# The Effects of Weather and Storage Shocks on Natural Gas Price in the UK

University of Stavanger

Universitetet i Stavanger

DET SAMFUNNSVITENSKAPELIGE FAKULTET,
HANDELSHØGSKOLEN VED UIS
MASTEROPPGAVE

| | |
|---|---|
| STUDIEPROGRAM:<br><br>Master – Økonomi og Administrasjon | OPPGAVEN ER SKREVET INNEN FØLGENDE SPESIALISERINGSRETNING:<br>Applied Finance<br><br>ER OPPGAVEN KONFIDENSIELL?<br>(NB! Bruk rødt skjema ved konfidensiell oppgave) |

TITTEL:
The Effects of Weather and Storage Shocks on Natural Gas Price in the UK

ENGELSK TITTEL:
The Effects of Weather and Storage Shocks on Natural Gas Price in the UK



| FORFATTER(E) | | VEILEDER: |
|---|---|---|
| Studentnummer: | Navn: | Bård Misund |
| 223370<br>………………… | Thor Alexander Landa Hansen<br>……………………………………. | |
| 222791<br>………………… | Jon Magne V. Sørholt<br>……………………………………. | |



OPPGAVEN ER MOTTATT I TO – 2 – INNBUNDNE EKSEMPLARER

Stavanger, ……/…… 2015 Underskrift administrasjon:……………………………

## Abstract:

In this thesis, we investigate the relationship between temperature deviations, storage level, and the price of natural gas in the United Kingdom. By applying these models, we expect to obtain a better understanding of the relationship between these factors, and be able to check the statistical relevance of our research problem. Increased comprehension about the relationship between weather, storage, and natural gas can assist market participants' decision-making. The analysis is based on daily data observations of 5 years from 2010 up to 2015.

We created the three main variables, natural gas returns, weather shock, and storage level deviation. Other variables such as Treasury bills, Brent oil, and S&P 500 are gathered and reported as the daily change to reflect the natural gas returns. We implemented a GARCH-model to estimate the volatility of the natural gas futures price. We then creates a VAR model to illustrate the tridimensional relationship between the main variables, enabling the use of IRF to simulate shocks and estimate the respond to changes in the economic environment.

The VAR model are unable to provide significant evidence of an integrated relationship, whereas the IRF model found results implying that weather and storage shock can affect natural gas returns, but the response may not materialize, rendering the results ambiguous.

This page intentionally left blank

## Preface

We created this thesis as a concluding part of our Master of Science in Business and Administration, with specialization in Applied Finance at the University of Stavanger Business School.

This master thesis include topics beyond the scope of the curriculum, which have both been struggling and exiting. The last six months have provided great challenges, tough defeats, slow progression through unknown territory, and inspiring victories, all leading up to the end result, namely this thesis. The famous words of Julius Caesar "Veni, Vidi, Vici" sum up how we both felt after finishing this thesis.

We would like to thank our supervisor Bård Misund for inspirational assistance throughout the process of the thesis. We would also like to thank William Gilje Gjedren for much appreciated help understanding R.

## Content

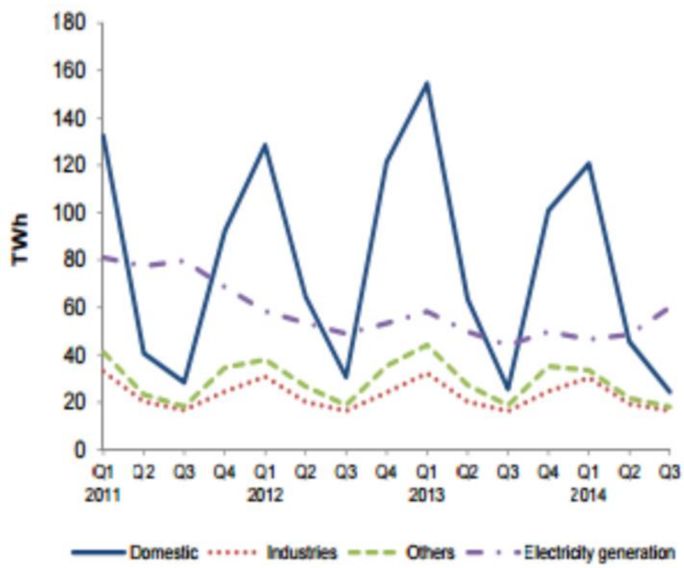
## Figures

## Tables

# 1. Introduction

We are going to examine the dynamics of natural gas price (NGP) in the UK market. As it is normal for the public to utilize natural gas for heating in the UK, and domestic usage are the largest demand group, in 2013 they represented approximately 40% of natural gas consumption within UK. Power stations are the second largest consumer, and represent approximately 23% of natural gas consumption within UK. A total overview of natural gas distribution for 2013 is in Figure 1-3.

*Figure 1-1: Demand for natural gas UK*




*Obtained from* (Department of Energy & Climate Change, 2014), *figure 4.6, the amount are listed in terawatt per hour (TWh)*

Basic economic theory state that increased demand results in increased price. The natural gas demand is highly cyclical, baring evidence of seasonality, illustrated in Figure 1-1. We can see that the demand increases during the winter months, when the temperature is low, and decreases during warmer periods of the year. This is a consequence of natural gas being primarily directed toward heating and thus affected by temperature variations in UK. Reduced temperatures result in increased demand, and thus increased NGP.

Natural gas storage reservoirs exhibit similar properties, displayed in figure 1-4, as the demand, being subject to cyclical changes based on seasonality. The storage reservoirs can compensate for unexpected increases in the demand, providing a mitigating effect on abnormal temperature behavior. If a weather shock occur, resulting abnormally low temperatures, there will be higher demand for natural gas, as the need for heating increases. The produced amount of natural gas are unable to meet this increased demand, resulting in excess demand, and increased willingness to pay for natural gas. This will result in higher NGP. The stored natural gas can be used to compensate for the excess demand, mitigating the NGP reaction.

With this in mind, we started wondering whether it was possible to elaborate this relationship, in a statistically meaningful manner, to isolate temperature and storage as contributing factors to changes in NGP.

After reading an article by Mu (2007), where a similar relationship were proposed and investigated in the U.S. market, we were inspired to do the same here in Europe. We were unable to find work focusing on this tridimensional price dynamic in the UK market, and as Mu (2007) argues, there is a lack of research done on the relationship between weather, storage, and the returns of natural gas. This thesis will contribute to increased understanding of factors affecting the returns in the European market, especially weather and storages contribution to changes in natural gas returns. This can prove to be useful when estimating price forecasts, as it may improve the precision of the estimate. This can in turn be used by day traders, speculates, and other participants in the futures market.

We have based our analysis on an article written by Mu (2007), where he isolates the effect of weather and storage shocks on NGP in the U.S. While his study examines the U.S. market, the results might not be the same here in Europe. According to the findings of Haff, Lindqvist, & Løland (2008) there is a distinct difference in the risk premium on natural gas forwards contracts traded in the U.S. and the UK. The risk premium on the forward price in the UK are poisitve, while it is negative in the U.S. market. UK has a good liquidity in their natural gas market but they are not at the level of US as stated in Heather (2010). UK had a churn rate[1] of 20 in 2007,

[1]Churn rate is a measurement of trading a commodity goes through from seller to final buyer. A market with a churn rate of 10, or above, is believed to have reached maturity.

while United States had an even higher churn rate at almost 30. The UK churn rate fell in the following years, but went back to approximately 20 in 2010.

We would therefore like to examine if this effect applies to the European natural gas market. For empirical analysis, it would be optimal to use a larger selection of gas markets in Europe, not just a single market such as the UK. However, the UK market is by far the largest and most liquid market in Europe, and we believe it is sufficient to focus on UK data. We have applied the same method as Mu (2007) to estimate our weather variable, in order to see if the weather effect in the UK coincides with the finding in the U.S. market.

To elucidate this relationship, we started by estimate a model, described in Table 4-1: Initial regression model, then we estimate a Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model for the return series, implement a Vector Autoregression (VAR) model, and at last apply the Impulse Response Function (IRF) to this model. Through these analyses, we aim to provide sufficient evidence elucidate the following research questions:

1. How do the natural gas price react to shocks in weather and storage variables in European Markets, exemplified by the UK gas market?
2. How does the NGP reaction compare with the findings of Mu (2007) for the U.S. market?

We found there to be no statistically significant reactions in return due to shocks in either weather or storage. However, we did find a relation where weather affect storage.

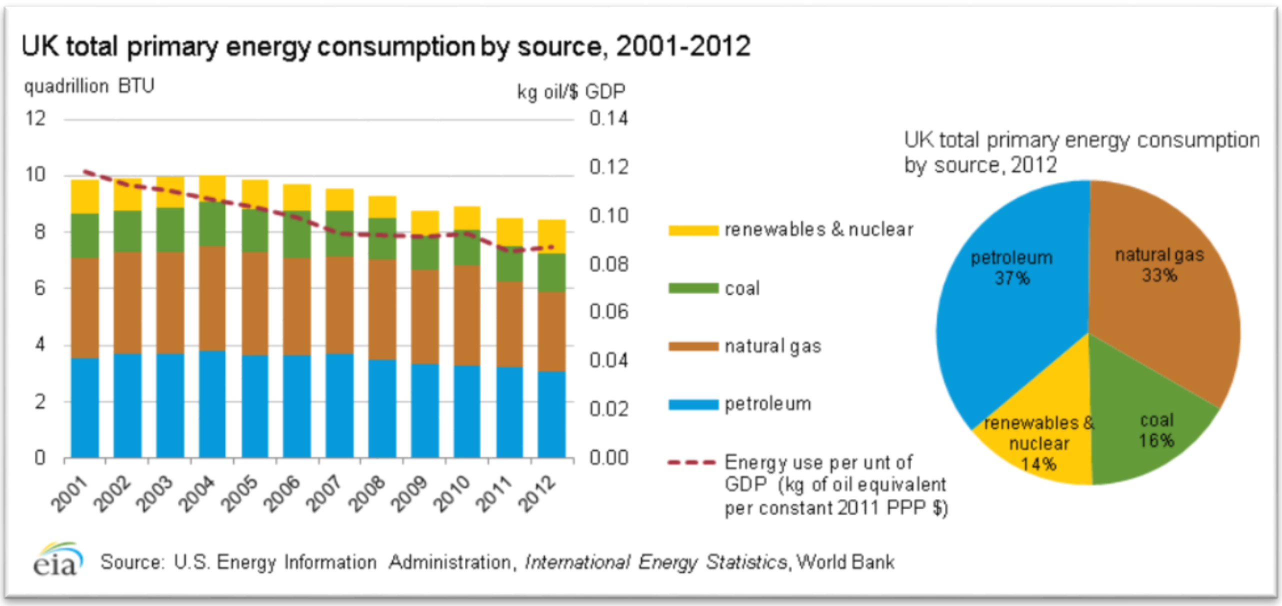
Mu (2007) found a significant weather effect on the conditional means of natural gas returns, whereas we were unable to provide statistically significant evidence in favor of this relationship

The remainder of this thesis is divided into 6 chapters; the first chapter is the introduction where we will discuss the UK natural gas market. Then we will provide some theoretical insight to the models used in this paper, which sums up chapter 2. Chapter 3 contains the methodology used to apply the theories introduced in chapter 2 to our data. The analysis, where we will list and interpret the results obtained from the different tests and model, is located in chapter 4. Chapter 5 consists of the conclusion. At last, we discuss possible improvements and limitations, which is located in chapter 6.

## 1.1.UK gas market

Gas currently forms an integral part of the UK's power generation mix and is a reliable, flexible source of electricity. Using gas as a fuel in the UK's power stations currently provides a significant proportion of the electricity generation, around 40% in 2011. Gas sets the electricity price for most of the year, as generation from gas is used to meet the peaks in the UK electricity demand. The government expects that gas will continue to play a major role in the UK electricity mix over the coming decades, alongside low-carbon technologies as they decarbonize their electricity system[2].

*Figure 1-2: UK total primary energy consumption 2001-2012*




As we can see from Figure 1-2, natural gas constituted 33% of the total energy consumption in the UK in 2012. This indicates that natural gas is an important source of energy, which seems to have withheld its position quite well during the last decade, holding a large market share during the period of 2001-2012.

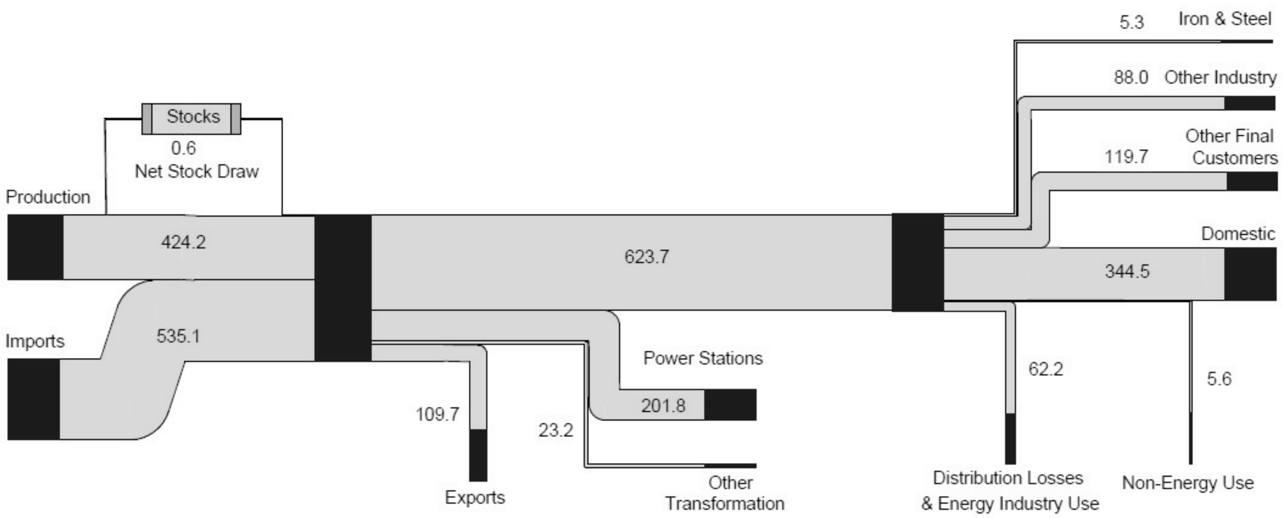
UK are the second largest contributor of natural gas in the European union but their supply have suffered a long term decline since its peak in 2000, and they are now only able to supply under

---

[2] http://www.eia.gov/countries/cab.cfm?fips=uk, 02.03.15

half of the demand themselves, which have made the UK reliant on importing gas in later years. They have also not invested in facilities to build up a large reserve of natural gas and are thus exposed toward disruptions in the supply line. As of 21 of February 2013 UK had a storage capacity for 15 days' supply, compared to other gas using countries' in Europe, where France have 99 days and Germany have 122.

*Figure 1-3: Acquisition and consumption of natural gas in the UK*




This figure show the natural gas flow chart for 2013, the amount are listed in TWh, excluding colliery methane (MacLeay, et al., 2014)

### 1.1.1. Demand

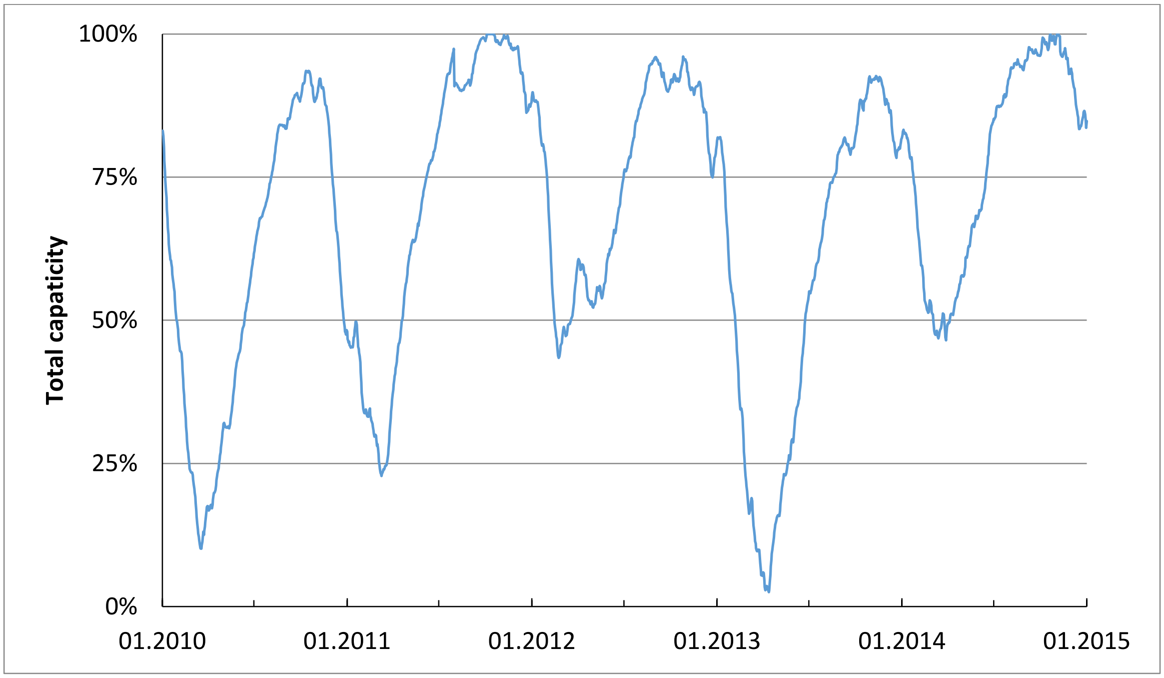
Figure 1-3 provides an overview of the consumption of natural gas, divided into sectors. Power stations generate electricity that can be used for private heating and cooling, and Domestic represents the private sector's consumption. It is reasonable to assume that both of these sectors contribute to the private sectors demand for heating and cooling, which stands for approximately 64.3%[3] of the total consumption of natural gas in 2013.

The demand for natural gas increases as the temperature decreases, which is a result of an increased need of heating. The amount of natural gas supplied during the winter are insufficient to account for the increased demand, which mean that the withdrawal rate surpasses the injection

[3] Excluding exports: $\frac{(Domestic + Power\ station)}{(Production + Imports - Exports)} = 64.3\%$

rate, and result in a reduced reservoir level. The storage level increases during warmer periods, where the demand decreases, resulting in an injection rate greater than the withdrawal rate. Figure 1-4 provides a graph where these effects are visible.

