity reports that in the period examined natural gas and electricity prices were led by oil prices, as the hypothesis for exogeneity for oil prices cannot be rejected. Finally to examine if there was a change in the relationship of the prices or the market integration the methodology used is applied the second period set. The cointegration argument still persists but results show that there might not be market integration after the Interconnector linked UK gas with the rest of Europe.

Another study that investigates the possibility of energy prices decoupling in UK is that of Panagiotidis & Rutledge (2006). They examine whether there is a long run cointegration relationship between wholesale natural gas prices, established in several markets, and Brent crude oil prices in the sample period of 1996 to 2003 at the time when natural gas market wan under liberalization regime and in 1998 the Bacton-Zeebrugge gas Interconnector was operational. Papagiotidis & Rutledge (2006) use three unit root tests to the series: the ADF the Breitung & Taylor (2003) and one by Saikkonen & Lutkepohl (2002) which allows structural breaks in the series. All tests conclude that series are stationary at first difference. Two methods for cointegration are applied: the trace test of Johansen and a method proposed by Breitung (2002). Both tests confirm the argument of the long run equilibrium relationship of oil and gas prices and provide evidence that this relationship also existed before the opening of the interconnector. The short-run dynamics are examined using a VECM and 4 tests for linearity were applied to the residuals of the models estimated: McLeod-Li, Engle, Tsay, and BDS so as to exclude the possibility of deviation from linearity. Finally, Generalized Impulse Response Function (GIRF) is estimated to trace the shock of one variable to the other endogenous variables. The GIRF is estimated using the Pesaran & Shin (1996) method and the results point out that a positive shock in oil prices result in a negative response of the natural gas price which fades out quickly.

Asche *et al.* (2013) is one more study that implicates the liberalization of natural gas market with the independent price determination its price. Through the liberalization of the market many believed that the natural gas price would be determined by demand and supply dynamics and it would decouple for oil prices. This papers aims to examine the integration of the natural gas market in Europe between the NBP in UK, the German contract price spot price in Zeebrugge and the Tide Transfer Facility (TTF) but also the existing link with the Brent crude oil prices. The biggest share of the natural gas traded in Europe is covered by long-term supply contracts but there are evidence supporting the argument that they are determined by oil prices and this might strengthen the linkage between the two prices not decouple them. The monthly data form 1999 to 2010 are tested for stationarity with the ADF, GLS and KPSS test to find out that all the series are integrated of order one. Since the series are non-stationary the researchers continue with the Johansen cointegration on a VECM test between the prices. The findings indicate that there are three cointegration vectors in the model. Evidence show that the market is integrated but it not proven if oil price still determines natural gas ones. Therefore, exogeneity tests are applied to all the prices. Hypothesis of weak exogeneity is rejected for all three natural gas prices but not for oil. These results are in consistence with the previous literature and once more it was found that oil prices determine natural gas ones. As a consequence it is proven that natural gas markets in Europe don't have an independent price determination.

Except from liberalization of the markets another reason that debates the decoupling of the prices is supply factors. Asche *et al.* (2012) study the impact of shale gas production on the gas-oil relationship. There has been a lot of development in the production of shale gas the past years. In Europe oil and natural gas prices seem to decouple more than ever due to the gas-to-gas arbitrage. However in Europe most long-term contracts of natural gas are priced based on oil and oil products and that indicates the relationship between the two prices is stable. The decrease of natural gas price the past years, may be attributed to the oversupply and the decreased demand as a result of the economic crisis, as most of types of energy. The authors state, that the increased use of natural gas in the power generation can be attributed to the Combined Cycle Gas Turbine (CCGT) plant's higher efficiency and lower cost. Shale gas production has been increasing in the US but in Europe it's still in transition because of the environmental concerns although there are reserves in over 32 countries. Taking that into consideration, other aspects such as the reduction of oil in electricity production and the proportion of oil in transport, one could argue that the price decoupling of natural gas and oil is probable. The empirical analysis of this paper uses Brent and NBP monthly quotations from September 1996 to March 2010. The ADF and KPSS test applied in the series to prove that all series are non-stationary. Then the Johansen method is applied to test for cointegration and exogeneity tests to a VECM. Findings prove that there is a long-term equilibrium relationship between natural gas and oil prices and exogeneity test show that oil prices determine natural gas ones in Europe. Finally a test for the stability of the relationship is conducted by implementing a recursive Chow test that support the argument of a stable long-run relationship between the prices despite the supply shocks and the liberalization of natural gas market that have effected prices the previous years.

A different study from Brown & Yücel (2009) examines how the gas-to-gas Atlantic Arbitrage that effects the natural gas and crude oil relationship. They study the relationship between natural gas in the two continents and try to support the argument that the linkage in the natural gas prices is through the integrated crude oil market rather than the gas-to-gas LNG arbitrage. The weekly data from June 1997 to May 2008 for Henry Hub, NBP, WTI and Brent are tested for stationarity with the ADF test. Results show that all the series used are I(1). The Johansen test shows that all combinations of prices are cointegrated. In order to capture the deviations from the long-run and which changes of the dependent variable are due to the independent or dependent variables an ECM is specified along with tan error-correction term. Findings reveal that there is a bi-directional causality between Henry-Hub and NBP prices that may be attributed to the LNG gas-to-gas arbitrage. Also oil prices are found weakly exogenous in both prices of natural gas but WTI was found significant in the Henry Hub equation. Causality was found between WTI-Henry Hub, Brent-NBP, WTI-NBP with the oil prices to effect natural gas ones not the other way around. An interesting finding is by estimating a multivariate model the two natural gas prices have no independent effect when WTI or Brent prices are included. That supports the argument that crude oil prices have a leading role in coordinating the two prices from two continents. Finally some exogenous variables are added to the multivariate model: heating degree days, cooling degree days and their deviations from the norm, shut-in production and inventories of natural gas. All multivariate models suggest that crude oil prices play the role of coordinator between the prices of natural gas across the Atlantic and gas-to-gas arbitrage may not be so important, even so strengthening the relationship between crude oil and natural gas because of the oil-index pricing of the European LNG shipment contracts.

The study of Erdos (2012) is another study that challenges the long-run relationship of oil and gas prices. Erdos (2012) is using a VECM and provides evidence that at the end of 2008 US gas prices decouple from European gas prices and oil prices, while UK gas to oil relationship is cointegrated in the long-run. Weekly spot WIT, Henry Hub and NBP are used from 1994-2011 to capture the dynamics of the relationship. Three tests for stationarity are applied, the ADF, PP and KPSS for the whole sample, for January 1994 to December 2008 and January 2009 to December 2011. The results reveal that WTI and NBP prices are non-stationary for all 3 sub-periods but the HH appears to be stationary for the last sub-period. Non-stationarity is a basic component of the cointegrating relationship. The Johansen test is performed for cointegration and for the full sample there is not such a relationship between WTI and HH probable because of the price trends of the last three years of the sample. Before 2008 there was a long-run cointegration relationship between the two but that has altered. A cointegration relationship stands between WTI and NBP, so as for the NBP with HH, however these relationships seem to weaken after 2009. A VECM model is constructed so as to find the causal links between the two commodities prices and to allow exogenous variables effecting natural gas prices such as cooling and heating degrees and their deviation from the normal degrees, the Gulf of Mexico production shut-ins and natural gas inventories deviations from the 5 year average. Different models are estimated using different sub-periods and changing some of the variables in order to extract more substantial information. The findings of these models support the argument that US natural gas prices have decoupled from the US crude and European natural gas after 2009. Until 2009 gas prices in US where higher than in Europe and that boosted LNG exports to US. These exports however boosted the supply in the US and price adjustments took place. The increasing production of shale gas could lead the US to export some of its gas but because of the lack of LNG and export capacities this didn't happen and thus gas prices decoupled.

This recent paper of Lin & Li (2015) examines the spillover effects between oil and natural gas markets based on a VEC-MGARCH (Multivariate Generalized Autoregressive Conditional Heteroscedasticity) framework that includes both price and volatility spillover. Previous literature has examined the relationship of the two commodities and the spillover affects across the markets but only from price or the volatility perspective, never combined. The paper uses monthly data from 1992 to 2012 from Europe, Japan and US. Unit root tests are applied to the data to determine the order of integration of the series. The ADF, the GLS de-trended test and the Lee-Strazichiz test that allows structural breaks indicate that the prices of natural gas and oil are integrated of order one. Using the Johansen approach evidence show that Brent crude oil prices are cointegrated with Europe's and Japans natural gas prices where in the US evidence show that the prices are decoupled. The authors of this paper state that the absence of cointegration among the US oil and gas prices could be because of the pricing mechanism in US and because the "soaring production of shale gas has not been mitigated by natural gas export for the limitation of liquefying and export capacity". By implementing a VEC specification it was found that there is a price spillover from the crude oil markets to the natural gas ones even at US where prices have been decoupled. The test for the volatility spillover was examined with the MGARCH-BEKK specification and it was shown that volatility in the oil market spillovers to the natural gas market in both Europe and US and vice versa but that is not the case for Japan mainly because of the price system in Japan is oil-indexed. Finally, as Erdos (2012) stated the only question that needs answering is if the separation of prices is a permanent situation or in the near future prices will again co-move.