*Figure 1-4: UK natural gas storage level from 2010 to 2015*

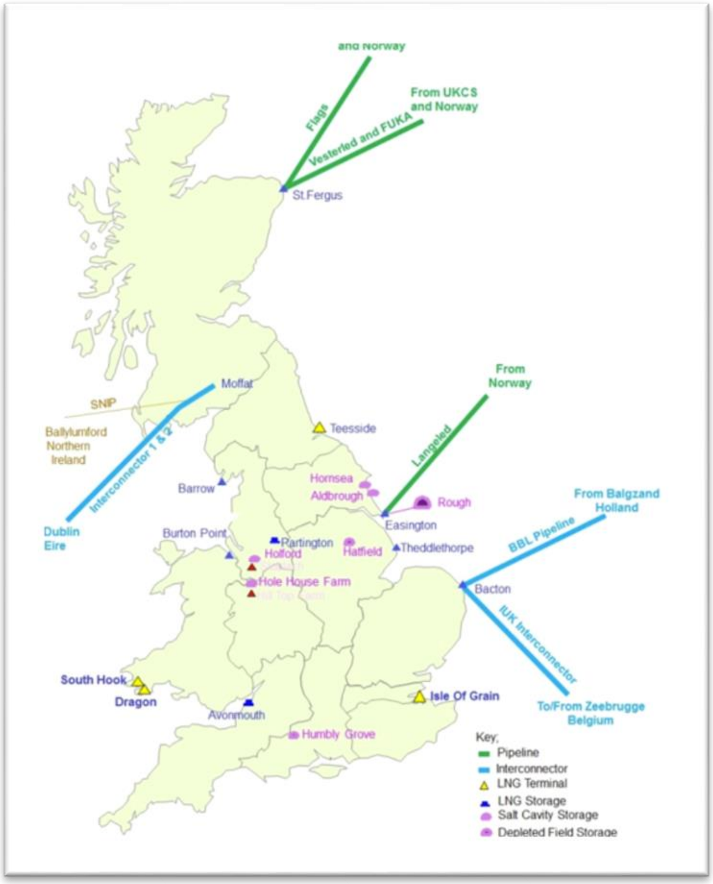



The graph shows reservoir levels as a percentage of maximum capacity.

### 1.1.2. Supply

UK has nine reception points where they receive natural gas for quality control and transportation throughout the country.

*Figure 1-5: UK pipelines and storage (gov.uk)*




The Norwegian Continental Shelf (NCS) deliver Natural gas to the terminals St. Fergus and Easington, which accounts for approximately 57% of the total UK imports. Balgzand Bacton line (BBL) is the interconnector line between the UK and Holland. These pipelines are used to for imports exclusively, and are referred to as one-way-pipelines.

The Interconnector UK (IUK) pipeline is a two-way pipeline that can deliver and receive natural gas, this pipeline is mostly used for exports during the summer and imports during winter.

As of 2013 UK produced 424 TWh and account 44.2% of the natural gas consumption, while the remaining 55.8%, 535 TWh, were imported (MacLeay, et al., 2014).

## 2. Theory

This chapter is designed to introduce models used in this thesis, and the theories they are based upon. Our dataset contains daily observations over several years, which implies that we use time series data, where the main focus is directed towards the price of natural gas.

In chapter 2.1 we will discuss different tests applied to ensure that the time series data meets requirements set forth to enable hypothesis testing. The objective is to obtain a dataset that includes variables with constant mean, variance, and covariance, construct robust standard errors that can correct for heteroskedasticity, and remove any presence of serial correlation.

Part 2.2 consist of the theory behind the models we have decided to use in this thesis. These models are the Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model, Vector Autoregression (VAR) model, and Impulse Response Function (IRF) model. We have also included the Forecast Error Variance Decomposition (FEVD) to help interpret the latter model.

*Figure 2-1: Spot and 1-month futures contract for NGP*

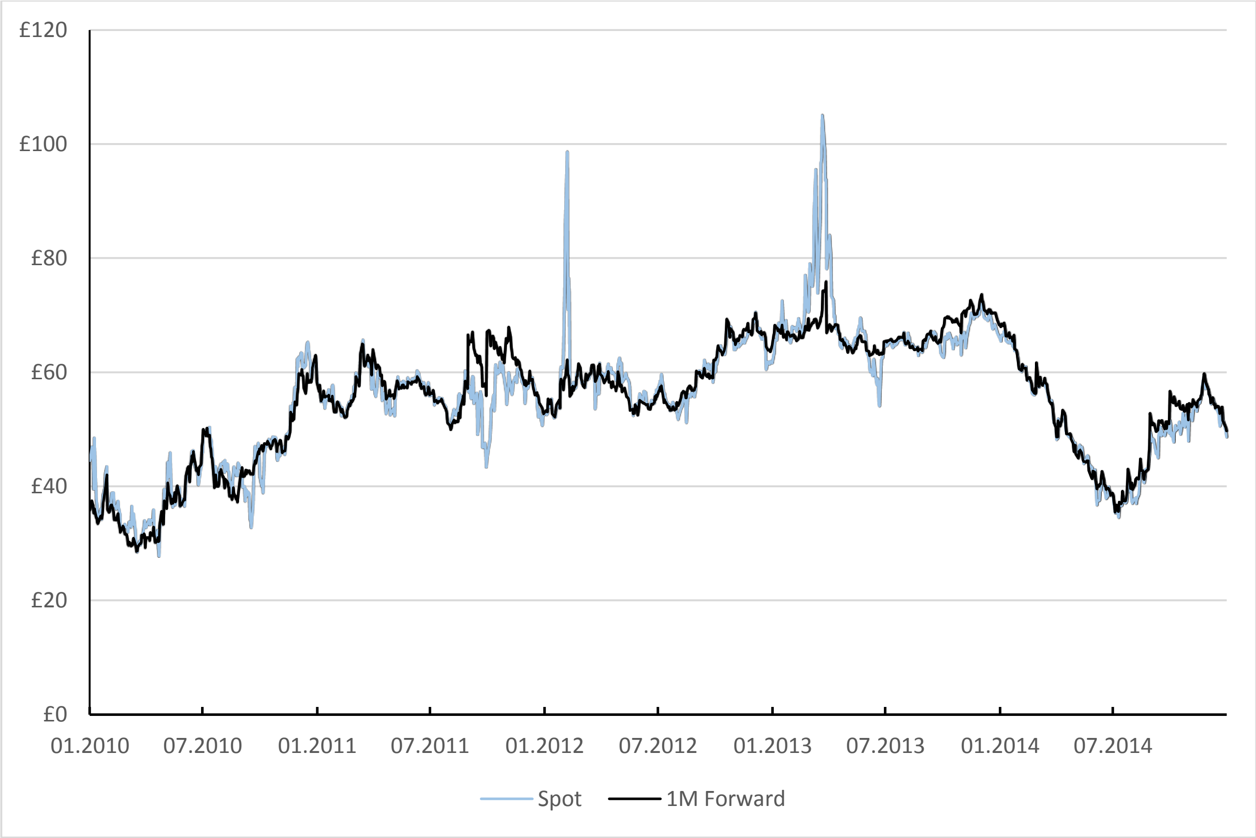


Figure 2-1 shows how the spot price and the 1-month futures price for NGP behaves over time, and we can see that the spot price displays some spiky behavior. These extreme spikes may be a result of incorrect reporting, as private firms, which are not obligated to report the correct spot prices, gather these data. The 1-month futures price, on the other hand, is reported by the stock exchange, and therefore represents the correct price level at the corresponding date. Due to this, we have decided to use the 1-month futures price, referred to as *M1*, in the following analysis.

The applied data consist of observations corresponding to a specific day, which means that it is time series data. Figure 2-1 illustrate this, where each point along the lines inside the graph represents both a value and a date.

These two price series appears to exhibit a mean reversion, which mean that there are a long-term mean that the price are reverting too.

## 2.1. Diagnostic Tests

In this part, we are going to do some preparatory work to ensure that the data are ready for further analysis. This is done through several tests, which are introduced below.

### 2.1.1. Stationary variables

Stationary data means that the variables included in the analysis have means, variances and covariance that are constant over time. This implies that each of these factors are equal, independent of what period they represent. This is necessary for being able to use the model to predict what will happen in the future.

If this assumption is violated, we have non-stationary data, resulting in unpredictable model outcome. The results obtained when using non-stationary data can become spurious in that they can indicate relationship between variables where it does not exist.

There are different forms of non-stationary time series data, and we need to be able to distinguish between these to apply the correct transformation of the data for the different variables.

First we have pure random walk as shown in equation (1). Where $Y_t$ is the estimated value at time $t$, and are equal to the value at time $Y_{t-1}$, plus a stochastic component containing white noise $\varepsilon_t$.

$$Y_t = Y_{t-1} + \varepsilon_t \tag{1}$$

The pure random walk model can be developed further to three different equations.

By adding a constant measurement for the drift, $\alpha_0$, into equation (1), we get equation (2). To account for the possibility of a non-stationary deterministic trend, we include a trend coefficient, $\beta t$, to obtain equation (3).

$$Y_t = \alpha_0 + Y_{t-1} + \varepsilon_t \tag{2}$$

$$Y_t = \alpha_0 + \beta t + \varepsilon_t \tag{3}$$

$$Y_t = \alpha_0 + Y_{t-1} + \beta t + \varepsilon_t \tag{4}$$

When combining equation (2) and (3) we get equation (4) that are a random walk with drift and deterministic trend.

Random walk is a non-mean reverting process that can move away from the mean either in a positive or negative direction, and the variance evolves over time, thus it cannot be predicted (Wooldridge, 2012).

### 2.1.2. Augmented Dickey-Fuller test

We implementer the Augmented Dickey-Fuller (ADF) test to check whether the variables are stationary or not. The test assumes that the variable is affected by unit root, which implies that the variable is non-stationary. This means that the alternative hypothesis is that the data is stationary, which is the desired result of the test.

The test is divided into three main components: Unit root, Unit root with drift, and Unit root with drift and trend displayed in equation (5) – (7).

$$\nabla Y_t = \delta Y_{t-1} + \varepsilon_t \tag{5}$$

$$\nabla Y_t = \alpha_0 + \delta Y_{t-1} + \varepsilon_t \tag{6}$$

$$\nabla Y_t = \alpha_0 + \delta Y_{t-1} + \beta t + \varepsilon_t \tag{7}$$

The null hypothesis, $H_0$, in all tests assumes that $\delta = 0$, which mean that there are unit root present in the data, thus the data are not stationary. The null hypothesis for unit root with drift assumes that there are unit root and no drift present at the same time ($\delta = \alpha_0 = 0$), while the last test assumes that there are unit root, no drift and no trend in the data ($\delta = \alpha_0 = \beta = 0$).

When the test statistic and the representative critical values are obtained, we can see whether the variable are stationary or not, through testing. If the absolute value of the test statistic is less than the absolute value of the critical value, we fail to reject the hypothesis of non-stationary data. If this is the case, we need to difference the data to obtain stationary variables (Enders, 2009).

### 2.1.3. Breusch-Pagan Test

The Breusch-Pagan (1979) test is a diagnostics test of a regression model, where the goal is to see if there is presence of heteroskedasticity. Heteroskedasticity is defined as a non-constant variance over a period of time. The test assumes that the model are homoskedastic, so if we fail to reject $H_0$, the test provides evidence supporting this hypothesis. If we end up rejecting the null hypothesis, we obtain evidence suggesting that there are heteroskedasticity in the regression model (Wooldridge, 2012).

### 2.1.4. Breusch-Godfrey Test

The Breusch-Godfrey (1978) test to detect presence of higher order serial correlation (AR(q)) illustrated in equation (8).

$$Y_t = \rho_1 Y_{t-1} + \cdots + \rho_q Y_{t-q} + e_t, \qquad t = 1, 2, \dots \tag{8}$$

The error, $e_t$, are white noise with variance $\sigma^2$ and $\rho_1 \dots \rho_q$ are parameters. The $q$ denotes the amount of lags included in the test.

Autocorrelation is present in the data series if the error terms in the regression are serially correlated across time. The test assumes that there is no serial correlation, and the result of the test has similar properties as the Breusch-Pagan Test (Wooldridge, 2012).

### 2.1.5. White correction

White (1980) proposed a method for correcting the standard errors of the coefficients in the regression model, to produce heteroskedasticity consistent standard errors (HCSE). Theory states that a regression model suffering from heteroskedasticity may produce incorrect significance level for the different variables, through a misleading estimate of the included variables' standard errors. These standard errors have a tendency to be under-predicted, resulting in increased chance of getting significant values, when this is not the case. Thus, the White correction produces robust standard errors, enabling hypothesis testing (Wooldridge, 2012).

## 2.2. Models used for analysis

When the requirements presented in **Error! Reference source not found.** are satisfied, we can move on to implement the models we plan to use to utilize. We are now going to present these models, and explain how they work.

### 2.2.1. The GARCH(1,1) Model

In our thesis we will apply the GARCH(1,1) model, developed by Bollerslev (1986), to estimate volatility. This model is a sophisticated, yet simple tool that allows for a flexible lag structure, accounts for long-term volatility, and conditional variance that may be dependent on own lag. In short, estimating a GARCH model consists of three steps:

1) Estimate fitted values for the autoregressive model
2) Compute autocorrelations of the error terms
3) Significance test

Equation (9) presents the GARCH(1,1) model (Hull, 2012).

$$\sigma_n^2 = \gamma V_L + \alpha u_{n-1}^2 + \beta \sigma_{n-1}^2 \quad (9)$$

The parameters gamma ($\gamma$), alpha ($\alpha$) and beta ($\beta$) are the weights assigned to each of the parts of the GARCH equation and will together sum to one, ($\gamma + \alpha + \beta = 1$). The first part $\gamma V_L$ displays the long-run average variance. The second part $\alpha u_{n-1}^2$ implements a time lag effect from the previous periods returns and the last section $\beta \sigma_{n-1}^2$ are the time lag effect from the previous period's volatility.

The simplified and most used GARCH(1,1) model, where alpha and beta sum to 1, ($\alpha + \beta = 1$), concentrate on the most recent observation of both return and variance to estimate volatility. If we set $\omega = \gamma V_L$, we can rewrite the model and get:

$$\sigma_n^2 = \omega + \alpha u_{n-1}^2 + \beta \sigma_{n-1}^2 \quad (10)$$

Using equation (10), we can use the parameters $\omega$, $\alpha$ and $\beta$ to obtain the long-run variance level ($V_L$) and mean reversion rate ($\gamma$).

$$V_L = {}^{\omega}/_{1 - \alpha - \beta} \quad (11)$$

$$\gamma = 1 - \alpha - \beta \quad (12)$$

The model is stable if alpha plus beta is less than one; if not then the long-term variance becomes negative.

### 2.2.2. Vector Autoregression (VAR)

The vector autoregression model is a flexible and reliable model for analyzing multivariate time series. This approach is used to estimate the linear cointegration among endogenous variables. The model estimates a regression for each of these variables, as a function of both endogenous variables, including a predetermined amount of lagged values for each variable, and a set of exogenous variables. This can provide a better estimate for $Y_t$, if the model suffers from autocorrelation or there are delayed effects between endogenous variables.

To illustrate this, we assume a two dimensional VAR(1)-Model, (Füss, 2007):

$$Y_{1,t} = \alpha_{11}Y_{1,t-1} + \alpha_{12}Y_{2,t-1} + \varepsilon_{1,t} \tag{13}$$

$$Y_{2,\mathrm{t}} = \alpha_{21}Y_{1,\mathrm{t}-1} + \alpha_{22}Y_{2,\mathrm{t}-1} + \varepsilon_{2,\mathrm{t}} \tag{14}$$

In a two dimensional model, equation (13) and (14), the dependent variables are $Y_{1,t}$ and $Y_{2,t}$ and move along the time series where $t = 1, 2, \ldots, T$.

To best explain the parameters and variables of equation (13) and (14), we create equation (15) with the variables *i* and *j*, where the dependent variable, $Y_{i,t}$, are based on it's own lag, $Y_{i,t-1}$, weighted by the parameter $\alpha_{1i}$ and the other endogenous variable $Y_{\mathrm{j,t}-1}$ with same amount of lag and a parameter, $\alpha_{1j}$. The equation (15) model is a VAR(p) model with *p* amount of lags.

$$Y_{i,t} = \alpha_{1i}Y_{i,t-1} + \alpha_{1j}Y_{j,t-1}, \ldots, \alpha_{pi}Y_{i,t-p} + \alpha_{pj}Y_{j,t-p} + \varepsilon_{\ t} \tag{15}$$

Matrix notation:

$$Y_t = A_1Y_{t-1} + \varepsilon_t \tag{16}$$

$$A_1 = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \tag{17}$$

The VAR(p) model are unable to conclude whether there are causal relationships between the endogenous variables, but it allow interpretation of the dynamic interaction of the explanatory variables, $Y_{1,t-1}$ and $Y_{2,t-1}$. The historic data are used to explain the development of the series.

The VAR(p) model can be extended to include additional components such as a constant term, trends or seasonality, and test whether these deterministic factors are significant.

### 2.2.3. Impulse Response Function (IRF)

An Impulse Response Function (IRF) is a methodology for investigating the dynamic effects of different variables with respect to the response variable. The IRF simulate a one standard deviation shock in endogenous variables, and then reports back how this shock affects the response variable over time. This simulated shock series are compared with the actual time series, without a shock, to give a graphical representation of the simulated shock. The impulse response sequence is then plotted as the discrepancies between these two series.

To illustrate this, we continue to use the VAR(1) model introduced in equation (13) and (14), and derive the IRF model similarly to Roland Füss:

> Initially, in $t = 1$ we assume a shock in the error term $\varepsilon_{1,1}$, of the first equation. This shock has a direct effect on $Y_{1,1}$, of exactly the same amount. Whereas $Y_{2,1}$, is not effected, assuming that $\varepsilon_{2,t} = 0$ with $t = 1, \dots, T$. In the second period (*t = 2*), the original shock has still an effect over the lagged value of $y_1$. The effect on $Y_{1,2}$, is $\alpha_{11}\varepsilon_{1,1}$, and the effect on $Y_2$, is $\alpha_{21}\varepsilon_{1,1}$. In the third period the effect on $Y_{1,3}$ is not only $\alpha_{11}(\alpha_{11}\varepsilon_{1,1})$, but also $\alpha_{12}(\alpha_{21}\varepsilon_{1,1})$. Accordingly, the effect on $Y_{2,3}$ is $\alpha_{21}\left(\alpha_{21}\varepsilon_{1,1}\right) + \alpha_{22}(\alpha_{21}\varepsilon_{1,1})$. Thus, it is possible to obsess the effect of a non-recurring shock in one variable, to all variables over time. (Füss, 2007) pp. 17.

One could summarize the result in:

$$Y_t = \sum_{k=0}^{\infty} C_k \varepsilon_{t-k} \tag{18}$$

With $C_0 = I$ (Vector-Moving-Average Process) and where $C_{\mathrm{k}}$ are the weight of past stocks.

In this approach to the IRF, one assumes that the error terms in the two different equations are uncorrelated, which is a restricted assumption. A shock in only one equation is not a realistic adjustment of the shock process, which can be controlled for by applying the orthogonal IRF sequence. The orthogonalized IRF approach implies that the model is modified to obtain uncorrelated, orthogonal, error terms, which is provided by equation (19).

$$Y_t = \sum_{k=0}^{\infty} \tilde{C}_k v_{t-k} \tag{19}$$

With$\tilde{C}_k = C_k * G$, where G is a matrix with the properties of the Cholesky decomposition. The error terms of the modified system are $v_{t-k} = G^{-1} * \varepsilon_{t-k}$.

#### 2.2.3.1. Forecast error variance decomposition (FEVD)

FEVD is a decomposition of the error variance which is a supplement designed to aid interpreting the fitted Vector Autoregression (VAR). The FEVD give insight to each endogenous variable's contribution of information in the autoregression. FEVD predict how variable *k* is affected by a shock in variable *j*. This decomposition expose which of the *j* variables that forces variable *k* to change. (Pfaff, 2008)

# 3. Method

The objective is to fit the theoretical models introduced in chapter 2 to our dataset, this approach is described in this chapter.

Our database consists of data from 2010 to the end of 2014. We would like to have used data for a longer period, but were unable to obtain storage data form the source National grid. They actually started collecting storage data in 2009, but many observations for this period were omitted, so we chose to exclude this year.

We decided on using excel for sorting and setting up the data, and the programming is done in R, with the following packages[4]:

*Table 3-1: R-packages*

| **R-Package** | **Package title** |
|---|---|
| foreign | Read Data Stored by Minitab, S, SAS, SPSS, Stata, Systat, Weka, dBase, ... |
| lmtest | Testing Linear Regression Models |
| stats | The R Stats Package |
| sandwich | Robust Covariance Matrix Estimators |
| car | Companion to Applied Regression |
| xts | eXtensible Time Series |
| portes | Portmanteau test for Univariate and Multivariate Time Series |
| urca | Unit root and cointegration tests for time series data, |
| fGarch | Rmetrics - Autoregression Conditional Heteroskedastic Modelling |
| vars | VAR Modelling |

[4] Available R-packages: http://cran.r-project.org/web/packages/available_packages_by_name.html, 08.05.2015

## 3.1.Data

The temperature data were acquired from the U.K. Met Office's web site, where we got the daily HadCET, from 1772 to February 2015, for mean (of min and max) temperature (Parker, et al., 1992). We decided to use Central England Temperature (CET) in our analysis, as this weather record provides a trustworthy estimate for general climate in the UK. A notion provided by (Subak, et al., 2000) proposes that CET captures a clear representation of the climate in the UK, and that individual station records are affected, or even contaminated, by local environmental conditions. Since we are investigating the effect in the UK market as a whole, we want to exclude the noise affiliated with local weather observations.

Storage level and natural gas demand data were provided by National Grid (National Grid, 2015). Spot prices, 1M, 2M, and 3M futures contracts for Natural Gas were obtained from ICIS Heren (ICIS Heren, 2015). Supplementary spot prices were sent to us by mail from Nick Grogan at Energy Solutions (Grogan, 2015). S&P500 data were gathered from www.finance.yahoo.com (Yahoo! Finance, 2015). TBills are provided by Federal Reserve Economic Data (Federal Reserve Bank of St. Louis, 2015), and are the 3-month treasury bills. Brent Oil data were obtained in USD from www.quandl.com (US Department of Energy, 2015). Currency exchange of USD to GBP there obtained from http://www.ozforex.com (OzForex, 2011).

### 3.1.1. Organizing the data

In structuring and organizing the final dataset, we choose to omit days where market data for 1-month futures price where not recorded such as weekends, holidays or other missing dates. We believe that this is the best approach, because these days may contain anomalies resulting in distorted relationships when computing our model.

## 3.2. Spot price, 1–3M Futures price on Natural Gas

Spot and M1 - M3 futures data are gathered from ICIS Heren, but we also got spot prices from Energy Solutions, to check for discrepancies on spot prices from different sources.

As Mu (2007) argues in his paper, the spot prices are not a good basis for these calculations because individual firms report prices, and data on spot prices are not readily available. These firms have no obligation to make sure they are giving correct or reliable information, and these spot prices may include discounts or premiums, resulting in discrepancies. Limited availability for spot prices is a typical problem in commodity price studies (Energy Information Administration, 2012). The literature suggests that the first nearby futures or futures prices is used as a proxy for the spot price. The futures prices are reported at stock exchanges, and are more reliable to reflect the real price process of natural gas. *RET1* and *RET2* are estimated in the same manner as Mu (2007), applying equation (20) and (21).

$$RET1 = ln(\frac{M1_t}{M1_{t-1}}) \tag{20}$$

$$RET2 = ln(\frac{M1_t}{M1_{t-2}}) \tag{21}$$

When creating the *RET1* series, we substituted the *RET1* corresponding to the first day of each month, with the second nearest observation, to account for the rollover of the contracts. This occurs when the market participants renew their contracts from the previous month to the coming month. “Traders are often forced to cover their positions at the last trading day of a contract's life such that trading volume and open interest decline, while price volatility increases substantially” (Mu, 2007) pp 50.

## 3.3. Weather data

For measuring the weather shocks we chose to use the same approach as (Mu, 2007), because this is the base used for weather derivatives. We use daily weather data (*DD*) that are composed of heating degree-days (*HDD*) and cooling degree-days (*CDD*):

$$\mathrm{CDD_t} = \mathrm{Max}(0\,, \mathrm{DailyTemp_t} - \mathrm{X°C}) \tag{22}$$

$$HDD_t = Max(0\,, X°\mathrm{C} - DailyTemp_t) \tag{23}$$

$$DD_t = CDD_t + HDD_t \tag{24}$$

*DailyTemp* are the temperature on a given day in our period, *X*°C are a base temperature reflecting the temperature commonly used in weather derivatives[5], which according to CME Group are set to 18℃ . As the temperature decreases (increases), and moves away from 18℃, the variable measures the need for heating (cooling), indicating an increased demand for natural gas.

$$\mathrm{W_tF} = \frac{1}{\mathrm{m}}\sum_{\mathrm{i}=1}^{\mathrm{m}}(\mathrm{DD_{t+i}} - \mathrm{DDnorm_{t+i}}) \qquad (25)$$

$W_tF$ are the weather shock variable we will focus upon in the subsequent analysis and consists of; *m* days of *DD* deviation ahead from the current day *t* and are our forecast horizon, set to 7 days, since this is the amount of days that are normally used for fairly accurate weather forecasts. $DD_{t+i}$ are the degree days used in the forecast period, while $DDnorm_{t+i}$ are the average temperature at day *t+1*, based upon daily data from January 1985 to January 2015.

When we remove the average temperature on each day during the forecast period, we design the variable *WtF* as a measurement of weather anomalies for the given day of the forecast.

## 3.4. Storage data

National Grid provided the storage data representing actual storage level and available storage capacity. Based upon these two, we calculated maximum storage capacity. This enabled us to estimate reservoir levels of total storage capacity at any given day in the dataset.

We then estimated the storage shock parameter[6] ($StDev_t$) as the deviation from an estimated average level of storage ($\overline{Storage_t}$). The latter estimate is based upon the sample average level of storage for each day, represented by observation *t*.

We used the expected storage level to find the actual change by taking the daily given amount and subtracting the expected amount of the corresponding date to obtain the size of the daily storage deviation. The *StDev* variable are calculated using equation (26):

[5] http://www.cmegroup.com/trading/weather/temperature-based-indexes.html#3, (11.05.2015)

[6] The sample size of storage data is limited to contain 5 years of daily observations.

$$\mathrm{StDev_t} = (\mathrm{Storage_t} - \overline{\mathrm{Storage_t}}) \tag{26}$$

Where *Storage*$_t$ is the level of natural gas storage at time *t*, and $\overline{Storage_t}$ is the average level of natural gas in storage at time *t*.

The storage data collected included an upward sloping trend, and we were interested checking if this trend affected the estimation, so we decided to detrend the storage data. This was done by identifying the slope parameter of the trendline, then subtracting the slope coefficient from each observation to get rid of the upward sloping trend.

These data, without trends, were used to estimate a new set of mean daily storage level. This new series were applied to make a detrended storage deviation function (*dStDev*).

$$\mathrm{dStDev_t} = (\mathrm{detrendStorage_t} - \overline{\mathrm{detrendStorage_t}}) \tag{27}$$

## 3.5. Additional variables

The variables Brent oil (*BOil*), T-bills (*TBill*), Demand (*Dem*) and S&P (*SP*) are included as exogenous variables to increase the fit of our economic models. Another argument for including these variables are that we intend to compare the results obtained in this analysis, to the findings of Mu (2007), which suggested a similar model.

Brent oil were acquired in USD. We multiplied every observation of the Brent oil with its currency exchange rate, USD to GDP, at the corresponding time. Through this transformation we retrieved the correct scaling of the variable.

To conduct the analysis we had to create stationary variables, which enable us to perform hypothesis testing. This transformation was done by applying equation (28), where $\Delta S_t$ are the new daily variable that are based on natural logarithm of today's value, $S_t$, divided by yesterdays value, $S_{t-1}$, to each individual variable, rendering the result of these timeseries data stationary.

$$\Delta S_t = ln\left(\frac{S_t}{S_{t-1}}\right) \tag{28}$$

## 3.6. Descriptive Statistic

*Table 3-2: Descriptive Statistics*

| | *RET1* | *RET2* | *dStDev* | *WtF* | *retBOil* | *retSP* | *retTBill* | *LnDem* |
|---|---|---|---|---|---|---|---|---|
| Mean | 0.0000642 | 0.000492 | -0.00304 | 0.21451 | -0.00026 | 0.000472 | -6.5E-05 | -0.00018 |
| Standard Error | 0.0005990 | 0.000848 | 0.005112 | 0.059342 | 0.000429 | 0.000283 | 9.51E-05 | 0.000431 |
| Median | -0.0009304 | -0.00091 | 0.015202 | 0.266144 | -0.00052 | 0.000686 | 0 | -0.00061 |
| St. Deviation | 0.0212637 | 0.030116 | 0.181451 | 2.106416 | 0.015226 | 0.010053 | 0.003375 | 0.015294 |
| Sample Variance | 0.0004521 | 0.000907 | 0.032924 | 4.43699 | 0.000232 | 0.000101 | 1.14E-05 | 0.000234 |
| Kurtosis | 9.2718313 | 5.391251 | 0.671222 | 1.257227 | 1.50255 | 4.815435 | 28.05092 | 17.82297 |
| Skewness | 1.0920790 | 0.781266 | -0.74811 | 0.586342 | -0.12804 | -0.47461 | -1.51866 | -1.18604 |
| Minimum | -0.1370371 | -0.16955 | -0.57654 | -4.74905 | -0.08394 | -0.06896 | -0.04032 | -0.18116 |
| Maximum | 0.1548765 | 0.186339 | 0.358031 | 8.982381 | 0.055847 | 0.046317 | 0.031643 | 0.070469 |
| Sum | 0.0808794 | 0.620492 | -3.83347 | 270.2824 | -0.32643 | 0.594201 | -0.08148 | -0.22827 |
| Largest(1) | 0.1548765 | 0.186339 | 0.358031 | 8.982381 | 0.055847 | 0.046317 | 0.031643 | 0.070469 |
| Smallest(1) | -0.1370371 | -0.16955 | -0.57654 | -4.74905 | -0.08394 | -0.06896 | -0.04032 | -0.18116 |
| Conf.Level (95,0%) | 0.0011752 | 0.001664 | 0.010029 | 0.116419 | 0.000842 | 0.000556 | 0.000187 | 0.000845 |

*Variable description available in chapter 3. RET1 and RET2 are in chapter 3.2, dStDev are in 3.3, WtF are in 3.4, and chapter 3.5 explains the remaining variables.*

The *RET1* series have a high kurtosis, which means that more of the variance is a result of infrequent extreme deviations, also referred to as the volatility of the volatility, suggesting that we need to implement a GARCH model.

The skewness of *RET1* and *RET2* are both positive, indicating that they have a longer right side distribution tails. From a standpoint where zero represent perfectly symmetrical data, the values of the skewness signify that *RET1* is highly skewed, while *RET2* is only moderately skewed (Bulmer, 1979). This skewness seem to be a result of the asymmetric properties of the demand variable.

*Table 3-3: Covariance matrix of residuals*

| | *RET1* | *WtF* | *dStDev* |
|---|---|---|---|
| *RET1* | 0.00044 | 0.00144 | -0.00001 |
| *WtF* | 0.00144 | 0.31547 | 0.00044 |
| *dStDev* | -0.00001 | 0.00044 | 0.00022 |

*Table 3-4: Correlation matrix of residuals*

| | *RET1* | *WtF* | *dStDev* |
|---|---|---|---|
| *RET1* | 1 | 0.12187 | -0.01966 |
| *WtF* | 0.12187 | 1 | 0.05297 |
| *dStDev* | -0.01966 | 0.05297 | 1 |

# 4. Analysis

In this chapter, we implement the models introduced in chapter 2. Part 4.1 contains analysis based on the tests introduced in chapter 2.1, and 4.2 consists of the analysis of the models introduced in chapter 2.2.

We started by estimating a preliminary model, to inspect the relationship between the return series and the other variables introduced in chapter 3.1. The results from the estimated regression model are in Table 4-1.

*Table 4-1: Initial regression model*

| | Estimate | Std. Error | t-value | p-value | |
|---|---|---|---|---|---|
| ***RET1*** | | | | | |
| *(Intercept)* | 0.01393 | 0.00834 | 1.67000 | 0.09510 | . |
| *WtF* | 0.00016 | 0.00030 | 0.52500 | 0.59960 | |
| *dStDev* | -0.00277 | 0.00432 | -0.64200 | 0.52090 | |
| *BOil* | 0.00001 | 0.00007 | 0.12300 | 0.90190 | |
| *TBill* | 0.09734 | 0.55410 | 0.17600 | 0.86060 | |
| *SP* | 0.00000 | 0.00000 | -1.67800 | 0.09360 | . |
| *Demand* | -0.00003 | 0.00001 | -2.18700 | 0.02890 | * |
| | | | | | |
| ***RET2*** | | | | | |
| *(Intercept)* | 0.02528 | 0.01178 | 2.14500 | 0.03212 | * |
| *WtF* | 0.00059 | 0.00042 | 1.41000 | 0.15874 | |
| *dStDev* | -0.00404 | 0.00610 | -0.66200 | 0.50791 | |
| *BOil* | -0.00003 | 0.00010 | -0.27100 | 0.78655 | |
| *TBill* | 0.27370 | 0.78300 | 0.35000 | 0.72672 | |
| *SP* | -0.00001 | 0.00000 | -1.87700 | 0.06073 | . |
| *Demand* | -0.00005 | 0.00002 | -2.66200 | 0.00787 | ** |

Regression of *RET1* with all variables and 1260 observations. Multiple R-squared: 0.005913, Adjusted R-squared: 0.001153, F-statistic: 1.242 on 6 and 1253 DF, p-value: 0.2818.
Regression of *RET2* with all variables and 1253 observations. Multiple R-squared: 0.01037, Adjusted R-squared: 0.005631, F-statistic: 2.188 on 6 and 1253 DF, p-value: 0.04174.
'***' '**' '*' '.' Denote significance at the 0.1%, 1%, 5% and 10% level.

The model in Table 4-1 fits the data poorly. The adjusted $R^2$ are quite small for both regressions, and we suspect that the model contains flaws. More specifically, we expect the data to be non-stationary. In the next chapter we will start out by testing whether the variables are stationary or not, by implementing an Augmented Dickey-Fuller (ADF) test.

## 4.1. Data assessment

We are now going to conduct tests presented in 2.1 to ensure that the time series data do not contains flaws. Through these tests, we obtain stationary data, that are unaffected by autocorrelation.

### 4.1.1. Augmented Dickey-Fuller test

We chose to run an ADF test on all our variables to check if we have stationary data. The output from the test reviled the following information about the variables:

*Table 4-2: ADF test*

| Variable | Test-statistic | |
|---|---|---|
| *M1* | 0.0529 | |
| *BOil* | -0.4899 | |
| *TBill* | -0.3118 | |
| *SP* | 2.1784 | * |
| *Demand* | -0.8006 | |
| | | |
| *RET1* | -13.6521 | ** |
| *RET2* | -12.1303 | ** |
| *WtF* | -7.8364 | ** |
| *dStDev* | -2.2059 | * |
| *retBOil* | -12.2692 | ** |
| *retT.Bill* | -6.5056 | ** |
| *retSP* | -14.388 | ** |
| *lnDem* | -8.9478 | ** |

The test is conducted with 1246 degrees of freedom and 6 lags. Critical values are '2.58' '1.95' '1.62', for test at the 1%, 5% and 10% level. '**' '*' '.' Denote significance at the 1%, 5% and 10% level.

1-month futures natural gas price (*M1*), Brent Oil (*BOil*), Treasury bills (*TBill*), and *Demand* provided evidence supporting our initial thought, which was that these variables were non-stationary, and the test concluded that the prices contain a unit root. As we can see from Table 4-2, these variables are insignificant.

We transformed the variables by first differencing the natural logarithm of the variables, and created the daily percentage change for the futures price, Brent oil, treasury bills, S&P 500, and demand. Even though *SP* did not require a transformation, we decided to do so, which resulted in higher significance level. Based on this, we chose to include the transformed variable.

These transformations create stationary variables in our time series analysis. This correction resulted in a new regression model that gave a better linear relationship.

*Table 4-3: Regression model*

| | Estimate | Std. Error | t-value | p-value | |
|---|---|---|---|---|---|
| <u>*RET1*</u> | | | | | |
| *(Intercept)* | 0.00014 | 0.00060 | 0.23500 | 0.81456 | |
| *WtF* | 0.00022 | 0.00029 | 0.77700 | 0.43715 | |
| *dStDev* | -0.00083 | 0.00333 | -0.24800 | 0.80384 | |
| *retBOil* | 0.15176 | 0.04183 | 3.62800 | 0.00030 | *** |
| *retTBill* | 0.15853 | 0.17759 | 0.89300 | 0.37222 | |
| *retSP* | -0.18769 | 0.06328 | -2.96600 | 0.00307 | ** |
| *LnDem* | -0.06073 | 0.03902 | -1.55600 | 0.11988 | |
| | | | | | |
| <u>*RET2*</u> | | | | | |
| *(Intercept)* | 0.00042 | 0.00085 | 0.48900 | 0.62500 | |
| *WtF* | 0.00079 | 0.00041 | 1.92900 | 0.05400 | . |
| *dStDev* | -0.00032 | 0.00474 | -0.06700 | 0.94600 | |
| *retBOil* | 0.05869 | 0.05954 | 0.98600 | 0.32400 | |
| *retTBill* | 0.34009 | 0.25278 | 1.34500 | 0.17900 | |
| *retSP* | -0.12891 | 0.09007 | -1.43100 | 0.15300 | |
| *LnDem* | -0.01062 | 0.05554 | -0.19100 | 0.84800 | |

Regression of *RET1* with all variables and 1260 observations. Multiple R-squared: 0.01599, Adjusted R-squared: 0.01128. F-statistic: 3.393 on 6 and 1253 DF, p-value: 0.002524.
Regression of *RET2* with all variables and 1260 observations. Multiple R-squared: 0.00614, Adjusted R-squared: 0.001381. F-statistic: 1.29 on 6 and 1253 DF, p-value: 0.2586.
'***' '**' '*' '.' Denote significance at the 0.1%, 1%, 5% and 10% level.

Even though the adjusted $R^2$ for *RET1* model is quite small, at 1.128%, we have obtained an improvement as it has increased from 0.1153% in Table 4-1Table 4-1: Initial regression model. Two of the variables are significant. It is also noteworthy that the two variables in focus, *WtF* and *dStDev*, are insignificant, with high p-values, representing low impact on the returns.

The *RET2* regression model still have a bad fit, with an adjusted $R^2$ of 0.1381%, which declined from Table 4-1, where it was 0.5631% in. The model contains no significant variables.

### 4.1.2. Breusch-Pagan (1979) test for heteroskedasticity

To check whether the data series suffered from heteroskedasticity, we ran a BP-test on each regression model, which provided strong evidence for the presence of heteroskedasticity for the *RET2* regression. The p-value is less than 0.05, thus we reject the null hypothesis and conclude that *RET2* are heteroskedastic.

*Table 4-4: Breusch-Pagan test*

| | BP stat | p-value |
|---|---|---|
| *RET1* | 12.2723 | 0.05616 |
| *RET2* | 14.0468 | 0.02912 |

We ran the Breusch-Pagan (BP) test on the two regressions described in Table 4-3: Regression model. The test is estimated with 6 degrees of freedom.