Ramberg & Parsons (2012) want to clarify through their study if prices have temporally shifted away or it's a permanent situation and whether cointegration exists. There are two facts that the previous literature didn't take into account or explained: the existence of a large portion of volatility in the price of natural gas and second that the cointegration relationship wasn't always stable: in early 2006 and the start of 2009. The energy industry has used through the years different rules of thumb such as a simple 10-to-1 ratio or energy content ratio and more specified ones like: based on natural gas-residual oil or natural gas-distillate oil competition. However these methods are far from explaining the relationship of the two commodities prices. For the purposes of this paper weekly day-ahead price quotations are used from January of 1997 to December 2010 for WTI and Henry Hub prices. A VECM is constructed to capture the changes of the lag price changes and a VAR model is fitted so as to include 6 exogenous variables that effect natural gas prices: cooling degree days heating degree days, deviations of both variables for the normal, shut-in production in Gulf of Mexico due to natural phenomenon and differences in the inventories of average natural gas storage. The Johansen test is applied for cointegration and it reveals a cointegration vector at significance of 0.05 level and no vector at significance of 0.01 level. So the authors get mixed signals over the relationship. After the estimation of the VECM, a conditional Error Correction Model is estimated so as to treat oil price as weakly exogenous. The conditional ECM captures the reversion of the natural gas prices to its equilibrium relationship and is estimated that the half-life for the reversion of the natural gas price is about eight weeks to its cointegration equilibrium relationship. The coefficients and the signs of each exogenous variable are consistent with the economic theory and past literature results. Finally the researchers consider the possibility of cointegration but with structural breakpoints: one at March 2006 and one February 2009. For the first period there is clear evidence that the long-run equation has shifted. For the second period the information is inconclusive. To sum up there is evidence that the cointegration relationship exists between the two prices but we also need to take into account the two major points that are outlined in this paper: that there is a 85% of volatility that is unaccounted for and of course that creates problems at predicting where the price of natural gas is going and that there shifts in the long-run relationship that created a decoupling in the prices but that didn't last long and the prices again reached their equilibrium.

Brigida (2014) also dispels the notion of the decoupling prices. He examines the cointegration relationship between natural gas and crude oil prices; however she endogenously includes shifts in the cointegration vector in the Markov-switching cointegration equation. The structural breaks that exist in the relationship between the two relative prices are modeled, by implementing them in the standard ECM. The data that this paper uses are weekly and monthly logged prices of natural gas and crude oil futures from July 1997 to September 2012. In order to choose the appropriate state that suits better the cointegration equation Brigida (2014) allows the states to range from one to three. Comparing the two-state model with the other two, evidence show that the two-state model has stable states with an obvious regime switching and it is the only one that rejects the null hypothesis of a unit root for the residuals in both the ADF and PP tests. This paper uses the state-weighted residuals to be able to estimate the ECM for the whole sample period including exogenous variables that effect prices such as cooling-heating degree day, extreme weather conditions, natural gas storage differential and rig count. Findings show that, ceteris paribus, an increase in crude oil prices will have an increasing effect on natural gas prices in subsequent periods and all exogenous variables are significant and with the appropriate sign. The final analysis indicates that the relationship of natural gas and crude oil prices is strengthened once there is control of the regime switching. Arguments that prices were decoupled in the beginning of the 20th century are dispelled and a shift is observed in August of 2000. The two-state cointegration equation is the one providing more and accurate information about the long run linkage of the two prices.

Wolfe & Rosenman (2014) examine the effect that inventory announcements of natural gas and oil have, on price volatility. The authors of the paper want to examine the bidirectional causal relationship of the two commodities. The dada set consist of intraday high-frequency future prices data of 10min interval, Wolfe & Roseman (2014) measure volatility using different regression equations for oil and natural gas. An important variable of the regression is the unexpected changes in the inventories. The "surprise" variable is constructed so that negative surprises reflect inventory shortages and positive ones to reflect inventory gluts. In order to exclude some non-related events that effect volatility some dummy and control variables are introduced in the regression such as dummies for the beginning-of-day, end-of-day and the first-trading-day, controls for three-month Treasury bill rate, Chicago Board Options Exchange Volatility Index, trader composition and S&P 500 return and absolute return. OLS method is used to estimate the regression. By graphical examination a spike in oil price volatility is observed, the day of gas announcements, an observation that has not been showed before. The empirical results using seemingly unrelated regression, shows that the effect of gas shortages and gluts is twice as strong to the effect of oil to gas price volatility. Furthermore it is shown that the bidirectional causality that was found holds across the maturity of the future contracts. Wolfe & Rosenman (2014) to provide robustness for their findings they contacted more checks including more dummy variables like LNG tech or discovery of shale gas, including structural breaks or even an estimation of the regressions using GARCH Models. Neither of those checks have altered the findings: there is a two-way causality in which inventory announcements affect future contracts prices. As the trading of futures contract is rising it is of importance to understand the linkage that exists so as to have better information about the price volatility spillovers and the sources of risk.

The importance of the pass-through of crude prices to natural gas and gasoline prices in an asymmetric and nonlinear manner is analyzed by Atil *et al.* (2014). They implement the cointegrating Nonlinear Autoregressive Distributed Lags (NARDL) model of Shin (2011) that allows asymmetries in the ECM both in long and short-run. Monthly spot closings prices of WTI, Henry-Hub and gasoline are used from 1997 to 2012. After estimating the NARDL models with short-run and long-run asymmetries, the Wald test is applied in order to study the short or long-run existence of symmetry. The presence of long-run symmetry is rejected for the oil-natural gas price case but cannot be rejected for the oil-gasoline price case. On the other hand short-run symmetry exists in the oil-natural gas price relationship and does not in the oil-gasoline price relationship. Taking into account the results from the Wald test and the AIC, SIC information criteria the model that best suits the oil-gasoline case is the NARDL (3,2) with long-run symmetry and short-run asymmetry. While for the oil-natural gas case the NARDL (2, 3) with short-run symmetry and long-run asymmetry is the most suitable one. So the empirical results show that gasoline reacts asymmetrically in the short-run where natural gas in the long-run. These results are consistent with the relevant theory because crude oil is the main input for gasoline production. Also crude oil prices are determined globally and can have an effect on natural gas prices over the long-run. So the adjustment of gasoline prices to crude oil price shocks is quicker than the natural gas adjustment. Finally another empirical finding shows that the negative oil shocks have an effect on both prices that is larger than positive shock. This larger asymmetric effect of the negative oil shocks can be linked to the bearish expectations that effect prices when economy is in recession.

3 Methodological Framework

3.1 Unit root and Stationarity Tests

The first step towards the empirical analysis is to determine the order of integration of our series. This step is crucial because it determines the framework to be used. A stationary series can be defined as a series with constant mean and constant variance, so if a series doesn't have those characteristics can be called non-stationary. To convert a non-stationary series to a stationary is by taking differences. The series then can be called integrated and are denoted as I(d) where d is the order of integration. The order of integration reveals the unit roots that a series may have or in other words the number of differences that we have to take in order to achieve stationarity. The stationary series are denoted as I(0) where the non-stationary that contain one unit root I(1). Most of the energy related financial time series as well as the most economic ones are known to be I(1). The non-stationarity is a prerequisite for our cointegration analysis. The stationarity of the variable is a very important characteristic that needs to be taken into account or else the researcher will conclude to spurious results.

In this empirical analysis we implement two unit root test: Augmented Dickey-Fuller test (ADF), the Phillips-Perron test (PP) and one stationarity test: Kwiatkowski, Phillips, Schmidt and Shit (KPSSS) test. With these three tests we will be able to establish the order of integration of our variables. In the following section we present the procedure that these tests follow as well as the mathematical equations that describe their function.

3.1.1 The Augmented Dickey-Fuller Test (ADF)

The Simple Dickey-Fuller Test was the first and has been the base-work for unit root testing, developed by Dickey and Fuller (1997). The simple DF test is implemented in an autoregressive process AR(1) while having a constant and a trend:

$$y_t = \rho y_{t-1} + x_t'\delta + \epsilon_t \tag{1}$$

The test can be implemented if we subtract y_{t-1} from both sides of the equation (1):

$$\varDelta y_t = \alpha y_{t-1} + x_t' + \epsilon_t \tag{2}$$

The two hypotheses that the results are tested are:

$$H_0\colon \alpha = 0\ (\rho = 1)$$
$$H_1\colon \alpha < 0\ (\rho < 1) \tag{3}$$

The null hypothesis implies that there is a unit root while the alternative that there isn't or that $\rho < 1$. The conventional *t*-ratio (3) is used against the critical values for the DF test.

$$t_a = \hat{a}/se(\hat{a}) \tag{4}$$

Where $\hat{a}$ is the estimated of a and $se(\hat{a})$ is the standard error of the coefficient. However if a series is correlated at higher order lags then we have to implement the ADF test. The ADF test can be converted to this equation:

$$\varDelta y_t = \alpha y_{t-1} + x_t'\delta + \beta_1 \varDelta y_{t-1} + \beta_2 \varDelta y_{t-1} + \cdots + \beta_p \varDelta y_{t-p} + u_t \tag{5}$$

With this specification we can test the hypothesis (3) using the *t*-ratio (4).