The *RET1* regression have a p-value just above 0.05, implying that we fail to reject the null hypothesis with 95% certainty, rendering the *RET1* regression model homoskedastic.

### 4.1.3. White Correction

To adjust the *RET2* model for heteroskedasticity, we needed to estimate robust standard errors. The White Correction is a commonly used approach for this purpose, and the corrected standard errors are listed in Table 4-5, with their corresponding estimated coefficients and significance levels. The *RET1* model is homoskedastic, but as it were close to the critical value we ran the white correction on it to see if it provided any significant changes.

*Table 4-5: White-test RET1 and RET2*

| | Estimate | Std. Error | t-value | p-value | |
|---|---|---|---|---|---|
| *RET1* | | | | | |
| *(Intercept)* | 0.00014 | 0.00060 | 0.23270 | 0.81607 | |
| *WtF* | 0.00022 | 0.00029 | 0.76000 | 0.44742 | |
| *dStDev* | -0.00083 | 0.00271 | -0.30480 | 0.76057 | |
| *retBOil* | 0.15176 | 0.03908 | 3.88330 | 0.00011 | *** |
| *retTBill* | 0.15853 | 0.17294 | 0.91670 | 0.35949 | |
| *retSP* | -0.18769 | 0.08344 | -2.24930 | 0.02467 | * |
| *LnDem* | -0.06073 | 0.05432 | -1.11800 | 0.26377 | |
| | | | | | |
| *RET2* | | | | | |
| *(Intercept)* | 0.00042 | 0.00085 | 0.49040 | 0.62397 | |
| *WtF* | 0.00079 | 0.00040 | 1.95630 | 0.05065 | . |
| *dStDev* | -0.00032 | 0.00387 | -0.08220 | 0.93448 | |
| *retBOil* | 0.05869 | 0.06194 | 0.94760 | 0.34350 | |
| *retTBill* | 0.34009 | 0.27521 | 1.23570 | 0.21679 | |
| *retSP* | -0.12891 | 0.11550 | -1.11610 | 0.26460 | |
| *LnDem* | -0.01062 | 0.07065 | -0.15030 | 0.88058 | |

The standard errors are the robust standard errors produced from the White test.
'***' '**' '*' '.' Denote significance at the 0.1%, 1%, 5% and 10% level.

When comparing the robust standard errors from Table 4-5, with the errors obtained from Table 4-3, we can see that the standard errors are slightly changed. The most notable change is that the variable representing the returns obtained from S&P 500, in the *RET1* regression, dropped form the 1% to the 5% significance level.

### 4.1.4. Breusch-Godfrey test for higher order serial correlation

We also decided to control the data for serial correlation by using another formula developed by Breusch and Godrey *(Godfrey, 1978)*. The test is applied to discover potential serial correlation in a regression. Selected output from the BG-test is in Table 4-6.

*Table 4-6: Breusch-Godfrey test*

| | BG stat | p-value |
|---|---|---|
| *RET1* | 9.6787 | 0.4691 |
| *RET2* | 555.2368 | 2.2e-16 |

The test is applied to the regression model in Table 4-3 with a $\chi^2$ distribution, and 10 lags.

The p-value of the BG-test for the *RET1* regression model is a high 0.4691 and supports the $H_0$ stating that there are no autocorrelation. This result supports the conclusion from the ADF tests in Table 4-2 for each variable. The *RET2* regression is, on the other hand, still exhibiting evidence that disproves $H_0$, which indicate that it still suffers from serial correlation.

### 4.1.5. Asymmetry

We now proceed to test the variables for asymmetric properties. To do this we added a binary variable, or a dummy variable, to see if there are different effects for positive changes as opposed to negative changes for each of the included variables. The dummy variables equals to one if the change is equal to or greater than zero, and 0 otherwise. This implies that the base group are a negative change, while the dummy variable takes the value of one when there have been reported an increase in the historic values of the variable.

In addition to this, we checked the data for seasonality, setting the shoulder months as base group. We found evidence of a significant winter effect, in both regression models, but the dummy variable representing the summer were insignificant. We choose to omit the other insignificant dummy variables[7]. The asymmetric winter effect decreases the *RET1* by approximately 0.3% and the *RET2* by 0.64%.

[7] All calculations regarding dummy variables are included in appendix 2 and 3

*Table 4-7: Regression including dummy variables*

| | Estimate | Std.Error | t-value | p-value | |
|---|---|---|---|---|---|
| <u>*RET1*</u> | | | | | |
| *(Intercept)* | 0.00171 | 0.00145 | 1.17700 | 0.23943 | |
| *dStDev* | 0.00720 | 0.00513 | 1.40400 | 0.16045 | |
| *WtF* | 0.00034 | 0.00029 | 1.17600 | 0.23968 | |
| *retBOil* | 0.15397 | 0.04164 | 3.69700 | 0.00023 | *** |
| *retTBill* | 0.13720 | 0.17950 | 0.76400 | 0.44480 | |
| *retSP* | -0.18654 | 0.06304 | -2.95900 | 0.00314 | ** |
| *LnDem* | -0.13275 | 0.04943 | -2.68600 | 0.00733 | ** |
| *dummydStDev* | -0.00411 | 0.00187 | -2.20000 | 0.02799 | * |
| *dummylnDem* | 0.00347 | 0.00155 | 2.24200 | 0.02512 | * |
| *dummyWinter* | -0.00338 | 0.00141 | -2.40100 | 0.01648 | * |
| | | | | | |
| <u>*RET2*</u> | | | | | |
| *(Intercept)* | 0.00741 | 0.00252 | 2.94600 | 0.00328 | ** |
| *dStDev* | 0.01445 | 0.00726 | 1.98900 | 0.04694 | * |
| *WtF* | 0.00226 | 0.00063 | 3.56800 | 0.00037 | *** |
| *retBOil* | 0.05058 | 0.05917 | 0.85500 | 0.39287 | |
| *retTBill* | 0.26812 | 0.25460 | 1.05300 | 0.29250 | |
| *retSP* | -0.12116 | 0.08933 | -1.35600 | 0.17525 | |
| *LnDem* | -0.10788 | 0.07017 | -1.53700 | 0.12445 | |
| *dummydStDev* | -0.00718 | 0.00265 | -2.71500 | 0.00672 | ** |
| *dummyWtF* | -0.00704 | 0.00267 | -2.63500 | 0.00851 | ** |
| *dummylnDem* | 0.00478 | 0.00220 | 2.17600 | 0.02977 | * |
| *dummyWinter* | -0.00639 | 0.00201 | -3.18600 | 0.00148 | ** |

Regression of *RET1* with all variables and dummies, and 1260 observations. Multiple R-squared: 0.02735 Adjusted R-squared: 0.02035. F-statistic: 3.906 on 9 and 1250 DF, p-value: 0.0000656.
Regression of *RET2* with all variables and dummies, and 1260 observations. Multiple R-squared: 0.02741, Adjusted R-squared: 0.01963. F-statistic: 3.52 on 10 and 1249 DF, p-value: 0.0001337.
'***' '**' '*' '.' Denote significance at the 0.1%, 1%, 5% and 10% level.

The three significant dummy variables in the *RET1* regression are the storage level (*dummydStDev*), the demand (*dummylnDem*), and the winter effect (*dummyWinter*). These are significant at the 5% level. The *RET2* regression includes three variables that are significant at the 1% level, which is the storage level (*dummydStDev*), the weather forecast (*dummyWtF*), and the winter effect (*dummyWinter*). The dummy variable for a positive change in the demand (*dummylnDem*) is significant at the 5% level.

From Table 4-7 we can see that the dummy variables are affecting the two return series (*RET1* and *RET2*) in the same direction, the coefficients have the same statistical implications for both return regressions. The dummy coefficients related to storage, winter, and weather are negative, meaning that a decrease in any of these three variables will result in a greater reduction in the returns, than an equivalent increase would affect the returns. The demand have an opposite effect on returns, where an increase in demand results in a greater increase in the returns, than what the equivalent demand reduction would decrease the returns.

The descriptive statistics, in Table 3-2 suggests that the data have a positive skewness of 1.09, which is in accordance with these findings.

### 4.1.6. Non-linear relationship

We also checked for non-linear relations in our data by creating the squared values for each variable and running a new regression with the squared variables along with the improved original variables. This test suggests that the demand have a non-linear effect on *RET1*, where the variable was significant at the 1% level. There were two significant squared variables in the *RET2* regression, *sqSP* and *sqLnDem*, where both of them are significant at the 5% level[8].

To avoid obtaining results that are in conflict with the assumption of no perfect multicollinearity[9], we had to remove the original values.

*Table 4-8: Final regression model*

| | Estimate | Std.Error | t-value | p-value | |
|---|---|---|---|---|---|
| <u>*RET1*</u> | | | | | |
| *(Intercept)* | 0.00278 | 0.00126 | 2.20100 | 0.02789 | * |
| *dStDev* | 0.00765 | 0.00509 | 1.50300 | 0.13303 | |
| *WtF* | 0.00032 | 0.00029 | 1.12100 | 0.26236 | |
| *retBOil* | 0.14905 | 0.04159 | 3.58400 | 0.00035 | *** |
| *retTBill* | 0.14718 | 0.17938 | 0.82000 | 0.41211 | |
| *retSP* | -0.18510 | 0.06296 | -2.94000 | 0.00334 | ** |
| *sqLnDem* | 1.67712 | 0.57047 | 2.94000 | 0.00334 | ** |
| *dummydStDev* | -0.00424 | 0.00186 | -2.28400 | 0.02256 | * |
| *dummyWinter* | -0.00291 | 0.00140 | -2.08200 | 0.03755 | * |
| | | | | | |
| <u>*RET2*</u> | | | | | |
| *(Intercept)* | 0.0088898 | 0.0022703 | 3.916 | 9.5e-05 | *** |
| *dStDev* | 0.0151841 | 0.0072163 | 2.104 | 0.035566 | * |
| *WtF* | 0.0023994 | 0.0006338 | 3.786 | 0.000161 | *** |
| *retBOil* | 0.0256009 | 0.0554494 | 0.462 | 0.644378 | |
| *retTBill* | 0.2039789 | 0.2563727 | 0.796 | 0.426396 | |
| *sqSP* | 6.2443760 | 3.2751234 | 1.907 | 0.056800 | . |
| *sqLnDem* | 1.8020509 | 0.8079516 | 2.230 | 0.025898 | * |
| *dummydStDev* | -0.0077058 | 0.0026278 | -2.932 | 0.003424 | ** |
| *dummyWtF* | -0.0079288 | 0.0026658 | -2.974 | 0.002993 | ** |
| *dummyWinter* | -0.0055247 | 0.0019989 | -2.764 | 0.005796 | ** |

Regression of *RET1* with improved variables and 1260 observations. Multiple R-squared: 0.02802, Adjusted R-squared: 0.02181. F-statistic: 4.508 on 8 and 1251 DF, p-value: 0.00002041.
Regression of *RET2* with improved variables and 1260 observations. Multiple R-squared: 0.02902, Adjusted R-squared: 0.02203. F-statistic: 4.151 on 9 and 1250 DF, p-value: 2.727e-5.
'***' '**' '*' '.' Denote significance at the 0.1%, 1%, 5% and 10% level.

[8] Insignificant squared variables were removed, but the main estimations are provided in appendix 4 and 5
[9] Multicollinearity is defined as a perfect correlation between explanatory variables (Wooldridge, J)

These two models are slightly improved through these corrections. The adjusted $R^2$ of *RET1* increased from 0.1153% in Table 4-1, to 2.181%, and more variables that are significant. The *RET2* model also provides significant variables, with an adjusted $R^2$ of 2.336%, which is higher than what Table 4-1 provided.

It is notable that the two variables we are focusing on are insignificant in the first regression, which can be interpreted as a failed attempt to capture this relationship, while the second regression provides a significant weather coefficient. We are interested in the short-term relationship between natural gas returns, weather and storage, represented by shocks to these variables, which we are unable to capture in a linear regression model. This part of the analysis is merely preparing the model for further analysis.

## 4.2. Analysis models

### 4.2.1. GARCH (1,1)

We estimate a GARCH(1,1) model to enhance our understanding of the underlying qualities of the return series. The model produces output that reveals the true nature of the series, and provides insight to the underlying behavior of the return series. The GARCH model will also detect whether there are spikes in the price series.

To find the parameters needed to calculate the volatility for the return on natural gas 1-month futures price, we ran a continuous GARCH(1,1) model.

*Table 4-9: GARCH Model*

| | Estimate | Std.Error | t-value | p-value | |
|---|---|---|---|---|---|
| *RET1* | | | | | |
| *mu* | -7.963e-05 | 4.389e-04 | -0.181 | 0.85603 | |
| *omega* | 8.333e-06 | 3.187e-06 | 2.615 | 0.00892 | ** |
| *alpha1* | 1.380e-01 | 3.229e-02 | 4.274 | 1.92e-05 | *** |
| *beta1* | 8.611e-01 | 3.045e-02 | 28.276 | < 2e-16 | *** |
| $V_L$ | 0.01111 | | | | |
| $\gamma$ | 0.00090 | | | | |
| | | | | | |
| *RET2* | | | | | |
| *mu* | -1.167e-03 | 6.434e-04 | -1.813 | 6,98E-02 | . |
| *omega* | 6.204e-05 | 1.505e-05 | 4.123 | 3.74e-05 | *** |
| *alpha1* | 3.893e-01 | 5.380e-02 | 7.237 | 4.60e-13 | *** |
| *beta1* | 6.028e-01 | 4.659e-02 | 12.939 | < 2e-16 | *** |
| $V_L$ | 0.007853 | | | | |
| $\gamma$ | 0.0079 | | | | |

The Model include 1260 observations, the standard errors are based on Hessian. The log likelihood are 3203.071 and 2843.241 for *RET1* and *RET2* respectively.
'***' '**' '*' '.' Denote significance at the 0.1%, 1%, 5% and 10% level.

The GARCH model estimated on *RET1* have a relatively high alpha value, 0.13800, which is higher than normal for daily data. The usual range for the reaction parameter, alpha, is between 0.05 and 0.1, where the size relates to the stability of the market *(Alexander, 2008)*. Higher alpha values indicate less stability. It also provides insight to how the data series reacts to market shocks. This means that a higher alpha value usually results in more spiky behavior, rather than a model with lower alpha values.

The persistence parameter, beta, usually ranges between 0.85 and 0.98, and so does the beta value of *RET1*, which is 0.86110. The beta value represents the persistence of the volatility after a shock occurs in the market.

The omega coefficient, combined with alpha and beta, represent the speed of which the mean reversion and the long-run volatility present in the data series.

The *RET1* series is quite stable, where the volatility are mostly less than 3%, but can spike up to approximately 6.5%. The long-run volatility is low, 1.11%, and the mean reversion rate is 0.09%. Figure 4-1 illustrate these findings, where the spikes are prominent.

*RET2* has a higher alpha, 0.3893, and a lower beta, 0.6028, indicating that the data series bare evidence of a more unstable market, with volatility spiking up to approximately 14%. The long-run volatility is even lower than for RET1, and is 0.7853%, with a mean reversion rate of 0.79%. Figure 4-2 illustrate this effect, where the spikes occur with the same frequency as for *RET1*, but the magnitude of the spikes are two times as high. This may be a result of the rollover effect, which increases the volatility, and are accounted for in the *RET1* series, but not in the *RET2* series.

*Figure 4-1: Volatility of Natural gas returns (RET1)*

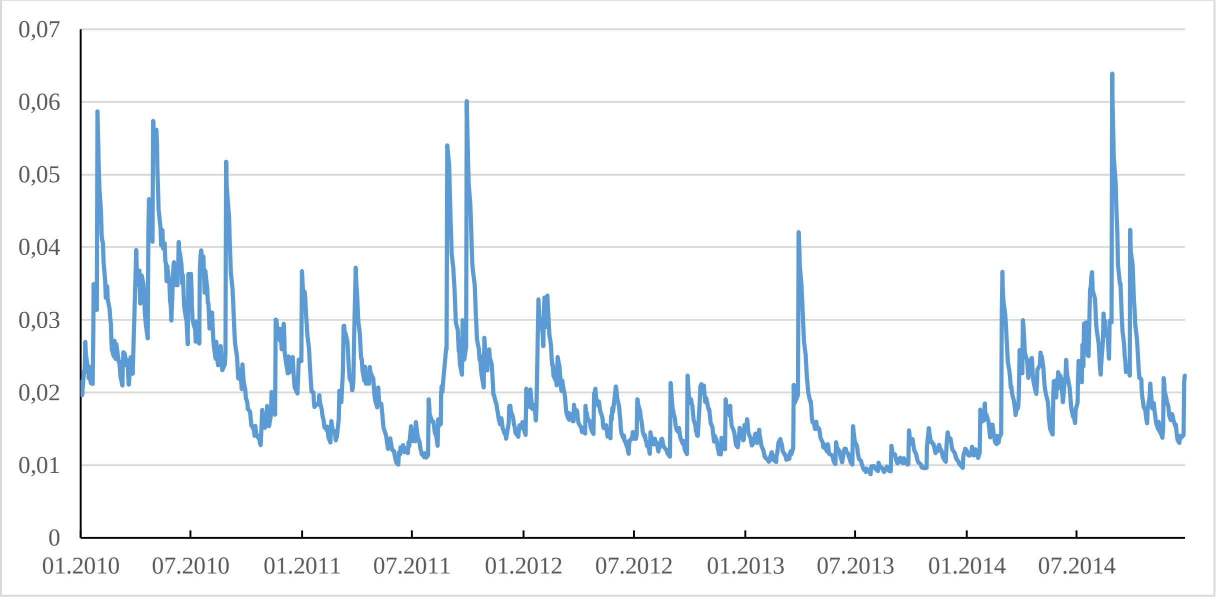


*Figure 4-2: Volatility of Natural gas returns (RET2)*

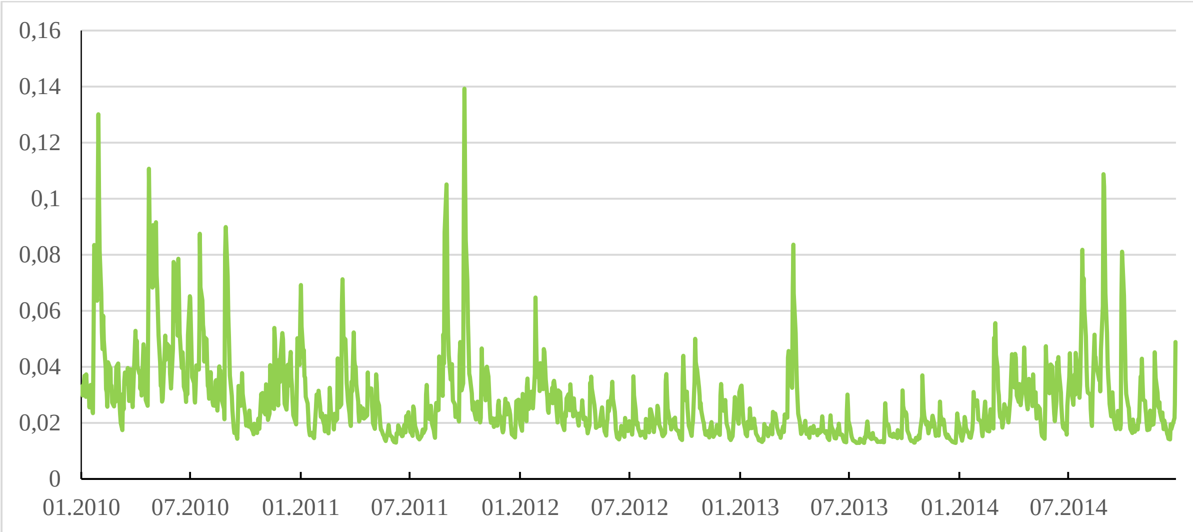



*Table 4-10: Standardized Residual Test GARCH Model*

| | | | Statistic | p-value |
|---|---|---|---|---|
| ***RET1*** | | | | |
| *Jarque-Bera Test* | R | Chi^2 | 1480.973 | 0 |
| *Shapiro-Wilk Test* | R | W | 0.928891 | 0 |
| *Ljung-Box Test* | R | Q(10) | 5.371607 | 0.8650133 |
| *Ljung-Box Test* | R | Q(15) | 11.03362 | 0.7502074 |
| *Ljung-Box Test* | R | Q(20) | 13.77553 | 0.841692 |
| *Ljung-Box Test* | R^2 | Q(10) | 11.25241 | 0.3381942 |
| *Ljung-Box Test* | R^2 | Q(15) | 15.34717 | 0.4267109 |
| *Ljung-Box Test* | R^2 | Q(20) | 22.06268 | 0.337119 |
| *LM Arch Test* | R | TR^2 | 13.4375 | 0.338059 |
| | | | | |
| ***RET2*** | | | | |
| *Jarque-Bera Test* | R | Chi^2 | 371.2895 | 0 |
| *Shapiro-Wilk Test* | R | W | 0.9720125 | 6.698163e-15 |
| *Ljung-Box Test* | R | Q(10) | 236.0304 | 0 |
| *Ljung-Box Test* | R | Q(15) | 240.9694 | 0 |
| *Ljung-Box Test* | R | Q(20) | 246.5269 | 0 |
| *Ljung-Box Test* | R^2 | Q(10) | 24.42135 | 0.006557204 |
| *Ljung-Box Test* | R^2 | Q(15) | 26.32866 | 0.03470679 |
| *Ljung-Box Test* | R^2 | Q(20) | 32.70703 | 0.03632308 |
| *LM Arch Test* | R | TR^2 | 22.04856 | 0.0369792 |

Tests conducted on a chi distribution, checking that the standardized errors are independent.