3.1.2 The Dickey-Fuller GLS Test (DF-GLS)

The Dickey-Fuller test with GLS detrending or GLS demeaning is a modification of the Dickey-Fuller test statistic that was proposed by Elliot *et al.* (1996). The DF-GLS test is mainly an augmented Dickey-Fuller test with the difference to be the time series transformation by a generalized least square method before executing the actual test. According to Elliot *et al.* (1996), this test can perform better that the previous forms of the ADF test.

3.1.3 The Phillips-Perron Test (PP)

Phillips and Perron (1998) developed a unit root test that is similar with the ADF one but with one important differentiation: the PP test includes an automatic correction to the DF method that allows residuals to be auto-correlated. This method estimates the non-augmented equation:

$$\varDelta_t = \alpha \mathrm{y}_{t-1} + x_t' + \epsilon_t \tag{6}$$

while modifying the *t*-ratio of the a a coefficient so that the serial correlation will not have an effect on the asymptotic distribution of the test statistic. The statistic that the PP is based on is:

$$t_a = t_a(\frac{\gamma_0}{f_0})^{\frac{1}{2}} - \frac{T(f_0 - \gamma_0)(se(\hat{a}))}{2f_0^{1/2}s} \tag{7}$$

3.1.4 The Kwiatkowski, Phillips, Schmidt and Shit Test (KPSS)

Another test that is commonly used to identify the order of integration of a time series is the Kwiatkowski, Phillips, Schmidt and Shit (KPSS) Test. One major difference between the previous tests (ADF, PP) is that the KPSSs null hypothesis is under the assumption that y_t is stationary where the alternative that is non-stationary. The estimation of the KPSS statistic is established on the residuals from the OLS regression of y_t:

$$y_t = x_t'\delta + u_t \tag{8}$$

where x_t' are the exogenous variables. The null hypothesis assumes that the variance of the error u_t is zero implying stationarity against the alternative that variance is greater than zero, meaning non-stationarity.

$$H_0: s_u^2 = 0$$
$$H_1: s_u^2 > 0 \tag{9}$$

The hypothesis is tested with the LaGrange Multiplier (LM) statistic:

$$LM = \sum_t S(t)^2 / (T^2 f_0) \tag{10}$$

where f_0 is the estimator of the residual spectrum at zero frequency and $S(t)$ is a function of cumulative residuals:

$$S(t) = \sum_{r=1}^{t} u_r \tag{11}$$

based on the $u_t = y_t - x_t'\delta(0)$ (12) residuals .

3.2 Cointegration

In order to establish if the is a long-run equilibrium relationship between our variables we need to implement cointegration tests. Engle and Granger (1987) were the first to introduce the concept of cointegration and state that if two or more non-stationary variables can form a linear combination that is stationary, then those variables are cointegrated. The the Johansen (1996) test will be applied in our empirical analysis for cointegration testing.

3.2.1 The Johansen Test for Cointegration.

In order to perform the Johansen method firstly, we must construct a VAR model of order p:

$$y_t = \mu + A_1 y_{t-1} + \cdots + A_p y_{t-p} + \varepsilon_t \tag{12}$$

where ε_t is an innovations vector and y_t is a n-vector of I(1) variables. The VAR model can be transformed to

$$\Delta y_t = \mu + \Pi y_{t-1} + \sum_{i=1}^{p-1} \Gamma_i \Delta y_{t-i} + \varepsilon_t \tag{13}$$

where Π is a matrix coefficient. In the case that the coefficient matrix Π has a reduced rank r<n then there are nxr matrices α and β with rank r that form stationary; $\Pi = \alpha\beta'$ and $\beta'\delta$. The rank of the coefficient matrix is the number of existing cointegrating relations and the columns in β are the cointergrationg vectors. Johansen, in his work provides two likelihood ratio tests: the trace test and the maximum eigenvalue test. The null hypothesis of the trace test is the r cointegrating vectors and the alternative of n cointegrationg vectors. On the other hand the eigenvalue test assumes the null hypothesis of r cointegration vectors against the hypothesis of r + 1 cointegrating vectors.

3.3 Vector Autoregression (VAR)

An econometric model that is widely used for forecasting models of interdependent time series is the Vector Autoregressive model. In the VAR approach each of the endogenous variables are considered as a function of lagged values of all the endogenous variables in the model. The VAR (p) model can be defined as:

$$y_t = A_1 y_{t-1} + \cdots + A_p y_{t-p} + \; Bx_t + u_t \quad (14)$$

where y_t is a vector of endogenous variables, x_t is a vector of exogenous variables, A and B are matrices containing coefficients, p is the number of lags and u_t is an innovations vector. Our system consists of I(1) and I(0) variables thus, we have to use the Structural Vector Autoregressive model (SVAR). The mathematical representation of the SVAR model is

$$Ay_t = \; A_1^* y_{t-1} + \cdots + A_p^* y_{t-p} + B\varepsilon_t \quad (15)$$

where the structural residuals ε_t are white noise and the A_i^* are structural coefficient matrices, for $i = 1, \dots, p$, that are different in general from the reduced form A coefficient matrices.

3.4 Granger Causality

After testing for cointegration an important procedure that we must follow is the Granger causality method. By this we will be able to find out the causal relationship between our variables. According to Granger (1969) if a X_1variable Granger-causes another X_2 variable then past values of the first variable contain information that helps predicting X_2 with better results than the information contained in the past values of X_2.The Granger causality method follows a linear autoregressive model of the two variables:

$$X_1(t) = \sum_{j=1}^{z} A_{11,j} X_1(t-j) + \sum_{j=1}^{z} A_{12,j} X_2(t-j) + E_1(t)$$
$$X_2(t) = \sum_{j=1}^{z} A_{21,j} X_1(t-j) + \sum_{j=1}^{z} A_{22,j} X_2(t-j) + E_2(t) \quad (16)$$

where, z is the total number of the lagged observations in our model, the A matrix includes the coefficients of the model and is E_1, E_2 are the residuals. The null hypothesis of no causality $A_{12} = 0$ can be tested using the *F*-test. The Granger method is a valid method to test causality however, has some limitations for several analyses. The framework that the test is based upon is by default linear. By this way we cannot take into account any non-linear relations in the model. Furthermore, the Granger causality test assumes that the causal relationship between the variables remains constant over time. So with the Granger method we cannot find out if the causal relationship is disrupted in some sub-periods or if it is indeed constant through our time frame. These are some reasons why will we also include in our empirical analysis causality testing in the frequency domain

3.5 Brock, Dechert and Scheinkman Test (BDS)

The BDS test developed by Brock *et al.* (1996) is used to identify if there is non-linear dependence in a time series. This test is applied to the estimated residuals in order to find out whether those residuals are independent and identically distributed (i.d.d.), as the null hypothesis assumes. If the BDS test rejects the null hypothesis then it is implied that in the structure of the time-series there is hidden a nonlinearity that cannot be neglected in the analysis.

3.6 Frequency Domain Causality Test

In the last part of our empirical analysis we apply the Frequency domain causality test of Breitung and Candelon (2006) (B&C hereafter). As stated in their work they adopted the earlier studies of Geweke (1982) and Hosoya (1991) to be able to develop their approach. With the B&C test we will be able to test for causal relation in the frequency domain and also to carry out several tests of causality over sub-frequencies. By this way we will have a clear overview of the causal relationship between our variables over the short-run and long-run.

Their study is based on a bivariate VAR model on a set of linear hypothesis on the autoregressive parameters but the procedure can be modified in order to allow cointegration and higher-dimensional systems. Firstly the $z_t = [x_t, y_t]$ is the two dimensional vector of the time series where z_t has a VAR representation of finite-order:

$$\Theta(L)z_t = \varepsilon_t \tag{17}$$

Where $\Theta(L)$ is a 2x2 matrix of polynomials with $L^k z_t = z_{t-k}$ and ε_t , the error vector, is white noise with $E(\varepsilon_t) = 0$ and $E(\varepsilon_t, \varepsilon_t') = \Sigma$ with Σ to be positive definite. If we assume that the system is stationary then the moving average representation of the system will be

$$z_t = \Phi(L)\varepsilon_t = \begin{bmatrix} \Phi_{11}(L) & \Phi_{12}(L) \\ \Phi_{21}(L) & \Phi_{22}(L) \end{bmatrix} \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix}$$

$$= \Psi(L)\eta_t = \begin{bmatrix} \Psi_{11}(L) & \Psi_{12}(L) \\ 21(L) & \Psi_{22}(L) \end{bmatrix} \begin{bmatrix} \eta_{1t} \\ \eta_{2t} \end{bmatrix} \tag{18}$$

with the implementation of the Cholesky factorization process we can obtain a structural representation of the system:

$$z_t = \Theta(L)C^{-1}u_t = \Psi(L)\mathrm{u}_t \, with \, \Theta(L)^{-1}C^{-1} = \Psi(L) \tag{19}$$

as the system can be modified in terms of the u_t structural innovations. Based on the process above, the spectral density z_t can be presented at frequency ω as

$$f_x(\omega) = (\frac{1}{2\pi})\left\{\left|\Psi_{11}(e^{-i\omega})\right|^2 + \left|\Psi_{12}(e^{-i\omega})\right|^2\right\} \tag{20}$$

So the hypothesis of non-causality in the framework of Geweke (1982) is checked from the Fourier transformation of the moving average coefficients:

$$M_{y\to x}(\omega) = log\left[\frac{2\pi f_s(\omega)}{|\Psi_{11}(e^{-i\omega})|^2}\right] = log\left[1 + \frac{\left|\Psi_{12}(e^{-i\omega})\right|^2}{|\Psi_{11}(e^{-i\omega})|^2}\right] \tag{21}$$

Within this transformation, if $\left|\Psi_{12}(e^{-i\omega})\right|^2$ is equal to zero then we can say that y does not cause x at ω frequency. As in our analysis, if the variables are I(1) and cointegrated then we can modify equation (20) to the form:

$$M_{y\to x}(\omega) = log\left[1 + \frac{\left|\widetilde{\Psi}_{12}(e^{-i\omega})\right|^2}{\left|\widetilde{\Psi}_{11}(e^{-i\omega})\right|^2}\right] \tag{22}$$

By this way we can work on a system that is cointegrated or with higher dimensional systems. As suggested by Hosoya (2001) the higher dimensional system can be re-transformed to a bivariate system by conditioning out the other variables of interest.

According the B&C procedure in order to test the null hypothesis $M_{y\to x}(\omega) = 0$ of no causality, some restrictions have to be imposed:

$$\sum_{k=1}^{p} \theta_{12,k} \, cos(k\omega) = 0$$

$$\sum_{\mathrm{k}=1}^{p} \theta_{12,k} \, sin(k\omega) = 0 \tag{23}$$

Those linear restrictions in order for y not to cause x at frequency ω, are the base of the Breitung and Cantelon (2006) method. So the final form of the VAR equation can be expressed as

$$x_t = a_1 x_{t-1} + \cdots + a_p x_{t-p} + \beta_1 y_{t-1} + \cdots + \beta_p y_{t-p} + \varepsilon_{1t} \tag{24}$$

for the x_t variable, where $a_j = \theta_{11,j}$ and $\beta_j = \theta_{12,j}$. As a consequence, the null hypothesis of no causality $M_{y\to x}(\omega) = 0$ can be written as the linear restriction:

$$H_0 \colon R(\omega)\beta = 0 \tag{25}$$

where

$$R(\omega) = \begin{bmatrix} cos\,\omega & \dots & cos\,p\omega \\ sin\,\omega & \dots & sin\,p\omega \end{bmatrix} \text{ and } \beta = \left[\beta_1, \dots, \beta_p\right] \tag{26}$$

The ordinary *F*-statistic for then null hypothesis is distributed as *F*($2, T - 2p$) for every frequency that takes values between (0, π). The significance of the obtained statistic will be compared with the 0.05 critical value of the χ^2 distribution with two degrees of freedom.

4 Data

4.1 Crude Oil and Natural Gas

In this section a brief analysis of our data are presented. The empirical analysis is based on daily spot prices of Western Texas Intermediate (WTI) for crude oil ($/bbl) and Henry Hub (HH) for natural gas ($/mmBtu) .Even though daily prices quotations for WTI were available from 1986, HH prices started to be documented in 1997. Take that into consideration WTI and HH daily prices were obtained from the U.S. Energy Information Administration (EIA) for the period from 7 January 1997 to 24 August 2015[1]. As stated in recent literature WTI is one of the basic crude oil benchmarks and its price is determined internationally. On the other hand HH prices are determined through regional demand and supply dynamics. The regional determination is the reason that many researchers dispute the long-run equilibrium relationship between them.

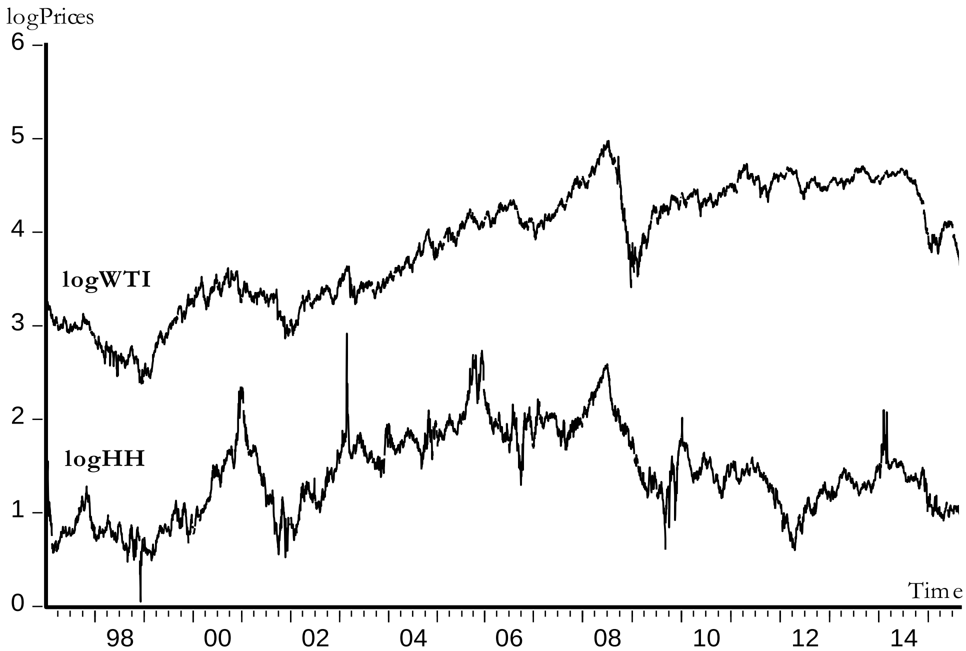


Figure 1: WTI and HH logarithmic prices.

In the above figure we display the log of WTI and log of HH prices. As a general observation we can say that the two prices seem not to move always together but that is under our empirical investigation.

---

[1] WTI and HH series are available at: http://www.eia.gov

4.2 Weather and Storage Shocks

To determine the relationship of the two commodities prices we had to select which other variables can affect natural gas price. The most common variables used in the relevant literature are cooling degree days (CDD), heating degree days (HDD) and working storage of natural gas. However in order to obtain the most useful connection between those variables and natural gas we must modify our data so as to subtract seasonality and get shocks that affect our variables. Daily data for cooling and heating degree days were obtained from the National Weather Service Climate Prediction Center for our time period[2]. The daily degree days (DD) are used so as to be able to construct the weather shocks (WS) by adopting a methodology introduced by Mu (2007). The equation to calculate the WS according to Mu (2007) is:

$$WS_t = \left(\frac{1}{k}\right) * \left\{\sum_{i=1}^{k} (DD_{t+i} - NDD_{t+i}\right\} \tag{27}$$

where DD_{t+i} are the actual degree days of the day t+i, NDD_{t+i} is the expected normally quotations of DD on that day t+i. We defined NDD_{i+1} as the sixteen-year average of the respective day values. As Mu (2007) states k can take values from one to seven as it stands for the weather forecast days. In this empirical analysis we use seven days as our forecast horizon. As seen in the Figure 2 below, in the Degree Days Index diagram seasonality is obvious. With the procedure that we followed we manage to remove seasonality and obtain Weather Shocks. Including this variable to our analysis we can take into consideration shocks from the normal weather that can affect natural gas prices.