We also include the standardized residual test from the GARCH model. The Jarque-Bera (1980) test is a goodness-of-fit test of whether the sample data have a kurtosis and skewness of a normal distribution. $H_0$ is a joint hypothesis of both the skewness and the excess kurtosis being equal to zero. As either the excess kurtosis or the skewness differs from zero, the test statistic increases.

Table 4-10 suggest that both return series violate the underlying hypothesis of the JB test.

The Shapiro-Wilk (1965) test is a test for a normally distributed population, which is $H_0$. We reject the null hypothesis if the p-value is less than the selected significance level (5%). Both test's p-values are small, and both series end up rejecting the null.

Next up is the Ljung-Box (Ljung & Box, 1978) test, which is a test to check for independent residuals in the fitted GARCH model. The test assumes that the residuals are independently distributed. As the Ljung-Box test conducted on the *RET1* series with 10 lags have a high p-value, we conclude that the standardized residuals obtained from the *RET1* series are independently distributed. The *RET2* series provide a Ljung-Box test statistic, with 10, 15, and 20 lags, of 0; implying that the standardized residuals are dependent on past values of itself. These results coincide with the conclusion from the BP test, in Table 4-4, that autocorrelation is present in *RET2*.

At last, we have the LM Arch test, introduced by Engle (1982). It is a test for autocorrelation in the squared residuals. $H_0$ is that the model has no form of autocorrelation, and a high critical value result in rejection of $H_0$. This test concludes that we fail to reject the null hypothesis for *RET1*, while *RET2* provide evidence against the null hypothesis.

Since the *RET2* series still suffer from autocorrelation, we have decided to omit this series from the remaining part of the analysis.

To investigate whether the short run volatility spikes, illustrated in Figure 4-1 and Figure 4-2, can be a result of weather or storage shocks, we will use a VAR model to elucidate this dynamic relationship.

### 4.2.2. Vector Autoregression (VAR) model

We then applied the Vector Autoregression (VAR) model to investigate the variables behavior when more than just one variable change over time. The model explains the evolution of the endogenous variables as a function of their own lags and the lagged values of the other variables in the model. We have decided to run a trivrariate VAR model, where we use three endogenous variables, the first one being the returns from natural gas (*RET1*), the second variable is the modeled weather forecast (*WtF*), and lastly the detrended storage variable (*dStDev*). The reason for including weather as an endogenous variable is to be able to see how the two other variables evolve as the weather changes[10]. We know that neither returns, nor storage levels can affect the weather, but we are interested in defining the dynamic effect of these three variables. Holding weather as an exogenous variable will exclude the delayed effect of weather in the analysis, which we are interested in measuring. The exogenous variables are; return on Brent Oil (*retBOil*), return on S&P500 (*retSP*), return on T-Bills (*retTBill*), and the change in demand (*sqlnDem*). We have chosen to exclude the dummy variable from the remaining analysis, both to obtain similarity in modeling compared to Mu (2007) and because they do not alter any of the conclusions.

[10]A structural VAR model might be a better approach to estimate this realtionship, as we whould be able to decide the direction of dynamic interaction. This means that we would have been able to let the weather affect storage and returns, but not the other way around.

Since the amount of lags is unknown, we first estimated the optimal number of lags using a set of information criterions. These information criterions are AIC[11] (Akaike, 1973), HQ[12] (Hanna & Quinn, 1979), SC[13] (Schwarz, 1978), and FPE[14] (Akaike, 1974).

*Table 4-11: Optimal number of lags*

| Lags | AIC(n) | HQ(n) | SC(n) | FPE(n) |
|---|---|---|---|---|
| 1 | -1.675395e+01 | -1.671229e+01 | -1.664313e+01 | 5.294811e-08 |
| 2 | -1.713074e+01 | -1.707519e+01 | **-1.698297e+01** | 3.632562e-08 |
| 3 | -1.715590e+01 | -1.708646e+01 | -1.697119e+01 | 3.542335e-08 |
| 4 | -1.716035e+01 | -1.707703e+01 | -1.693870e+01 | 3.526598e-08 |
| 5 | -1.715305e+01 | -1.705584e+01 | -1.689445e+01 | 3.552456e-08 |
| 6 | -1.721549e+01 | -1.710439e+01 | -1.691995e+01 | 3.337442e-08 |
| 7 | **-1.723093e+01** | **-1.710594e+01** | -1.689845e+01 | **3.286337e-08** |
| 8 | -1.722785e+01 | -1.708898e+01 | -1.685843e+01 | 3.296475e-08 |
| 9 | -1.722124e+01 | -1.706847e+01 | -1.681487e+01 | 3.318383e-08 |
| 10 | -1.722073e+01 | -1.705408e+01 | -1.677742e+01 | 3.320099e-08 |

| AIC(n) | HQ(n) | SC(n) | FPE(n) |
|---|---|---|---|
| 7 | 7 | 2 | 7 |

Table 4-11 provides four different selection criterions, where three of them, AIC, HQ, and FPE, suggests the use of p=7 lags, while SC suggests the use of p=2 lags. The weather variable inflates the optimal amount of lags suggested by the three first selection criterions.

When we ran the VAR model with p=7 lags, and p=2 lags we discovered $R^2$ were displaying slight variations and F-statistics increased when employing 2 lags. We decided to proceed with two lags as this did not omit any significant variables, gave a better F-statistic, and made it easier to interpret the VAR tables[15]. The output from the weather regression of the VAR model in Appendix 8 and appendix 9 illustrate that these two models produce similar results. These weather regressions are unaffected by any other variable than itself, but bare evidence of serial correlation within the weather variable, as the weather is the only significant variable in these regressions.

[11] Akaike's Information Criterion
[12] Hannan & Quinn
[13] Schwarz Criterion
[14] Akaike's Final Prediction Error
[15] VAR models with 7 lags are provided in appendix 6-8

#### 4.2.2.1. The VAR model output

We start by running the VAR model, and obtain a tridimensional output. The endogenous variables have an additional designation in the output, where the variable name includes "*.l1*" and "*.l2*", representing the first and second lag of the variables.

Table 4-12 displays *RET1* as a function of each included variable, both exogenous and endogenous. The difference between these variables is how we allowed them to behave within the model. The exogenous variables are treated as constants, and the endogenous variables can have a form of integrated relationship, where historic values are taken into account.

*Table 4-12: VAR Model RET1*

| | Estimate | Std. Error | t-value | p-value | |
|---|---|---|---|---|---|
| *RET1.l1* | 2.960e-02 | 2.825e-02 | 1.048 | 0.294898 | |
| *RET1.l2* | -2.887e-02 | 2.813e-02 | -1.027 | 0.304817 | |
| *dStDev.l1* | 3.133e-02 | 3.768e-02 | 0.831 | 0.405953 | |
| *dStDev.l2* | -3.334e-02 | 3.743e-02 | -0.891 | 0.373297 | |
| *WtF.l1* | -5.734e-04 | 8.902e-04 | -0.644 | 0.519589 | |
| *WtF.l2* | 5.204e-04 | 9.082e-04 | 0.573 | 0.566743 | |
| *const* | 8.112e-04 | 1.286e-03 | 0.631 | 0.528170 | |
| *trend* | -1.548e-06 | 1.769e-06 | -0.875 | 0.381876 | |
| *retBOil* | 1.513e-01 | 4.199e-02 | 3.602 | 0.000328 | *** |
| *retTBill* | 1.368e-01 | 1.792e-01 | 0.763 | 0.445482 | |
| *retSP* | -1.924e-01 | 6.329e-02 | -3.040 | 0.002418 | ** |
| *sqLnDem* | 1.696e+00 | 5.725e-01 | 2.963 | 0.003105 | ** |

VAR estimation with *RET1*, *WtF*, and *dStDev* as endogenous variables, with respect to *RET1*. 1260 observations. Log likelihood: 5456.175, Multiple R-squared: 0.02386, Adjusted R-squared: 0.01525. F-statistic: 2.769 on 11 and 1246 DF, p-value: 0.001472.
'***' '**' '*' '.' Denote significance at the 0.1%, 1%, 5% and 10% level.

When the lags are set to p=2, we find that the dynamic relationship is insignificant. None of the endogenous variables are significant, and the model is able to explain 1.5% of the change in natural gas returns. These finding suggests that the returns are unaffected by the endogenous variables, providing evidence against our hypothesized relationship.

*Table 4-13: VAR model dStDev*

| | Estimate | Std. Error | t-value | p-value | |
|---|---|---|---|---|---|
| *RET1.l1* | -1.941e-02 | 2.021e-02 | -0.961 | 0.336853 | |
| *RET1.l2* | 8.396e-04 | 2.012e-02 | 0.042 | 0.966718 | |
| *dStDev.l1* | 6.916e-01 | 2.695e-02 | 25.657 | < 2e-16 | *** |
| *dStDev.l2* | 2.950e-01 | 2.677e-02 | 11.019 | < 2e-16 | *** |
| *WtF.l1* | 2.819e-03 | 6.367e-04 | 4.428 | 1.03e-05 | *** |
| *WtF.l2* | -6.123e-03 | 6.496e-04 | -9.426 | < 2e-16 | *** |
| *const* | 3.577e-03 | 9.195e-04 | 3.890 | 0.000105 | *** |
| *trend* | -5.189e-06 | 1.266e-06 | -4.100 | 4.40e-05 | *** |
| *retBOil* | 5.512e-02 | 3.004e-02 | 1.835 | 0.066740 | . |
| *retTBill* | -2.928e-01 | 1.282e-01 | -2.284 | 0.022535 | * |
| *retSP* | -2.568e-02 | 4.527e-02 | -0.567 | 0.570633 | |
| *sqLnDem* | 7.003e-01 | 4.095e-01 | 1.710 | 0.087510 | . |

VAR estimation with *RET1*, *WtF*, and *dStDev* as endogenous variables, with respect to *dStDev*. 1260 observations. Log likelihood: 5456.175, Multiple R-squared: 0.9931, Adjusted R-squared: 0.9931. F-statistic: 1.64e+04 on 11 and 1246 DF, p-value: <2.2e-16.
'***' '**' '*' '.' Denote significance at the 0.1%, 1%, 5% and 10% level.

The VAR model is able to confirm that the storage variable is affected by both its own lagged values and the weather variable. These coefficients are all highly significant. The model also have a great fit, as the adjusted $R^2$ is 0.9931, meaning that 99.31% of a change in the storage variable can be explained by this regression model.

It is also worth mentioning that both the constant and the trend coefficient are highly significant, which means that the model has a significant intercept, and that there is a significant linear time trend in the data. The magnitude of the trend is marginal, implying that we were unable to remove the trend completely through the detrending.

Appendix 10 to Appendix 12 contains graphs illustrating the fitted values and the residuals of the endogenous variables. The return series' fitted values are equal to the error plot, due to the low $R^2$ of the *RET1* model. The two other series have a high $R^2$, which result in fitted and residual plots that differentiate from one another.

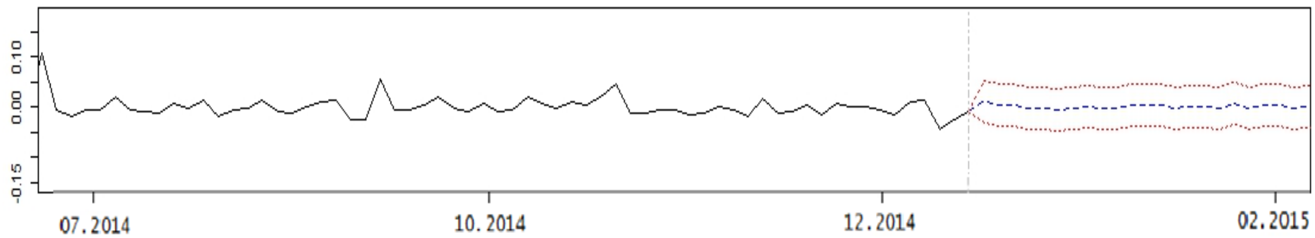
The exogenous variables are the only significant variables in the *RET1* VAR model. We have found evidence of weather affecting the storage, and that the storage variable is affected by lagged values of itself. This means that the VAR model fail to provide significant evidence favoring our hypothesized relationship.

### 4.2.3. Prediction forecast

To begin with, we start by applying the predict formula to predict the future values of *RET1*, *WtF*, and *dStDev* with a 95% confidence interval. We applied the prediction function to the VAR model, introduced in chapter 4.2.2. The exogenous variables imposes restrictions, in the sense that the prediction period must be equivalent to the length of the exogenous time series, which is five years of daily observations.

We decided to narrow down the graph, by restricting the x-axis, to contain the prediction for the first two months following the end of our data. Figure 4-3 to Figure 4-5 contains predictions for each endogenous variable, where the stippled line in the middle is the expected values of the corresponding variable, and the upper and lower dotted lines represents the 95% upper and lower bound of the confidence interval.

*Figure 4-3: Forecast of RET1*




The historical values and the predicted forecast in Figure 4-3 indicates that the volatility is quite low but it can at any time spike up or down. The fluctuations are maintained within a small area of -0.05 to 0.05. The forecast are also consistent with historical returns that mainly display less deviation, with a tendency of mean reversion behavior.

*Figure 4-4: Forecast of WtF*

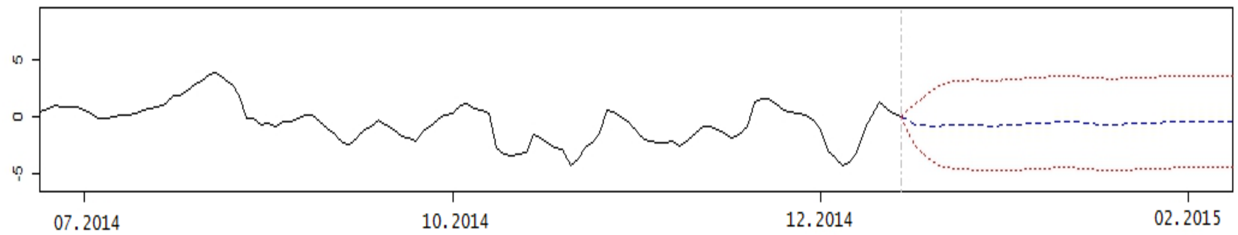



Figure 4-4 displays a relatively smaller spike volatility compared to the *RET1* series, which can be seen as the *WtF* series' confidence interval do not immediately reaches its maximum and minimum limits, but steadily increases until stabilizing with an upper bound and a lower bound of 5 and -5 degrees, respectively. This buildup of the confidence interval represents a less volatile time series, compared to the *RET1* series, which reaches its stabilized confidence interval limits almost immediately.

*Figure 4-5: Forecast of dStDev*

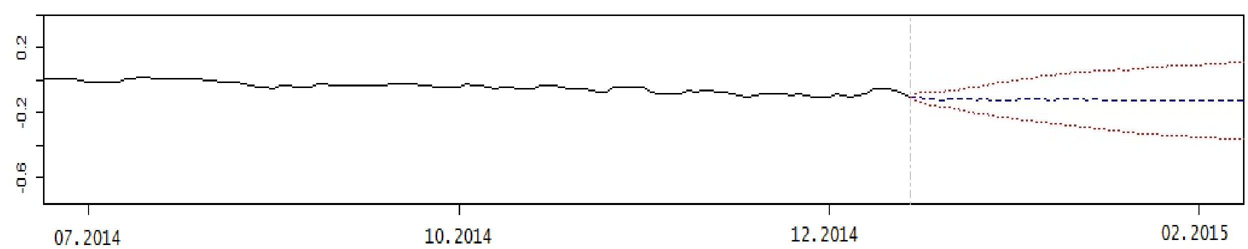



Figure 4-5 contains the endogenous variable with the lowest volatility. This are visualized through the low expansion rate of the confidence interval, which do not seem to reach its maximal limits within the forecast horizon set in this graph. A narrow confidence interval can be interpreted as high prediction precision. This implies low volatility, which corresponds to the observed values of storage level, as it do not provide large fluctuations and are stable.

An additional note is that the storage variable is estimated as the percentage level of capacity, which implies that there is an upper and lower limit imposed to the variable. This imposes further constrictions through injection and withdrawal rate restrictions that limit possible depletion and filling of the storage facilities.

### 4.2.4. Impulse Response Functions

The impulses response model is a method to analyze the response of a variable to a unit shock of another variable. The generalized IRF model is a tool to replicate the effect of a shock in a variable that may be correlated with another endogenous variable, this model are supposed to replicate the repercussions created by such a shock, simulating this effect, and plotting this in a graph. The IRF is based on the VAR model introduced earlier.

The specified relationship between our endogenous variables are represented in:

Table 3-3: Covariance matrix of residuals provide an insight to how our endogenous variables behave, these covariance's may give some greater understanding of how each of them will react when the IRF are applied to the data.