[2] CDD and HDD series are available at: http://www.cpc.ncep.noaa.gov/

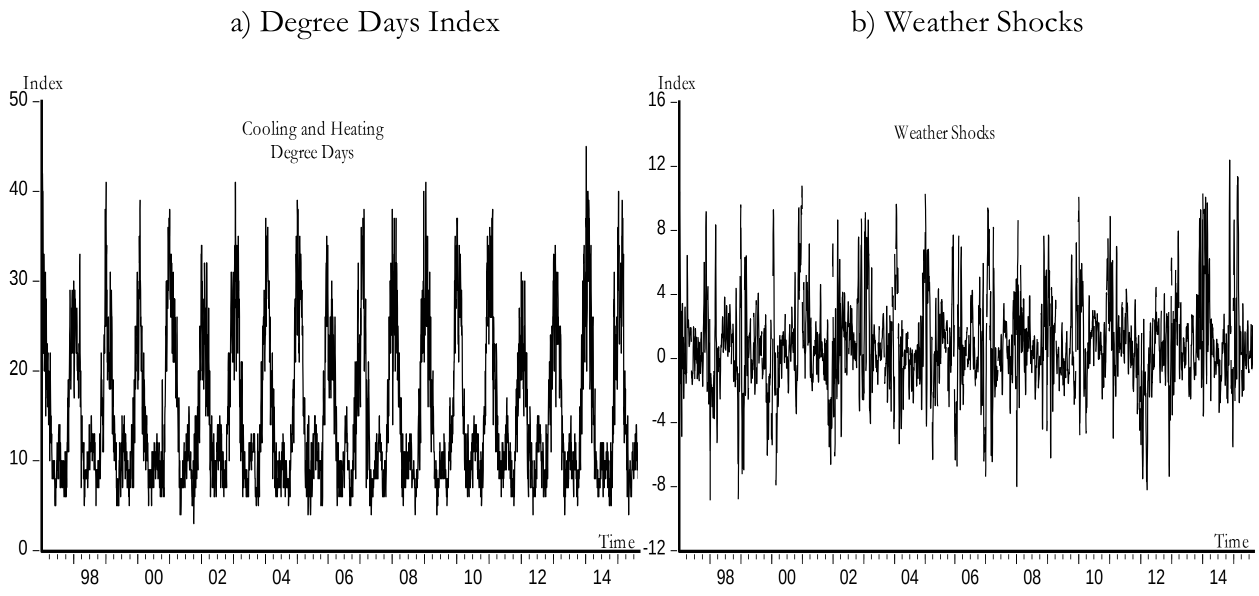


Figure 2: Degree days and weather shocks

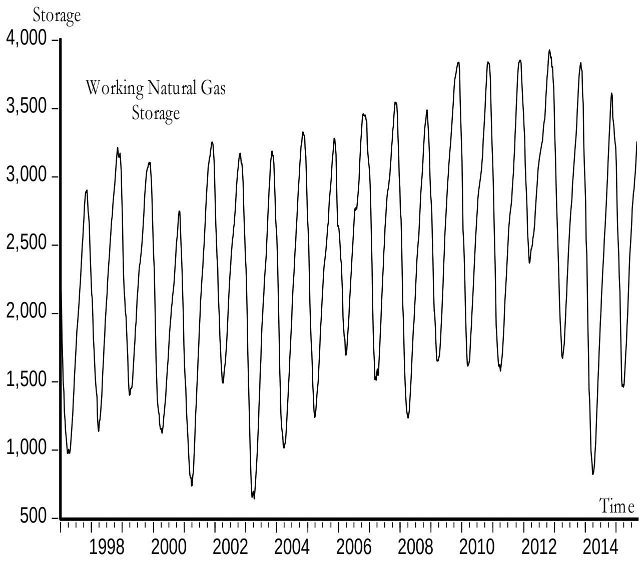
Weekly working natural gas storage data (Billion cubic feet) were obtained from EIA[3]. Based on another methodology of Mu (2004), we are able to construct another variable that reflects storage shocks. In order to find which week best describes the cycle of storage we must estimate equation 28:

$$y = c_1 + c_2 \sin \frac{2\pi}{\text{period}} t + c_3 \cos \frac{2\pi}{\text{period}} t + c_4 t^2 \tag{28}$$

where y is storage, t is the time trend and *period* is a scalar that describes the duration of the cycle. The cycle that best fits the above equation is the cycle with duration of 52 weeks. After that we estimate again equation 2 using the best fit and then to generate the residuals.[4] The difference between the raw data of working natural gas storage and storage shocks are displayed in Figure 3:

[3] Working natural gas storage data available at: http://www.eia.gov/dnav/ng/ng_stor_wkly_s1_w.htm
[4] Storage data were available at weekly frequency; we converted the weekly data to daily using the quadratic-match average method.

a) Working natural gas storage



b) Storage Shocks

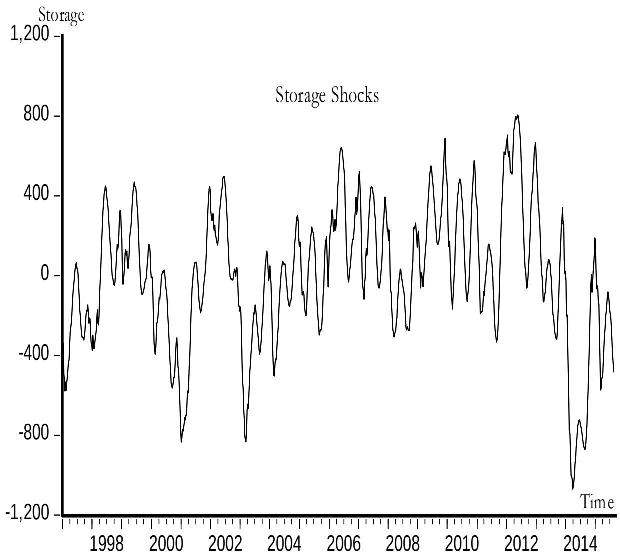


Figure 3: Working natural gas storage and storage shocks

As pointed out previously we transform some of our data. Weather shocks and storage shocks are major variables that can affect natural gas prices. The selection of those two is done because they have an important effect that can't be ignored. In the following section of empirical analysis WTI stands for crude oil prices, HH for natural gas prices, WS for the weather shocks and STOR for storage shocks.

5 Empirical Analysis and Results

5.1 Unit Root-Stationarity Tests

As the first part of our empirical analysis we must identify the order of integration of our variables. Currently we have four variables which are: WTI for Western Texas Intermediate crude oil prices, HH for Henry Hub natural gas prices, WSHOCK for weather shocks and STOR for storage shocks. We implement the ADF, DF-GLS and PP unit root test which are fundamental among the literature and finally, the KPSS stationarity test. In this section the results of the tests are presented in Table 1. The optimal lag length is based on the Schwartz criterion for the ADF, DF-GLS tests. For the PP test the Newey-West approach is used. The KPSS uses the Bartlett Kernel spectral estimation method and the selection of the bandwidth is according to the Newey-West procedure. Rejection of the null hypothesis is represented by *, ** and *** at 0.1, 0.05, 0.01 significance level respectively.

As the three unit root tests are applied to the logWTI we fail to reject the null hypothesis at the 0.01 significance level. The same result applies also for the KPSS test or in other words the H_0: stationarity is rejected. At first differences we reject the null hypothesis for all three tests for all levels of significance. This indicates strong evidence that crude oil prices are a non-stationary series. For the natural gas prices the DF-GLS rejects the null hypothesis for all levels of significance.

Table 1. Unit Root/Stationarity test for WTI and HH.

Variable/Test Method	logWTI(k)	dlogWTI(k)	logHH(k)	dlogHH(k)
	t-statistic			
ADF	-1.461110(0)	-69.70994(0)***	-2.760545(2)	-57.75392(1)***
DF-GLS	-0.927660(0)	-7.025910(15)***	-2.740989(2)	-56.54311(1)***
PP	-1.397573	-69.82338***	-2.913715	-65.97147***
	LM-statistic			
KPSS	7.270507***	0.182531	1.802788***	0.036456

Notes: The asterisks *, ** and *** denote the rejection of the null hypothesis at the 10%, 5% and 1% significance levels, respectively. For the ADF and DF-GLS tests k represents the selected lag-length based on the on the Schwarz criterion with k_{min}=0 and k_{max}=30.

On the other hand the ADF test indicates that there is a unit root at 0.1 significance level and the PP test at 0.05 significance level. Furthermore stationarity is rejected through the

KPSS test at 0.01 significance level. The same results as WTI are found for first differences. We can conclude that WTI and HH prices are integrated of order I(1).

Table 2 presents the results for the remaining two variables. For WSHOCK and STOR the unit root tests reject the null hypothesis of a unit root at 0.01 significance level. The KPSS test confirms that the variables are stationary as the null hypothesis cannot be rejected at 0.01 and 0.05 significance level for WSHOCK and STOR, respectively. After implementing the tests we can safely say that both variables are stationary I(0) as the majority of the test provide such evidence. Now that we have concluded about the order of integration for our series and we may argue that WTI and HH prices are non-stationary we can test if those two prices have a long-run equilibrium relationship. The remaining variables appear to be stationary, as expected, since they are considered as exogenous variables affecting natural gas prices.

Table 2. Unit Root-Stationarity tests for WSHOCK and STOR.

Variable/Test Method	WSHOCK(k)	STOR(k)
	t-statistic	
ADF	-14.02854(11)***	-5.568937(6)***
DF-GLS	-3.113580(14)***	-4.031481(6)***
PP	-10.04665***	-4.831100***
	LM-statistic	
KPSS	0.108722	0.475143

Notes: The asterisks *, ** and *** denote the rejection of the null hypothesis at the 10%, 5% and 1% significance levels, respectively. For the ADF and DF-GLS tests k represents the selected lag-length based on the on the Schwarz criterion with k_{min}=0 and k_{max}=30.

5.2 Cointegration

Now that we have established the order of integration for our variables we can proceed with cointegration tests. Using the standard Johansen approach we need to examine the presence of an equilibrium relationship between natural gas and crude oil prices. However our system is a compilation of I(1) and I(0) variables since weather and storage shocks can be treated as exogenous ones . As a consequence we implement the Rahbek and Mosconi (1999) specification so as to include in the cointegration vector the accumulated sum of weather and storage shocks along with a constant and restricted linear trend. Our model can be represented as:

$$\Gamma(L)\Delta y_t = \delta_0 + \alpha\mu_1 = \alpha\beta' y_{t-1} + \Phi(L)x_{0.t} + \varepsilon_\tau \quad (29)$$

Our model is estimated using three lags for the endogenous variables and one lag for the exogenous variables as the Schwarz information criterion indication. Cointegration test results are presented in Table 3 below:

Table 3. Cointegration Rank Test.

Null	Alternative	Trace Statistic	Critical Values	Eigenvalues
r =0	r =1	93.2441**	35.7764	0.0187
r ≤1	r =2	5.2372	18.1226	0.0011

Notes: The asterisks *, ** and *** denote the rejection of the null hypothesis at the 10%, 5% and 1% significance levels, respectively.

It is clear that the test provides evidence of the existence of one cointegration relation. Thus crude oil prices and natural gas prices have a long-run equilibrium relationship while conditioning on weather and storage shocks. The evidence we find is consistent with the majority of the past literature. There were indications that the two prices have been decoupled the last years but that hasn't taken place yet.

The estimated cointegration model is displayed in the Table 4 below for one cointegrating vectors.

Table 4. Estimated Cointegration Relations.

Coint-Relations	HH	WTI	Cum(WS)	Cum(STOR)	Trend
1	-4.5981	3.5845	-0.0045	0	0.033

5.3 Causality

The next step to our empirical analysis is to examine if there is a causal relationship between crude oil and natural gas prices. The presence of cointegration between WTI and HH implies causality but does not show the direction of it. We will conduct Granger causality tests to examine the direction of causality. In order to conduct this test we will implement the Wald-type test statistics.

Table 5. Granger Causality Tests 1.

H_0: WTI is Granger non-causal for HH H_1: WTI is Granger causal for HH		
Wald-statistic	DF	P-value
135.0113	3	0.0000

Notes: The Wald test statistic is asymptotical χ^2 with degree of freedom DF being the number of constrains.

The results of the first causality test are reported in Table 5. The null hypothesis of WTI no Granger cause HH is rejected as the *p*-value is zero. Thus the causality test provides evidence that past values of crude oil prices include information that can predict future prices of natural gas. The alternative direction of causality $(\text{HH} \longrightarrow \text{WTI})$ is presented in the Table 6 and in this case where the *p*-value is 0.3361, the null hypothesis of no-causality cannot be rejected.

Table 6. Granger Causality 2.

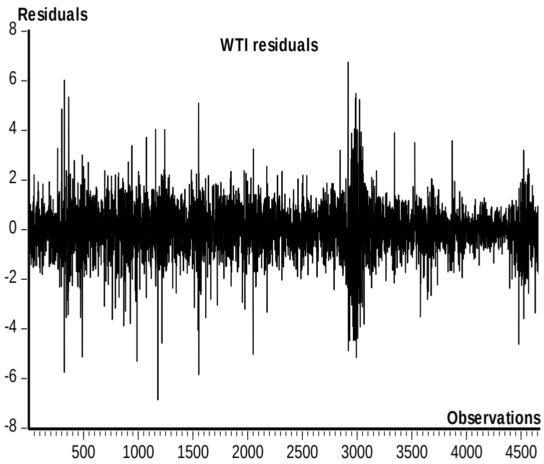
H_0: HH is Granger non-causal for WTI H_1: HH is Granger causal for WTI		
Wald-statistic	DF	P-value
3.3841	3	0.3361

Notes: The Wald test statistic is asymptotical χ^2 with degree of freedom DF being the number of constrains.

The two Granger causality tests support the argument that there is one-way causality between the commodities prices that of $\text{WTI} \longrightarrow \text{HH}$. However as we stated in an earlier section, Granger causality method has some limitations. It does not take into account the existence of nonlinearities in our data and assumes that causality is the same between short-run and long-run periods. To overcome those limitations we test for nonlinearities using the BDS test and then we implementing the B&C (2006) frequency domain causality test.

In order to perform the BDS test we extracted the normalized residuals for WTI and HH from our VAR model. Graphic representations of the residuals can be seen in the Figure 4 below:

a) Normalized WTI residuals



b) Normalized HH residuals

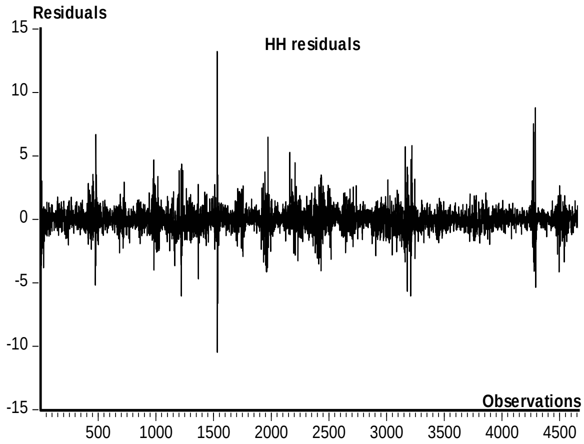


Figure 4. Normalized WTI and HH residuals

Now that we extracted our residuals we can perform the BDS to test the hypothesis of i.d.d. residuals. The findings from the BDS test are displayed in Table 7. As we can see the null hypothesis is rejected for both series. Thus the presence of nonlinearities is verified and must not be ignored.

Table 7. BDS Test.

Variable	BDS Statistic	Probability
WTI	0.015086	0.0000
HH	0.023458	0.0000

Notes: The BDS statistics presented in this table account for a dimension equal to 2. The same results were found for dimensions up to 6

5.4 Frequency Domain Causality

The establishment of the causal relationship between crude oil and natural gas prices contains valuable information for several market participants. The implementation of the standard Granger causality test provides evidence of unidirectional causality running from WTI to HH prices. However there is no evidence whether this causal relationship is valid the long-run. In order to obtain more valuable information for the nature of causality, we apply the B&C (2008) frequency domain causality test. We use the obtained residuals, for the bivariate VAR model that has been used to delinearize the series. Three different cases are presented based on the hypothesis that crude oil prices cause natural gas prices (WTI $\longrightarrow$ HH): first with conditioning on weather shocks, then with conditioning on storage shocks and finally with conditioning on both variables. We choose to include these three cases so as to incorporate the effect the two exogenous variables have to natural gas prices that may alter the causal relationship. Furthermore in order to verify causality running from HH prices to WTI prices, we also present the $(HH \longrightarrow WTI)$ case while conditioning both variables. For all causality tests the frequency interval is between 0 and π. This frequency interval can be converted to time by using: $2\pi/\omega$.

As we stated earlier, in Figure 5 the hypothesis of HH prices causing WTI prices ($HH \longrightarrow WTI$) is presented. As in the linear Granger causality test, there is absence of causality. Through the whole frequency domain, the test fails to reject the null hypothesis of no causality without conditioning any variables. The two exogenous, weather shocks and storage shocks, affect natural gas prices and including them in this test is be appropriate way to proceed.

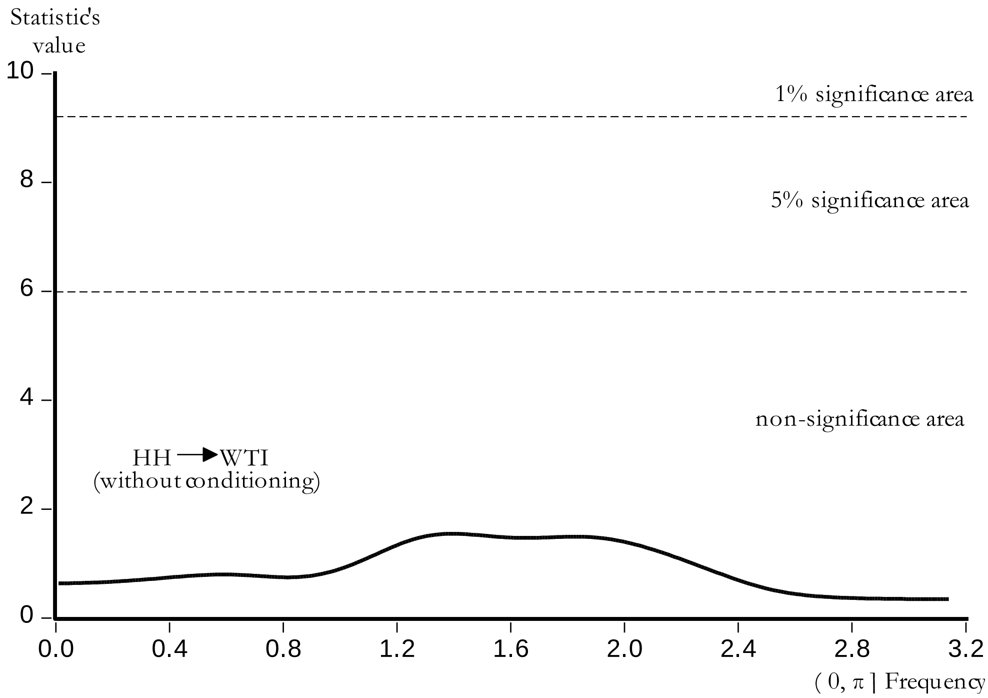


Figure 5. HH prices cause WTI prices (without conditioning)

Notes: The optimal lag-length used for this frequency domain causality test is 6 as indicated in the VAR model using the lag-length information criteria.

In the second case, presented in Figure 6, we examine the hypothesis that crude oil prices cause natural gas prices $(WTI \longrightarrow HH)$ while conditioning on weather shocks.