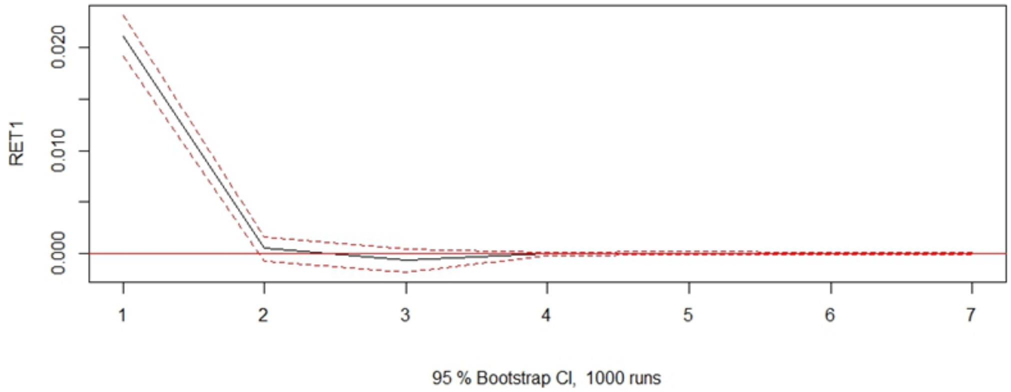
Table 3-4: Correlation matrix of residuals implies a small, but existing correlation between the variables, implying that a generalized IRF (GIRF) model may include the ripple effects of a shock throughout the model, rather than looking at the isolated effect of the shock variable on the response variable.[16]

We chose to use the orthogonalized IRF model, where we isolate the effect of a shock to the response and shock variable, resulting in a less complicated model. We have also decided to shock all endogenous variables with respect to the natural gas returns, and simulate a weather shock with respect to the storage variable, to look for dynamic effects. This resulted in three different shock-simulations with respect to *RET1*, and one with respect to *dStDev*. These estimations are conducted with a 95% confidence interval and the process was simulated 1000 times.

The results from these simulations are provided in Figure 4-6 to Figure 4-9, where the graphs illustrate the ceteris paribus reaction of the response variable as a shock occurs in an endogenous variable. The scaling along the y-axis is the numeric value of a percentage change in the response variable.

[16] When trying to implement this GIRF, we were unable to locate any R packages containing this model, but we were able to detect an undocumented approach to the GIRF. This model are located in the appendix, but the results from this model differed too much from the results from the IRF model. An unknown source proposed the the GIRF model, so we did not trust the results and chose to continue with OIRF.

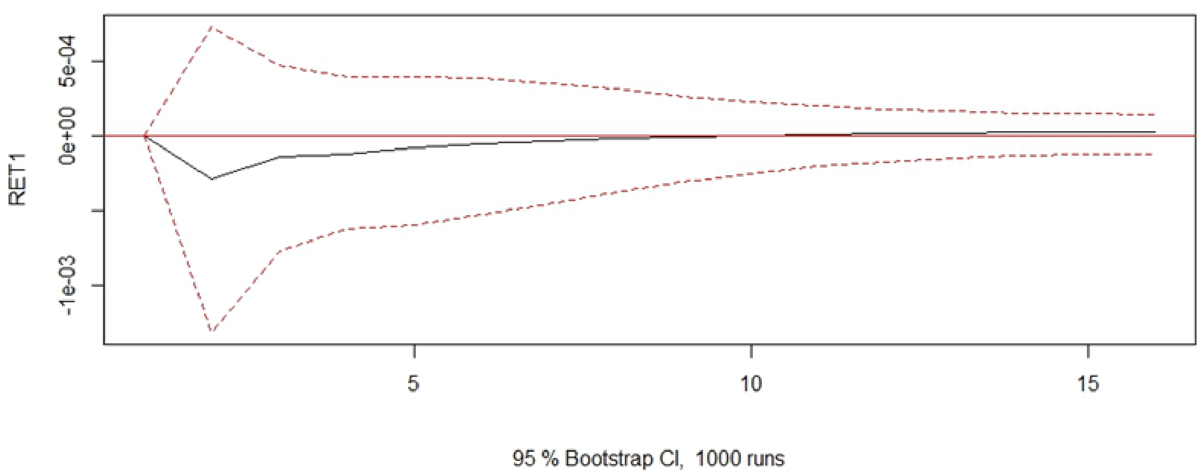
*Figure 4-6: The Orthogonal Impulse Response from RET1*




The graph in Figure 4-6 shows the response of *RET1* illustrate the response of *RET1* as a shock occurs within itself. The immediate effect is that the returns drops back to normal, and the effect of the shock is completely gone after 4 days.

This rapid disappearance of the shock effect, may be caused by market participants holding 1-month futures contracts, which will maximize their own profit, making them sell out their positions when the returns increases by one standard deviation. As the market participants wants to close their positions, there will be an abundance of 1-month futures contracts, forcing the price to drop, canceling out the initial effect of the shock.

*Figure 4-7: The Orthogonal Impulse Response from WtF*




The weather shock, shown in Figure 4-7, involves a longer period of influence before *RET1* stabilizes. When a shock occurs in the weather variable, the effect on return is delayed until the next day. The effect is weighted heavier on the negative side as seen on the middle line, but the effect can be either positive or negative. Some of the next day effect on *RET1* revert fast on day 3, while the rest dissipate slowly over the remaining period.

The IRF model suggests that a weather shock are, in fact, influencing the natural gas returns, but the magnitude of the response are marginal, and may also not occur. This contradicts the findings provided in both the linear regression model and the VAR model without shocks. These findings are quite interesting, as they can support our believes.

*Figure 4-8: The Orthogonal Impulse Response from dStDev*

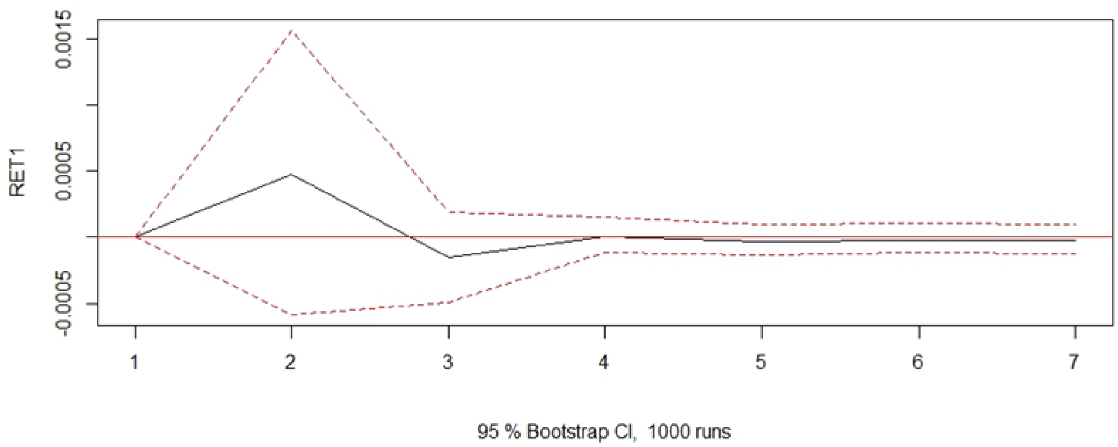



Figure 4-8 illustrate the response of natural gas returns to a shock in the storage level. The initial effect of a storage shock will most likely cause *RET1* to increase. The confidence interval include both positive and negative possible responses, rendering the actual direction of the response unknown. If an impact toward *RET1* occurs, the effect is short and reverts toward zero fast but retain a small rippling effect form day 4.

*Figure 4-9: The Orthogonal Impulse Response of dStDev to a shock in WtF*

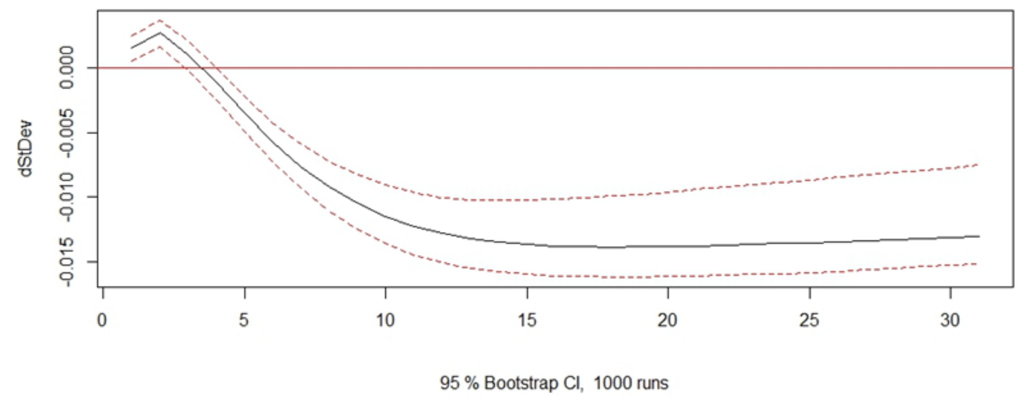



Figure 4-9 illustrates that a weather shock initially increases the amount of natural gas in storage. The effect reverses after approximately three days, resulting in a vast drop in the storage level, which continues to decrease until day 10, where it starts to flatten out and the confidence interval becomes wider, increasing the uncertainty of the reaction.

Based on the findings above, we conclude that weather and storage shocks can affect natural gas return. The magnitude of the response is small, when the shock equals one standard deviation. The y-axis contains the numeric value of *RET1*'s response. This either means that the expected response within a 95% confidence interval of a weather shock is that the returns can decrease by 0.1%, or increase by 0.05%. The effect of a storage shock is that the returns can decrease by 0.05 or increase by 0.15%. The most significant discovery is that the storage level has a negative response to a weather shock, draining the storage after the shock occurs, with a lasting effect.

The IRF hint toward an integrated relationship between the three endogenous variables, both weather and storage shocks causes the return to change, but we are unable to conclude whether these responses are positive or negative. The evidence are unfortunately ambiguous, since the response of natural gas returns can both be positive or negative, which also indicate that the response may not materialize. This is just suggesting that the relationship is exciting, so we are unable to conclude that it is statistically significant.

#### 4.2.4.1. Forecast Error Variance Decomposition (FEVD)

The forecast error variance decomposition (FEVD) tables include values that are normalized to sum to unity, this means that the magnitude of the values indicates how much of the change in variable *k* are driven by the different *j* variables, as we apply the IRF to the VAR model.

*Table 4-14: FEVD of k=RET1*

| | *RET1* | *WtF* | *dStDev* |
|---|---|---|---|
| [1,] | 100,00 % | 0,00 % | 0,00 % |
| [2,] | 99,93 % | 0,02 % | 0,05 % |
| [3,] | 99,92 % | 0,02 % | 0,05 % |
| [4,] | 99,92 % | 0,03 % | 0,05 % |
| [5,] | 99,92 % | 0,03 % | 0,06 % |
| | ⋮ | ⋮ | ⋮ |
| [16,] | 99,91 % | 0,03 % | 0,06 % |
| [17,] | 99,91 % | 0,03 % | 0,06 % |
| [18,] | 99,91 % | 0,03 % | 0,06 % |
| [19,] | 99,91 % | 0,03 % | 0,06 % |
| [20,] | 99,91 % | 0,03 % | 0,06 % |
| [30,] | 99,91 % | 0,03 % | 0,06 % |

*The first column represents the response time, in days forward in time, and the three remaining column represent the contribution to the response in RET1.*

Table 4-14 lists how a shock in either one of the three endogenous variables affect *RET1*, and we can see that neither weather, nor storage has an impact on the return. These findings indicate that the response of *RET1* is almost exclusively driven by itself.

*Table 4-15: FEVD of k=dStDev*

| | *RET1* | *WtF* | *dStDev* |
|---|---|---|---|
| [1,] | 0,04 % | 1,02 % | 98,94 % |
| | ⋮ | ⋮ | ⋮ |
| [5,] | 0,32 % | 3,27 % | 96,40 % |
| [6,] | 0,54 % | 6,32 % | 93,13 % |
| [7,] | 0,80 % | 10,65 % | 88,56 % |
| [8,] | 1,06 % | 15,51 % | 83,43 % |
| [9,] | 1,31 % | 20,35 % | 78,34 % |
| [10,] | 1,53 % | 24,84 % | 73,64 % |
| [11,] | 1,72 % | 28,82 % | 69,47 % |
| [12,] | 1,88 % | 32,27 % | 65,85 % |
| [13,] | 2,02 % | 35,24 % | 62,75 % |
| [14,] | 2,13 % | 37,77 % | 60,10 % |
| [15,] | 2,23 % | 39,93 % | 57,83 % |
| | ⋮ | ⋮ | ⋮ |
| [20,] | 2,55 % | 47,02 % | 50,43 % |
| [30,] | 2,82 % | 52,94 % | 44,24 % |
| [40,] | 2,93 % | 55,43 % | 41,64 % |
| [50,] | 3,00 % | 56,78 % | 40,22 % |

The first observations in Table 4-15 also starts out with the same interpretation as *RET1* and *WtF*, where the main force driving a change in storage level is a shock to itself. As opposed to the previous relationships, we can see that after day 5 the force of change shift to include the weather variable as well. The weather variables contribution to the storage change increases to eventually be the main force driving the change, while return contribute to the smallest extent.

The main result from the FEVD analysis is that the storage level is more and more affected by the weather variable. These findings provide evidence supporting the notion that a weather shock will affect the level of storage, with a lagged effect, which Figure 4-9 in the IRF chapter illustrates.

Figure 4-10 to Figure 4-12 provides another presentation of the FEVD output, where we can observe how each variable *k*, in each figure, are affected by shock in variables *j*.

*Figure 4-10: FEVD k=RET1*

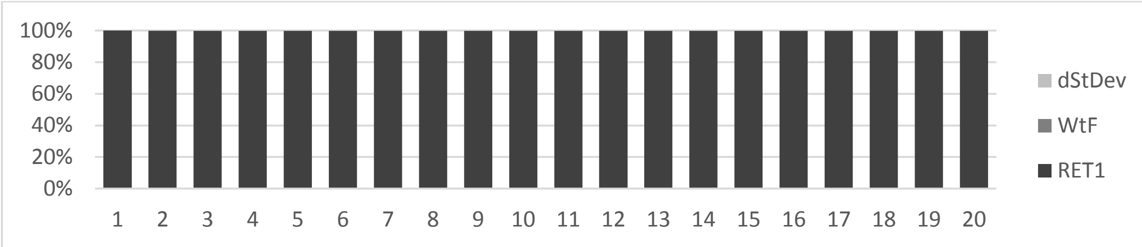



*Figure 4-11: FEVD k=WtF*

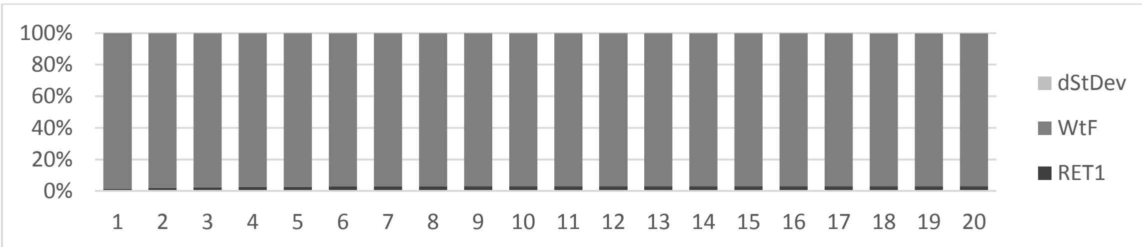
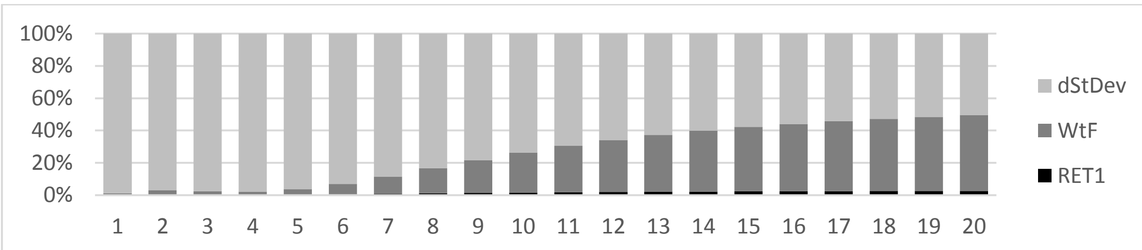



*Figure 4-12: FEVD k=dStDev*



## 4.3. Summary

To detect the volatility of the return series, we estimated a GARCH model. This resulted in an alpha value for *RET1* of 0.13, implying a jumpy market, and a beta value of 0.86, representing the persistence of the volatility after a shock occurs. These two combined are consistent with a moderately spiky market. We constructed the *RET1* series in a way that will account for the rollover effect, which occur when traders takes a similar position before closing out their old position in the futures market, resulting in excess retunes and volatility. This resulted in a reduced volatility in the *RET1* series, and the GARCH model fitted the data well. The residual tests conclude that we have successfully removed all traces of autocorrelation.

*RET2* did not go through the same precautions as *RET1*, resulting in a more volatile series. This were confirmed through the GARCH estimation, where we obtained a high alpha of 0.39 and low beta of 0.60. Both of these values fall outside their usual ranges and represent a much more spiky market than *RET1*. The residual tests revealed that *RET2* suffered from autocorrelation.

From the GARCH analysis we can conclude that there are a good amount of spikes at irregular intervals, this can make it problematic when making precision forecasting or estimating a good fit regression as the variables included need to fit the unique behavior displayed in return. As a direct consequence of these results, we decided to concentrate on *RET1* in the remaining part of the analysis.

We choose to employ two lags in the applied VAR model. The *RET1* variable consists of insignificant endogenous variables, but most of the exogenous variables were significant, resulting in a low adjusted $R^2$, of 1.5%. The model with respect to storage contains significant coefficients for both storage and weather at the 0.1% level, rendering a good fit, with an adjusted $R^2$, of approximately 99%. The first weather lag coefficient is positive, increasing the storage level. The second lag coefficient is negative, reducing the storage level.

The regression model explaining the weather variable produced an adjusted $R^2$, of approximately 92%, implying that it predict the change in the dependent variable quite well. This model's good fit is a result of including itself as explanatory variables, which were statistically significant at the 0.1% level. Weather is a force of nature and is not affected by economic variables, thus including other variables gives no meaning. In spite of this, we include the weather as an endogenous variable, since we expect the weather to affect both return and storage.

From the VAR models, we could not conclude that there is any significant influence affecting the returns based on our weather or storage variables. However, we did discover a statistically significant weather effect on storage.

We included the IRF model to see how simulated shocks affect the VAR model. The difference between the results from the VAR mode and the IRF analysis, is that the VAR model aim to define an integration between the endogenous variables, while the IRF applies shocks to see how this cointegration behave when introduced to extreme values.

The IRF model is unable to contribute with significant results differing from results previously discovered by the VAR model. We found that shocks simulated to each endogenous variable resulted in an ambiguous response in the return. The response had possibility of turning out to be either positive or negative, including the possibility of no response in the aftermath of the shock.

When simulating a shock in the weather variable, the IRF model found the response of the storage variable to be significant, where the storage's initial response is a small buildup, before depleting over the following periods. This supports the results from VAR, where the weather coefficients provided similar traits for the respective lags.

# 5. Conclusion

The objective in this thesis is to investigate whether natural gas prices in the UK are affected by weather or storage. We also aim to compare our findings with the findings from Mu (2007).

For our analysis of this relationship we made us of GARCH, VAR, and IRF. The models have been studied individually and simultaneously in an effort to enhance our understanding of their cohesion. The design of the VAR model were unable to acquire significant results in favor of cointegration between *RET1* and the two other variables, consequently failing to reject that there are no interaction between them. We found significant evidence favoring that the weather is affecting the storage levels. The IRF found that both weather and storage shocks resulted in a response in the returns, but the evidence were equivocal, as there is no guarantee that the response actually will materialize. Thus, the IRF results are similar to the findings in VAR.

During the development of this paper we found there to be no statistically significant reactions in return due to shocks in either weather or storage. However, we did find evidence suggesting that weather influence storage.

Compared to the findings of Mu (2007) the effects of weather and storage were not as extensive in UK as in the U.S. When comparing the data we found that the American market had a greater variety in temperature, shown as degree-days. This may be a result of discrepancies in the demographic properties of these two countries, rendering them difficult to compare. Due to limited amount of time we decided to construct our storage variable without using furrier series, causing some discrepancies. There is also the fact that storage levels are not published at regular weekly intervals, but rather in a continuous lagged stream in the UK, increasing the distance between the papers. Mu (2007) found a significant weather effect on the conditional means of natural gas returns, whereas we were unable to provide statistically significant evidence in favor of this relationship.

Other academics or students interested in researching the effect of weather and or storage on natural gas prices can utilize our findings. We have created an orderly and detailed overview of our approach and findings through this paper, that when applied can help toward speeding up future research.

## 6. Limitations, and Improvements

Throughout the process of working on this thesis, we have gained enhanced knowledge of both how theories and analysis adopted works. This has given us greater insight into the subjects, and we have arrived at some possible procedures that may improve the thesis, which may be of interest for continued or later study.

When comparing the weather variables for UK and U.S., we can see that the American weather gives a much larger variety in temperatures. As opposed to the UK, where the weather anomalies are less frequent, visualized through a lower variation ranging between 5 to -4 degrees. This is graphically represented in

Appendix 13 in our thesis and as Fig 4.A on page 57 in Mu (2007).

The sample size of storage data are limited to contain 5 years of daily observations, which result in restrictions to the storage shock variable. The estimated storage shock variable ($dStDev_t$) is calculated using equation (27). With a sample size of 5 years this estimation will be based upon 5 observations, increasing the amount of historic data would improve the estimated coefficients.

We were unable to obtain the demand of natural gas accumulated by the domestic sector in UK, and decided to use the total demand for natural gas as a proxy for private demand. Total demand includes sectors such as industry, that are mostly unaffected by seasonality compared to domestic demand, resulting in a less evident effect.

The VAR model could have been improved by estimating a Structural VAR (SVAR) model, which would enable us to decide the direction of cointegration. The weather variable can be endogenous, where the direction of the dynamics would eliminate the spurious results obtained from the VAR model.

A possible limitation of the OIRF, where only a single variable are shocked at a time, is that we are only looking at the big picture piecemeal as it is likely that a shock on one of the variables will cause repercussions on the other variables. This could be improved by applying a generalized IRF (GIRF) to calculate how a shock will affect the return in a more realistic manner. But due to a lack of available computerized models (not available in R) we were unable to do this.

# Bibliography

Akaike, H., 1973. Information theory and an extension of the maximum likelihood principle. *2nd International Symposium on Information Theory*, p. 267–281.

Akaike, H., 1974. A new look at the statistical model identification. *IEEE Transactions on Automatic Control,* ac-19(6), pp. 716-723.
Akaike, H., 1974. A new look at the statistical model identification. *IEEE Transactions on Automatic Control,* ac-19(6), pp. 716-723.
Alexander, C., 2008. *Market Risk Analysis: Pricing, Hedging and Trading Financial Instruments.* Volume III ed. West Sussex: Wiley.

Bollerslev, T., 1986. Generalized autoregressive conditional Heteroskedasticity. *Journal of Econometrics*, Volume 31, pp. 307-327.

Breusch, T. S. & Pagan, A. R., 1979. A Simple Test for Heteroscedasticity and Random Coefficient Variation. *Econometrica,* 47(5), pp. 1287-1294.

Bulmer, M. G., 1979. *Principles pf Statistics.* New York: Dover Publications Inc..

Davey, E., 2012. *Gas Generation Strategy,* Norwich: The Stationery Office.

Department of Energy & Climate Change, 2014. *Energy Trends: December 2014,* London: National Statistics.

Enders, W., 2009. *Applied Econometric Time Series.* 3 ed. Hoboken: Wiley.

Energy Information Administration, 2012. *Derivatives and Risk Management in the Petroleum, Natural Gas, and Electricity Industries,* Washington DC: U.S. Department of Energy.

Engle, R. F., 1982. Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica,* 50(4), pp. 987-1007.

Godfrey, L. G., 1978. Testing for Higher Order Serial Correlation in Regression Equations when the Regressors Include Lagged Dependent Variables. *Econometrica,* 46(6), pp. 1303-1310.

Grogan, N., 2015. *Gas data 2003-*. Penicuik: Energy Solutions.

Haff, I. H., Lindqvist, O. & Løland, A., 2008. Risk premium in the UK natural gas forward market. *Energy Economics 30,* p. 2420–2440.

Hanna, E. J. & Quinn, B. G., 1979. The Determination of the Order of an Autoregression. *Journal of the Royal Statistical Society. Series B (Methodological),* 41(2), pp. 190-195.

Heather, P., 2010. *The Evolution and Functioning of the Traded Gas Market in Britain,* Oxford: Oxford Institute for Energy Studies.

Hull, J. C., 2012. *Options, futures, and other derivatives.* Boston: Pearson Education, Inc..

Jarque, C. M. & Bera, A. K., 1980. EFFICIENT TESTS FOR NORMALITY, HOMOSCEDASTICITY AND SERIAL INDEPENDENCE OF REGRESSION RESIDUALS. *Economics Letters,* Volume 6, pp. 255-259.

Koop, G., Pesaran, M. H. & Potter, S. M., 1996. Impulse response analysis in nonlinear multivariate models. *Journal of Econometrics*, Volume 74, pp. 119-147.

Ljung, G. M. & Box, G. E. P., 1978. On a measure of lack of fit in time series models. *Biometrika,* 65(2), pp. 297-303.

Lütkepohl, H., 2011. *Vector Autoregressive Models,* Florence: European University Institute.

MacLeay, I., Harris, K. & Annut, A., 2014. *Digest of United Kingdom Energy Statistics 2014,* London: Department of Energy & Climate Change.

Mu, X., 2007. Weather, storage, and natural gas price dynamics: Fundamentals and volatility. *Energy Economics ,* Volume 29, p. 46–63.

Parker, D. E., Legg, T. P. & Folland, C. K., 1992. A new daily Central England Temperature Series, 1772-1991. *International Journal of Climatology,* pp. 317-342.

Pesaran, M. H. & Shin, Y., 1998. Generalized Impulse Response Analysis in Linear Multivariate Models. *Economics Letters 58,* p. 17–29.

Pfaff, B., 2008. VAR, SVAR and SVEC Models: Implementation Within R Package vars. *Journal of Statistical Software*, 27(4).

Schwarz, G., 1978. Estimating the dimension of a model. *Annals of Statistics*, Volume 6, pp. 461-464.

Shapiro, S. S. & Wilk, M. B., 1965. An Analysis of Variance Test for Normality (Complete Samples). *Biometrika*, 52(3/4), pp. 591-611.

Subak, S. et al., 2000. The Impact of the Anomalous Weather of 1995 on the U.K. Economy. *Climatic Change*, Volume 44, p. 1–26.

Wooldridge, J. M., 2012. *Introductory Econometrics: A modern approach.* s.l.:South-Western Cengage Learning.

Zivot, E. & Wang, J., 2006. Vector Autoregressive Models for Multivariate Time Series. In: *Modeling Financial Time Series with S-Plus.* s.l.:Springer, pp. 385-428.

## Web pages

Chicago Mercantile Exchange Group, 2015. *CME Group.* [Online]

[Accessed 10 May 2015].

Federal Reserve Bank of St. Louis, 2015. *Economic Research.* [Online]

Available at: http://research.stlouisfed.org/

[Accessed 11 March 2015].

Füss, R., 2007. *University of Feriburg.* [Online]

Available at: https://www.empiwifo.uni-freiburg.de/lehre-teaching-1/winter-term/dateien-financial-data-analysis/chapter4.pdf

[Accessed 16 March 2015].

ICIS Heren, 2015. *ICIS.* [Online]

Available at: http://www.icis.com/energy/gas/

[Accessed 10 February 2015].

National Grid, 2015. *National Grid.* [Online]

Available at: http://www2.nationalgrid.com/uk/industry-information/gas-transmission-operational-data/data-item-explorer/

[Accessed 08 February 2015].

OzForex, 2011. *OzForex.* [Online]

Available at: http://www.ozforex.com.au/about-us/contact-us

[Accessed 14 March 2015].

US Department of Energy, 2015. *Quandl.* [Online]

Available at: https://www.quandl.com/data/DOE/RBRTE-Europe-Brent-Crude-Oil-Spot-Price-FOB

[Accessed 14 March 2015].

Yahoo! Finance, 2015. *Yahoo! Finance.* [Online]

Available at:

https://finance.yahoo.com/q/hp?s=%5EGSPC&a=00&b=1&c=2010&d=02&e=13&f=201

5&g=d
[Accessed 13 03 2015].

# 7. Appendix

*Appendix 1-Table: ADF output*

| | *Test-statistic* | *Adjusted* $R^2$ | *F-statistic* | *p-value* |
|---|---|---|---|---|
| *M1* | *0.0529* | *-0.001865* | *0.6667* | *0.7005* |
| *RET1* | *-13.6521* | *0.488* | *171.6* | *2.2e-16* |
| *RET2* | *-12.1303* | *0.4457* | *144.9* | *2.2e-16* |
| *WtF* | *-7.8364* | *0.33* | *89.17* | *2.2e-16* |
| *dStDev* | *-2.2059* | *0.08635* | *17.92* | *2.2e-16* |
| *BOil* | *-0.4899* | *0.002096* | *1.376* | *0.2115* |
| *retBOil* | *-12.2692* | *0.5345* | *206.5* | *2.2e-16* |
| *TBill* | *-0.3118* | *0.4041* | *122.4* | *2.2e-16* |
| *retT.Bill* | *-6.5056* | *0.3011* | *78.12* | *2.2e-16* |
| *SP* | *2.1784* | *0.01295* | *3.348* | *0.001532* |
| *retSP* | *-14.388* | *0.5372* | *208.8* | *2.2e-16* |
| *Demand* | *-0.8006* | *0.407* | *123.9* | *2.2e-16* |
| *lnDem* | *-8.9478* | *0.744* | *521.3* | *2.2e-16* |
| | | 1% | 5% | 10% |
| Critical value for test-statistic | | \|2.58\| | \|1.95\| | \|1.62\| |

*1247 degrees of freedom conducted on 1260 observations.*

*Appendix 2-Table: Dummy Variable output RET1*

| | Estimate | Std. Error | t-value | p-value | |
|---|---|---|---|---|---|
| *(Intercept)* | 0.0035347 | 0.0024712 | 1.430 | 0.15286 | |
| *dStDev* | 0.0071111 | 0.0051497 | 1.381 | 0.16757 | |
| *WtF* | 0.0007813 | 0.0004508 | 1.733 | 0.08329 | . |
| *retBOil* | 0.1124095 | 0.0618989 | 1.816 | 0.06961 | . |
| *retTBill* | 0.1940009 | 0.2201376 | 0.881 | 0.37834 | |
| *retSP* | -0.1171473 | 0.0853126 | -1.373 | 0.16995 | |
| *LnDem* | -0.1320692 | 0.0497276 | -2.656 | 0.00801 | ** |
| *dummydStDev* | -0.0041555 | 0.0018753 | -2.216 | 0.02688 | * |
| *dummyWtF* | -0.0025365 | 0.0019259 | -1.317 | 0.18807 | |
| *dummyretBOil* | 0.0014255 | 0.0018033 | 0.791 | 0.42937 | |
| *dummyTBill* | -0.0006779 | 0.0016409 | -0.413 | 0.67960 | |
| *dummyretSP* | -0.0018052 | 0.0016575 | -1.089 | 0.27632 | |
| *dummylnDem* | 0.0034730 | 0.0015773 | 2.202 | 0.02786 | * |
| *dummySummer* | 0.0008548 | 0.0015239 | 0.561 | 0.57494 | |
| *dummyWinter* | -0.0032473 | 0.0014814 | -2.192 | 0.02856 | * |

Residual standard error: 0.02106 on 1245 degrees of freedom. Multiple R-squared: 0.03026. Adjusted R-squared: 0.01936. F-statistic: 2.775 on 14 and 1245 DF, p-value: 0.0004448.
'***' '**' '*' '.' Denote significance at the 0.1%, 1%, 5% and 10% level

*Appendix 3-Table: Dummy Variable output RET2*

| | Estimate | Std. Error | t-value | p-value | |
|---|---|---|---|---|---|
| *(Intercept)* | 0.0079369 | 0.0034988 | 2.268 | 0.023472 | * |
| *dStDev* | 0.0140272 | 0.0072910 | 1.924 | 0.054596 | . |
| *WtF* | 0.0023015 | 0.0006382 | 3.606 | 0.000323 | *** |
| *retBOil* | -0.0484040 | 0.0876377 | -0.552 | 0.580829 | |
| *retTBill* | 0.4610458 | 0.3116753 | 1.479 | 0.139326 | |
| *retSP* | -0.0278702 | 0.1207873 | -0.231 | 0.817556 | |
| *LnDem* | -0.1140397 | 0.0704053 | -1.620 | 0.105537 | |
| *dummydStDev* | -0.0074152 | 0.0026551 | -2.793 | 0.005306 | ** |
| *dummyWtF* | -0.0073945 | 0.0027267 | -2.712 | 0.006783 | ** |
| *dummyretBOil* | 0.0038674 | 0.0025531 | 1.515 | 0.130086 | |
| *dummyTBill* | -0.0021719 | 0.0023232 | -0.935 | 0.350046 | |
| *dummyretSP* | -0.0022824 | 0.0023467 | -0.973 | 0.330937 | |
| *dummylnDem* | 0.0051046 | 0.0022332 | 2.286 | 0.022432 | * |
| *dummySummer* | 0.0015474 | 0.0021576 | 0.717 | 0.473397 | |
| *dummyWinter* | -0.0058369 | 0.0020974 | -2.783 | 0.005469 | ** |

Residual standard error: 0.02981 on 1245 degrees of freedom. Multiple R-squared: 0.03094. Adjusted R-squared: 0.02004. F-statistic: 2.839 on 14 and 1245 DF, p-value: 0.0003266.
'***' '**' '*' '.' Denote significance at the 0.1%, 1%, 5% and 10% level

*Appendix 4-Table: Squared Variables output RET1*

| | Estimate | Std. Error | t-value | p-value | |
|---|---|---|---|---|---|
| <u>*RET1*</u> | | | | | |
| *(Intercept)* | -8.995e-04 | 8.748e-04 | -1.028 | 0.30401 | |
| *sqdStDev* | -7.941e-03 | 1.109e-02 | -0.716 | 0.47426 | |
| *sqWtF* | 7.045e-05 | 7.187e-05 | 0.980 | 0.32719 | |
| *sqBOil* | 1.082e-01 | 1.421e+00 | 0.076 | 0.93933 | |
| *sqTBill* | 9.224e+00 | 9.611e+00 | 0.960 | 0.33740 | |
| *sqSP* | 3.770e+00 | 2.354e+00 | 1.601 | 0.10959 | |
| *sqLnDem* | 1.702e+00 | 5.758e-01 | 2.956 | 0.00318 | ** |
| <u>*RET2*</u> | | | | | |
| *(Intercept)* | -0.0006250 | 0.0012386 | -0.505 | 0.6139 | |
| *sqdStDev* | -0.0204552 | 0.0157073 | -1.302 | 0.1931 | |
| *sqWtF* | 0.0001729 | 0.0001018 | 1.699 | 0.0896 | . |
| *sqBOil* | -1.1226785 | 2.0113927 | -0.558 | 0.5768 | |
| *sqTBill* | 16.0418086 | 13.6083436 | 1.179 | 0.2387 | |
| *sqSP* | 6.7717935 | 3.3336024 | 2.031 | 0.0424 | * |
| *sqLnDem* | 1.7471013 | 0.8152873 | 2.143 | 0.0323 | * |

Residual standard error: 0.0212 on 1253 degrees of freedom. Multiple R-squared: 0.01116. Adjusted R-squ ared: 0.006421. F-statistic: 2.356 on 6 and 1253 DF, p-value: 0.02877.
Residual standard error: 0.03001 on 1253 degrees of freedom. Multiple R-squared: 0.01176. Adjusted R-squared: 0.007027. F-statistic: 2.485 on 6 and 1253 DF, p-value: 0.0215
'***' '**' '*' '.' Denote significance at the 0.1%, 1%, 5% and 10% level

*Appendix 5-Table: Squared Variables output RET2*

| | Estimate | Std. Error | t-value | p-value | |
|---|---|---|---|---|---|
| *(Intercept)* | 0.0001121 | 0.0016881 | 0.066 | 0.94707 | |
| *WtF* | 0.0007293 | 0.0004518 | 1.614 | 0.10678 | |
| *dStDev* | -0.0070425 | 0.0053727 | -1.311 | 0.19017 | |
| *retBOil* | 0.0523880 | 0.0593565 | 0.883 | 0.37762 | |
| *retTBill* | 0.4807734 | 0.2656357 | 1.810 | 0.07055 | . |
| *retSP* | -0.1200409 | 0.0903894 | -1.328 | 0.18441 | |
| *LnDem* | -0.0638283 | 0.0755479 | -0.845 | 0.39834 | |
| *dummylnDem* | 0.0037141 | 0.0022931 | 1.620 | 0.10556 | |
| *dummyWinter* | -0.0059788 | 0.0020752 | -2.881 | 0.00403 | ** |
| *sqdStDev* | -0.0377334 | 0.0184213 | -2.048 | 0.04073 | * |
| *sqWtF* | 0.0001295 | 0.0001150 | 1.126 | 0.26026 | |
| *sqBOil* | -1.0926952 | 2.0144739 | -0.542 | 0.58762 | |
| *sqTBill* | 19.5842547 | 14.1798223 | 1.381 | 0.16749 | |
| *sqSP* | 5.1396848 | 3.4062700 | 1.509 | 0.13158 | |
| *sqLnDem* | 1.3787042 | 0.8722363 | 1.581 | 0.11421 | |

Residual standard error: 0.02989 on 1245 degrees of freedom. Multiple R-squared:  0.02607. Adjusted R-squared:  0.01512. F-statistic: 2.381 on 14 and 1245 DF, p-value: 0.002833.
'***' '**' '*' '.' Denote significance at the 0.1%, 1%, 5% and 10% level