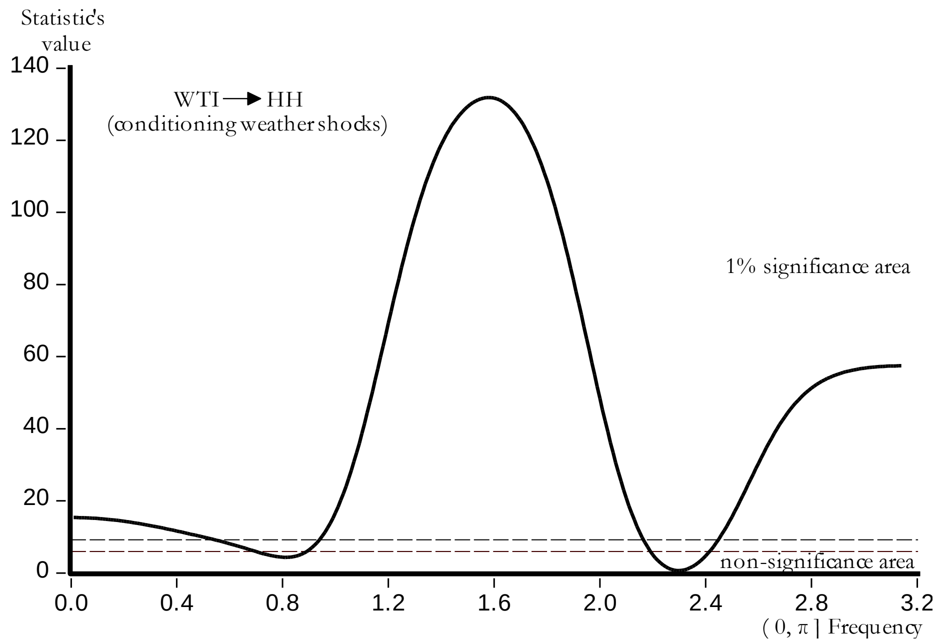


Figure 6. WTI prices cause HH prices (weather shocks conditioning)

Notes: The optimal lag-length used for this frequency domain causality test is 7 as indicated in the VAR model using the lag-length information criteria.

The results provide evidence that there is causality running from crude oil prices to natural gas prices in the long-run and definitely in the short-run. The short-run interval implies predictability between 6.6 to 2.9 days and 2.4 to 2 days while the long run for wave lengths of 11.6 days and above

We test again the hypothesis of WTI ⟶ HH but this time we condition on storage shocks. The null hypothesis of no causality can be rejected between [0.18π, 0.51π] for the 0.01 significance level as observed in Figure 7. However at the same time at the 0.05 significance level, the frequency interval, in which causality is confirmed, is larger and provides evidence of causality from 0.085π (0.27) to 0.87π (2.74). So the components of crude oil prices that may contain information about the predictability of natural gas prices are between 2.2 to 23 days.

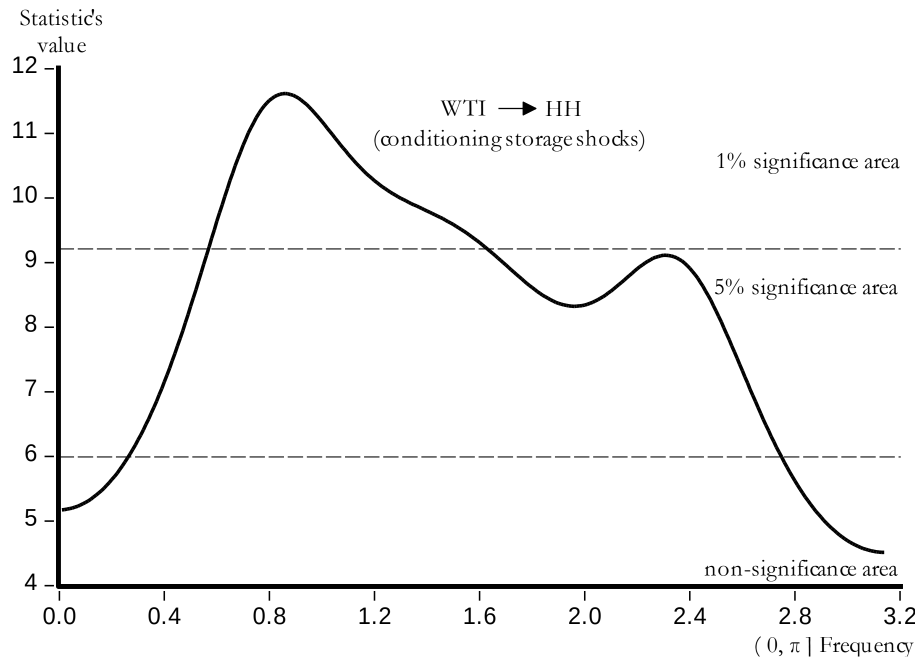


Figure 7. WTI prices cause HH prices (storage shocks conditioning)

Notes: The optimal lag-length used for this frequency domain causality test is 6 as indicated in the VAR model using the lag-length information criteria.

In the final case presented in Figure 8 we examine the causal relationship while conditioning both exogenous variables. In this model the rejection of the null hypothesis takes place in two different intervals; in [0.06π, 0.17π] and [0.61π, 0,84π] at the 0.05 significance level. As a consequence the predictability is implied in the long-run for wave length of 1 month to 11 days and in the short-run for 3.25 to 2.37 days. Until now the two cases presented support the argument that causality seems to exist. In addition to that predictive power can be mostly found in the short-run.

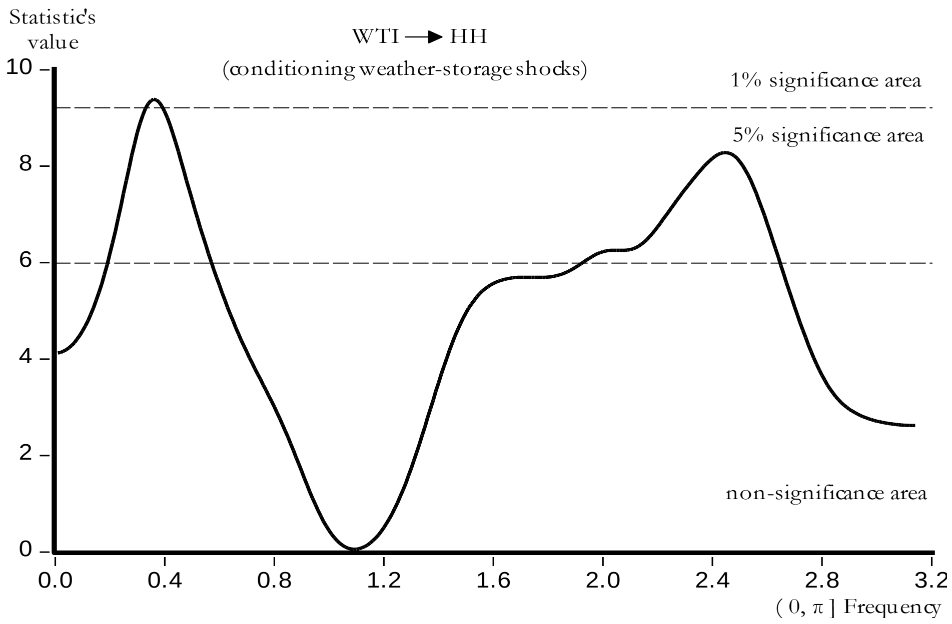


Figure 8. WTI prices cause HH prices (weather shocks and storage shocks conditioning)

Notes: The optimal lag-length used for this frequency domain causality test is 12 as indicated in the VAR model using the lag-length information criteria.

In conclusion we can support the argument that there is causality running from WTI crude oil prices to HH natural gas prices. This causality given the de-linearization process we have conducted can be attributed mainly to the non-linear components of the series. A similar pattern observed is that causality is present in the short-run in all three cases. Thus the valuable information that WTI prices contain, correspond to time interval between 2 to 6 days. On the other hand, as presented in the case with the two variables conditioning, there is also predictive power in the long-run components of crude oil prices. This finding may be attributed to the fact that the two prices are cointegrated. Overall, the B&C test provides useful information about causality where the standard Granger causality method didn't. Below in Table 8 we present the intervals in which the null hypothesis is rejected with their corresponding frequency and time frame based on the frequency domain causality test of B&C.

Table 8: Summary Table for B&C causality tests.

Causality Interval	Corresponding Frequency	Corresponding time frame(days)
Case 1: WTI ⟶ HH (conditioning weather shocks)		
(0, 0.17π]***	(0, 0.54]	Over a year – 11.6
[0.29π, 0.68π]***	[0.94, 2.16]	6.6 – 2.9
[0.78π, π]***	[2.45, 3.14]	2.4 - 2
Case 2: WTI ⟶ HH (conditioning storage shocks)		
[0.18π, 0,51]***	[0.57, 1.62]	11 - 3.8
[0.085π, 0.87π]**	[0.27, 2.74]	23 – 2.2
Case 3: WTI ⟶ HH (conditioning weather and storage shocks)		
[0.06π, 0.17π]**	[0.2, 0.57]	31.4 – 11.01
[0.61π, 0,84π]**	[1.93, 2.64]	3.25 – 2.3

Notes: The asterisks *, ** and *** denote the rejection of the null hypothesis at the 10%, 5% and 1% significance levels, respectively.