*Appendix 6-Table: VAR RET1 p=7*

| | Estimate | Std. Error t | t-value | Pr(>\|t\|) | |
|---|---|---|---|---|---|
| *RET1.l1* | 2.971e-02 | 2.851e-02 | 1.042 | 0.297539 | |
| *RET1.l2* | -3.567e-02 | 2.848e-02 | -1.253 | 0.210558 | |
| *RET1.l3* | -2.549e-02 | 2.848e-02 | -0.895 | 0.370817 | |
| *RET1.l4* | -3.971e-02 | 2.857e-02 | -1.390 | 0.164740 | |
| *RET1.l5* | 1.213e-02 | 2.853e-02 | 0.425 | 0.670699 | |
| *RET1.l6* | -3.528e-02 | 2.851e-02 | -1.237 | 0.216197 | |
| *RET1.l7* | 3.034e-02 | 2.832e-02 | 1.071 | 0.284280 | |
| *dStDev.l1* | 2.269e-02 | 4.052e-02 | 0.560 | 0.575654 | |
| *dStDev.l2* | -7.535e-02 | 4.855e-02 | -1.552 | 0.120892 | |
| *dStDev.l3* | 5.740e-02 | 5.025e-02 | 1.142 | 0.253545 | |
| *dStDev.l4* | 3.484e-02 | 5.026e-02 | 0.693 | 0.488359 | |
| *dStDev.l5* | -3.108e-02 | 5.022e-02 | -0.619 | 0.536078 | |
| *dStDev.l6* | -6.697e-02 | 4.854e-02 | -1.380 | 0.167962 | |
| *dStDev.l7* | 5.650e-02 | 3.984e-02 | 1.418 | 0.156452 | |
| *WtF.l1* | 4.198e-04 | 1.066e-03 | 0.394 | 0.693797 | |
| *WtF.l2* | -1.032e-03 | 1.754e-03 | -0.588 | 0.556343 | |
| *WtF.l3* | 3.869e-04 | 1.761e-03 | 0.220 | 0.826127 | |
| *WtF.l4* | 2.436e-04 | 1.752e-03 | 0.139 | 0.889443 | |
| *WtF.l5* | -1.370e-03 | 1.746e-03 | -0.784 | 0.432949 | |
| *WtF.l6* | 3.552e-03 | 1.741e-03 | 2.040 | 0.041601 | * |
| *WtF.l7* | -2.219e-03 | 1.121e-03 | -1.980 | 0.047965 | * |
| *const* | 1.122e-03 | 1.359e-03 | 0.826 | 0.409057 | |
| *trend* | -1.896e-06 | 1.862e-06 | -1.018 | 0.308705 | |
| *retBOil* | 1.626e-01 | 4.234e-02 | 3.839 | 0.000130 | *** |
| *retTBill* | 1.237e-01 | 1.806e-01 | 0.685 | 0.493330 | |
| *retSP* | -2.126e-01 | 6.400e-02 | -3.323 | 0.000918 | *** |
| *sqLnDem* | 1.647e+00 | 5.721e-01 | 2.879 | 0.004056 | ** |

VAR estimation with *RET1*, *WtF*, and *dStDev* as endogenous variables, with respect to *RET1*. 1260 observations. Log likelihood: 5550.364. Residual standard error: 0.02107 on 1226 degrees of freedom. Multiple R-Squared: 0.03666. Adjusted R-squared: 0.01623. F-statistic: 1.795 on 26 and 1226 DF, p-value: 0.008578. '***' '**' '*' '.' Denote significance at the 0.1%, 1%, 5% and 10% level

*Appendix 7-Table: VAR dStDev p=7*

| | Estimate | Std. Error | t-value | Pr(>\|t\|) | |
|---|---|---|---|---|---|
| *RET1.l1* | -1.345e-02 | 2.004e-02 | -0.671 | 0.502149 | |
| *RET1.l2* | -3.684e-03 | 2.002e-02 | -0.184 | 0.854032 | |
| *RET1.l3* | 1.063e-02 | 2.002e-02 | 0.531 | 0.595667 | |
| *RET1.l4* | 2.756e-03 | 2.008e-02 | 0.137 | 0.890890 | |
| *RET1.l5* | 9.845e-03 | 2.006e-02 | 0.491 | 0.623641 | |
| *RET1.l6* | 9.784e-04 | 2.005e-02 | 0.049 | 0.961078 | |
| *RET1.l7* | 1.976e-02 | 1.991e-02 | 0.992 | 0.321240 | |
| *dStDev.l1* | 6.568e-01 | 2.849e-02 | 23.053 | < 2e-16 | *** |
| *dStDev.l2* | 3.276e-01 | 3.413e-02 | 9.598 | < 2e-16 | *** |
| *dStDev.l3* | 6.763e-02 | 3.533e-02 | 1.914 | 0.055817 | . |
| *dStDev.l4* | 1.199e-03 | 3.534e-02 | 0.034 | 0.972945 | |
| *dStDev.l5* | 4.995e-02 | 3.531e-02 | 1.415 | 0.157462 | |
| *dStDev.l6* | -6.380e-02 | 3.413e-02 | -1.869 | 0.061816 | . |
| *dStDev.l7* | -5.336e-02 | 2.801e-02 | -1.905 | 0.057028 | . |
| *WtF.l1* | 8.733e-04 | 7.495e-04 | 1.165 | 0.244168 | |
| *WtF.l2* | -2.186e-03 | 1.233e-03 | -1.773 | 0.076558 | . |
| *WtF.l3* | -1.984e-04 | 1.238e-03 | -0.160 | 0.872722 | |
| *WtF.l4* | -1.448e-03 | 1.232e-03 | -1.175 | 0.240155 | |
| *WtF.l5* | -3.005e-03 | 1.228e-03 | -2.448 | 0.014516 | * |
| *WtF.l6* | 2.751e-03 | 1.224e-03 | 2.247 | 0.024840 | * |
| *WtF.l7* | 1.273e-04 | 7.882e-04 | 0.162 | 0.871716 | |
| *const* | 3.683e-03 | 9.553e-04 | 3.855 | 0.000122 | *** |
| *trend* | -5.307e-06 | 1.309e-06 | -4.054 | 5.36e-05 | *** |
| *retBOil* | 4.235e-02 | 2.977e-02 | 1.422 | 0.155141 | |
| *retTBill* | -2.908e-01 | 1.270e-01 | -2.290 | 0.022168 | * |
| *retSP* | -2.535e-02 | 4.500e-02 | -0.563 | 0.573288 | |
| *sqLnDem* | 6.565e-01 | 4.023e-01 | 1.632 | 0.102960 | |

VAR estimation with *RET1*, *WtF*, and *dStDev* as endogenous variables, with respect to *dStDev*. 1260 observations. Log likelihood: 5550.364. Residual standard error: 0.01481 on 1226 degrees of freedom. Multiple R-Squared: 0.9935. Adjusted R-squared: 0.9934. F-statistic: 7210 on 26 and 1226 DF, p-value: < 2.2e-16. ‘***’ ‘**’ ‘*’ ‘.’ Denote significance at the 0.1%, 1%, 5% and 10% level

*Appendix 8-Table: VAR WtF p=7*

| | Estimate | Std. Error | t-value | p-value | |
|---|---|---|---|---|---|
| *RET1.l1* | 1.161e+00 | 7.642e-01 | 1.520 | 0.1288 | |
| *RET1.l2* | 1.108e-01 | 7.634e-01 | 0.145 | 0.8846 | |
| *RET1.l3* | 1.387e+00 | 7.634e-01 | 1.817 | 0.0695 | . |
| *RET1.l4* | 1.233e+00 | 7.658e-01 | 1.610 | 0.1076 | |
| *RET1.l5* | 4.385e-01 | 7.648e-01 | 0.573 | 0.5665 | |
| *RET1.l6* | 1.923e-02 | 7.643e-01 | 0.025 | 0.9799 | |
| *RET1.l7* | 6.914e-01 | 7.593e-01 | 0.911 | 0.3627 | |
| *dStDev.l1* | -1.743e+00 | 1.086e+00 | -1.605 | 0.1088 | |
| *dStDev.l2* | -4.872e-01 | 1.301e+00 | -0.374 | 0.7082 | |
| *dStDev.l3* | 2.018e+00 | 1.347e+00 | 1.498 | 0.1344 | |
| *dStDev.l4* | 2.561e+00 | 1.347e+00 | 1.901 | 0.0576 | . |
| *dStDev.l5* | 2.937e-02 | 1.346e+00 | 0.022 | 0.9826 | |
| *dStDev.l6* | -2.504e+00 | 1.301e+00 | -1.924 | 0.0545 | . |
| *dStDev.l7* | 1.316e-02 | 1.068e+00 | 0.012 | 0.9902 | |
| *WtF.l1* | 1.360e+00 | 2.858e-02 | 47.576 | < 2e-16 | *** |
| *WtF.l2* | -3.255e-01 | 4.701e-02 | -6.924 | 7.06e-12 | *** |
| *WtF.l3* | -7.067e-02 | 4.721e-02 | -1.497 | 0.1347 | |
| *WtF.l4* | -6.048e-02 | 4.698e-02 | -1.287 | 0.1982 | |
| *WtF.l5* | -2.819e-01 | 4.681e-02 | -6.022 | 2.27e-09 | *** |
| *WtF.l6* | 4.410e-01 | 4.668e-02 | 9.446 | < 2e-16 | *** |
| *WtF.l7* | -1.441e-01 | 3.005e-02 | -4.796 | 1.81e-06 | *** |
| *const* | 7.315e-02 | 3.642e-02 | 2.008 | 0.0448 | * |
| *trend* | -8.954e-05 | 4.991e-05 | -1.794 | 0.0731 | . |
| *retBOil* | 8.772e-01 | 1.135e+00 | 0.773 | 0.4398 | |
| *retTBill* | -7.047e+00 | 4.841e+00 | -1.456 | 0.1457 | |
| *retSP* | -2.773e-01 | 1.716e+00 | -0.162 | 0.8716 | |
| *sqLnDem* | -7.921e+00 | 1.534e+01 | -0.516 | 0.6056 | |

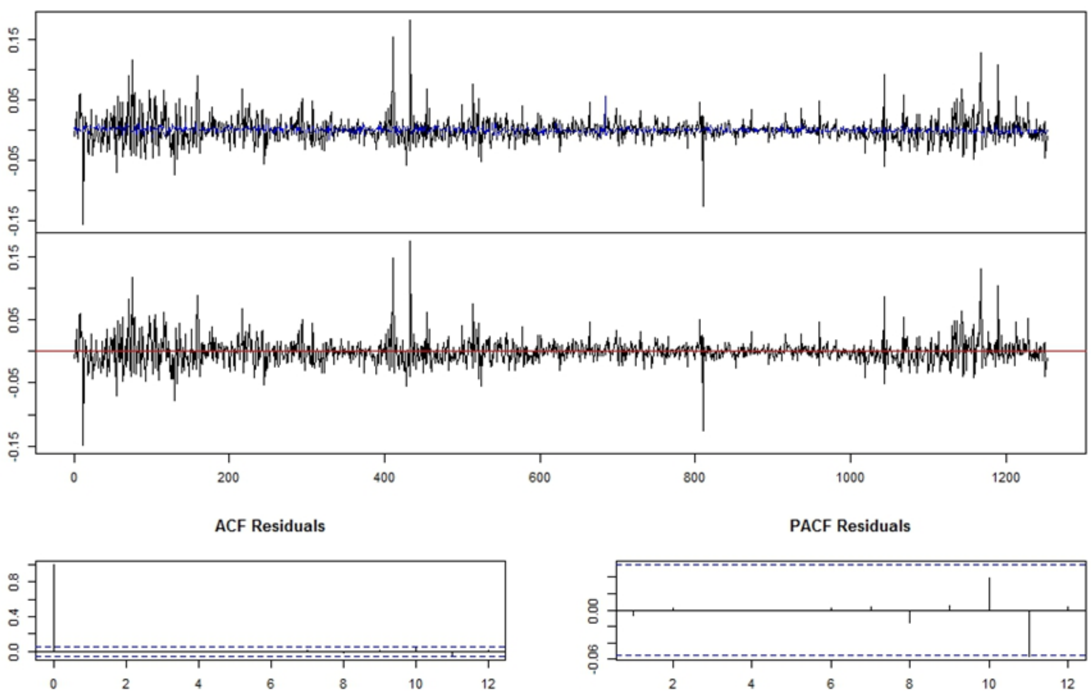
VAR estimation with *RET1*, *WtF*, and *dStDev* as endogenous variables, with respect to WtF. 1260 observations. Log likelihood: 5550.364. Residual standard error: 0.5648 on 1226 degrees of freedom. Multiple R-squared: 0.9286, Adjusted R-squared: 0.9271. F-statistic: 613.6 on 26 and 1226 DF, p-value: <2.2e-16. '***' '**' '*' '.' Denote significance at the 0.1%, 1%, 5% and 10% level.

*Appendix 9-Table: VAR WtF p=2*

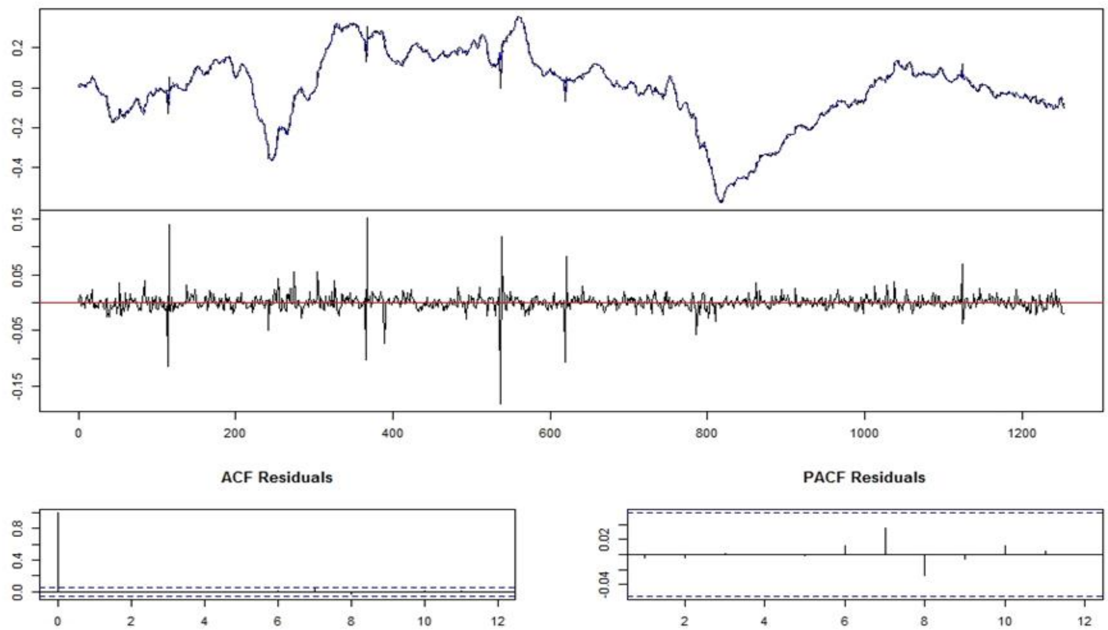
| | Estimate | Std. Error | t-value | p-value | |
|---|---|---|---|---|---|
| *RET1.l1* | 9.820e-01 | 7.990e-01 | 1.229 | 0.2193 | |
| *RET1.l2* | 5.396e-01 | 7.956e-01 | 0.678 | 0.4977 | |
| *dStDev.l1* | -9.495e-01 | 1.066e+00 | -0.891 | 0.3732 | |
| *dStDev.l2* | 7.922e-01 | 1.059e+00 | 0.748 | 0.4544 | |
| *WtF.l1* | 1.405e+00 | 2.518e-02 | 55.805 | <2e-16 | *** |
| *WtF.l2* | -4.979e-01 | 2.569e-02 | -19.383 | <2e-16 | *** |
| *const* | 8.560e-02 | 3.636e-02 | 2.354 | 0.0187 | * |
| *trend* | -1.091e-04 | 5.004e-05 | -2.181 | 0.0294 | * |
| *retBOil* | 9.739e-01 | 1.188e+00 | 0.820 | 0.4124 | |
| *retTBill* | -9.690e+00 | 5.069e+00 | -1.912 | 0.0562 | . |
| *retSP* | 1.659e-01 | 1.790e+00 | 0.093 | 0.9262 | |
| *sqLnDem* | -4.026e+00 | 1.619e+01 | -0.249 | 0.8037 | |

VAR estimation with *RET1*, *WtF*, and *dStDev* as endogenous variables, with respect to *WtF*. 1260 observations. Log likelihood: 5456.175, Multiple R-squared: 0.9194, Adjusted R-squared: 0.9187. F-statistic: 1293 on 11 and 1246 DF, p-value: <2.2e-16.
'***' '**' '*' '.' Denote significance at the 0.1%, 1%, 5% and 10% level.

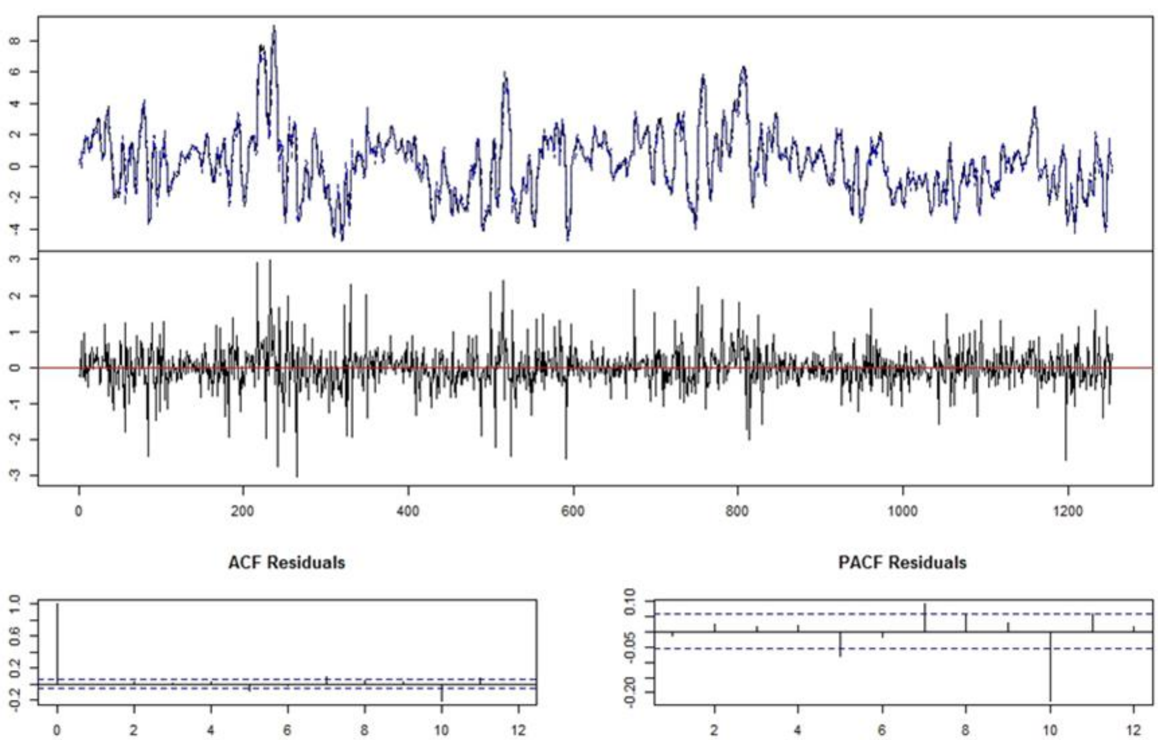
*Appendix 10-Figure: VAR Graph of fit and residuals for RET1*



*Appendix 11-Figure: VAR Graph of fit and residuals for dStDev*




*Appendix 12-Figure: VAR Graph of fit and residuals for WtF*



*Appendix 13-Figure: Degree-Day*