6 Conclusions

In this dissertation the relationship between Western Texas Intermediate crude oil and Henry Hub natural gas prices is examined. In order to make an extensive study of the relationship we use daily price quotations from January 1997 to August 2015 (4666 observations), as HH prices were available from 1997. In the recent years changes in the energy markets appear to have altered the relationship between some of its commodities. At this point the US natural gas market is "isolated" from the rest of the world markets and WTI crude oil has no longer the price leader role. So it is important to identify the linkages that connect them. Cointegration tests were implemented and except from standard Granger causality tests we apply the B&C causality test in the frequency domain to identify the true causal relationship.

To begin our empirical analysis we investigate the order of integration for the variables of our study (WTI, HH, weather shocks and storage shocks). The ADF, DF-GLS, PP unit root test and the KPSS stationarity test, all are implemented to reveal that: crude oil and natural gas prices are non-stationary; weather and storage shocks are stationary. To address the issue the set of our variables is mix of I(1) and I(0) variables, we applied the Rahbek & Mosconi (1999) approach to test for cointegration. The test provides evidence of one cointegration vector between the two prices, conditioning on weather and storage shocks

As cointegration is confirmed we proceed with the standard Granger causality test. Our findings suggest causality, running from WTI to HH prices and not vice versa. The standard Granger causality method implies that causality is constant, in the short-run and by definition is linear. To overcome those issues the Breitung and Cantelon (2006) method is applied. We extracted the residuals, for WTI and HH, from our SVAR model to test them for nonlinearities. The BDS test (1996) supports our initial argument of nonlinear structure presence in our series. We perform four B&C frequency domain causality tests. The first test shows that there is no causality flowing from natural gas price to crude oil prices ($\text{HH} \longrightarrow \text{WTI}$). The next three cases were based upon the hypothesis of $\text{WTI} \longrightarrow \text{HH}$ causality while conditioning on weather shocks and storage shocks separately and then the two variables together. All three tests provide evidence that there is causality. The information in the WTI prices that can predict HH prices are between 2 to 6 days, implying short-run causality. Adding to that, the long-run components may contain valuable information where causality was present in a time interval between 11 days to a month. Overall by performing different B&C tests we can summarize that: a) there is unidirectional causality running from WTI prices to HH prices b) causality isn't constant through the whole time interval c) by conditioning on different variables we can see that results a relatively different d) predictive power in crude oil prices can be found in the short-run as well as in the long-run e) the identified causality can be attributed to the non-linear components of the series.

These linkages between crude oil and natural gas prices are of major importance for policy makers, hedgers and stockholders in order to adopt the appropriate strategy. These strategies are bases upon estimation of these prices. So, the identification and estimation of the relationship is a task that affects decision making. In this dissertation we attempt to examine the current state in U.S. However, recent developments in the US could be used for future work. The US plans to lift the multi-year ban on oil and natural gas exports and attempts to build LNG export infrastructure. This is a development that could change the global energy market scheme and will be interesting to see its implications to the relationship between the world's energy commodities.

Acknowledgments

"I would like to express my sincere gratitude to my supervisor, Dr. Theologos Dergiades, for his guidance, suggestions and time he spend throughout the Master's course. Also I would like to thank my family for the continuous support this whole time."

7 References

Asche, F., B. Misund and M. Sikveland (2013). The relationship between spot and contract gas prices in Europe. *Energy Economics* 38, 212-217.

Asche, F., A. Oglend and P. Osmudsen (2012). Gas versus oil prices the impact of shale gas. *Energy Policy* 47, 117-124.

Asche, F., P. Osmundsen and M. Sandsmark (2006). The UK market for natural gas, oil and electricity: Are the prices decoupled? *The Energy Journal* 27(2), 27-40.

Atil, A., A. Lahiani and D.K. Nguyen (2014). Asymmetric and nonlinear pass-through of crude oil pricces to gasoline and natural gas prices. *Energy Policy,* 65, 567-573.

Bachmeier, L.J. and J.M. Griffin (2006). Testing for market integration. Crude Oil, Coal, and Natural Gas. *The Energy Journal* 27(2), 55-71.

Bekiros, D.K. and C. Diks (2008). The relationship between crude oil spot and future prices: Cointegration, linear and nonlinear causality. *Energy Economics* 30, 2673-2685.

Bencivenga, C., G. Sargenti and R.L. D'Ecclesia (2010). Energy Markets: crucial relationship between prices. *Mathematical and Statistical Methods for Actuarial Sciences and Finance* 23-32.

Breitung, J. and B. Candelon (2006). Testing for short and lon-run causality: a frequency domain approach. *Journal of Econometrics* 132, 363-378.

Brigida, M. (2014). The switching relationship between natural gas and crude oil prices. *Energy Economics* 43, 48-55.

Brown, S.P.A. and M.K. Yücel (2008). What Drives Natural Gas Prices. *The Energy Journal* 29(2), 45-60.

Brown, S.P.A. and M.K. Yücel (2009). Market Arbitrage: European and North American Natural Gas Prices. *The Energy Journal, Special Issue. World Natural Gas Markets and Trade: A Multi-Modeling Perspective* 30, 167-185.

Dickey, D.A. and W.A. Fuller (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association* 74, 427-431.

Elliot, G., T.J. Rothenberg, and J. Stock, (1996). Efficient Tests for an Autoregressive Unit Root. *Econometrica* 64(4).

Erdős, P. (2012). Have oil and gas prices got seperated? *Energy Policy* 49, 707-718.

Granger, C.W.J. (1969). Investigating Causal Relations by Econometric Models and Cross-spectral Methods. *Econometrica* 37(3), 424-438.

Hartley, P.R. and K.B. Medlock III (2014). The Relationship between Crude Oil and Natural Gas Prices: The Role of the Exchange Rate. *The Energy Journal* 35(2).

Hartley, P.R., K.B. Medlock III and J. Rosthal (2008). The Relationship of Natural Gas to Oil Prices. *The Energy Journal* 29(3), 47-65.

He, Y., S. Wang and K.K. Lai (2010). Global economic activity and crude oil prices: A cointegration analysis. *Energy Economics* 32(4), 868-876.

Ji, Q. and Y. Fan (2015). Dynamic integration of world oil prices: A reinvestigation of globalisation vs. regionalisation. *Applied Energy* 155, 171-180.

Ji, Q., J.-B. Geng and Y. Fan (2014). Seperated influence of crude oil prices on regional natural gas import prices. *Energy Policy* 70, 96-105.

Johansen, S. (1995). *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models.* Oxford: Oxford University Press.

Kwiatkowski, D., P. Phillips, P. Schmidt and Y. Shin (1992). Testing the null hypothesis of stationarity against the alternative of a unit root. *Journal of Econometrics* 54, 159-178.

Lin, B. and J. Li (2015). The spillover effects across natural gas and oil markets: Based on the VEC-MGARCH framework. *Applied Energy* 155, 229-241.

Mjelde, J.W. and D.A. Bessler (2009). Market integration among electricity markets and their major fuel source markets. *Energy Economics* 31, 482-491.

Mohammadi, H. (2011). Long-run relations and short-run dynamics among coal, natural gas and oil prices. *Applied Economics* 43(2), 129-137.

Mu, X. (2007). Weather, storage and natural gas price dynamics: Fundamentals and volatility. *Energy Economics* 29(1), 46-63.

Nick, S. and S. Thoenes (2014). What drives natural gas prices? - A structural VAR approach. *Energy Economics* 45, 517-527.

Panagiotidis, T. and E. Rutledge (2007). Oil and gas markets in the UK: Evidence from a cointegrating approach. *Energy Economics* 29(2), 329-347.

Pfaff, B. (2008). VAR, SVAR and SVEC models: Implementation within R package vars. *Journal of Statistical Software* 27(4), 1-32.

Phillips, P.C.B. and P. Perron (1988). Testing for a Unit Root in Time Series Regression. *Biometrika* 75(2), 335-346.

Rahbek, H. and R. Mosconi (1999). Cointegration rank inference with stationary regression in VAR model. *Econometrics Journal* 2, 76-91.

Ramberg, D. and J. Parsons (2012). The Weak Tie Between Natural Gas and Oil Prices. *The Energy Journal* 33(2).

Serletis, A. and J. Herbert (1999). The message in North American energy prices. *Energy Economics* 21(5), 471-483.

Siliverstovs, B. G. L'Hegaret, A. Neumann and C. von Hirschhausen (2005). International market integration for natural gas? A cointegration analysis of prices in Europe, North America and Japan. *Energy Economics* 27(4), 603-615.

Villar, J.A. and F.L. Joutz (2006). *The relationship between crude oil and natural gas prices.* Washigntpn D.C.: Energy Information Administration, Office of Oil and Gas.

Wolfe, M.H. and R. Rosenman (2014). Bidirectional causality in oil and gas markets. *Energy Economics* 42, 325-331.