Universitetet i Stavanger

DET TEKNISK-NATURVITENSKAPELIGE FAKULTET

Studieprogram/spesialisering: Industriell Økonomi	Vårsemesteret, 2018
Forfatter: Kristian Gaard Magne Opstad Mellemstrand	…………………………………………
Fagansvarlig: Atle Øglend Veileder: Atle Øglend	
Title: Relationship between futures and spot market prices – an empirical study of the U.S Natural gas market before and after the Shale gas revolution	
Studiepoeng: 30	
Emneord: • Efficient market hypothesis • Johansen cointegration test • Shale gas revolution • Futures contracts	Sidetall: 51 + vedlegg/annet: 9 Stavanger, 24.05.18

Forside for masteroppgave

Det teknisk naturvitenskapelige fakultet

Abstract

In 2009, a massive increase in production of natural gas started due to new technology. Based on this increase in production, this study examines the efficient market hypothesis in the US natural gas market for the period Jan 1997-Dec 2017 for futures contracts with 1-, 2-, 3-, and 4-month to maturity. 2009 is the point of separation for the two periods analysed and compared in this thesis. Jan 1997- Dec 2008 is defined as before shale gas revolution, Jan 2009- Dec 2017 is defined as after shale gas revolution. The efficient market hypothesis is tested by using Johansen cointegration test and by imposing restrictions on α and β. An efficient market is in this thesis defined as significant cointegration while at the same time restrictions on $\alpha=0$, and $\beta=1$ cannot be rejected. If the efficient market hypothesis holds, the futures contract is an unbiased estimator of future spot price.

The results from this thesis shows a change in the US natural gas futures market. US natural gas prices suffered a decline after the shale gas revolution and the volatility decreased, indicating that predicting prices was more difficult before 2009. Before the shale gas revolution, the market was not efficient for any of the contracts. After the shale gas revolution, the 1-month contract was efficient while the contracts with 2-, 3-, and 4-month to maturity was not efficient. The 1-month contract after the shale gas revolution has neither over- or underestimated the spot price, but for every other contract the futures price has overestimated spot price. As the futures price has been overestimating the spot price it is fair to assume that bias occurs due to a risk premium of going short.

Preface

This thesis has been done in cooperation by Kristian Gaard and Magne Opstad Mellemstrand and the workload has been shared equally (50/50). This thesis concludes our master's degree in industrial economics at the University of Stavanger. We both have a bachelor's degree in petroleum technology from the University of Stavanger and the combination of courses from the master program founded our interest in the chosen topic.

Working with this thesis has let us apply our acquired knowledge from several courses during our study. In addition, we have through collaboration gained valuable experience in teamwork and we have felt an increasing sense of achievement throughout this period.

We would like to thank our supervisor Atle Øglend for all the good advices during the writing process and for always keeping his door open.

Kristian Gaard and Magne Opstad Mellemstrand
24.05.2018.

Table of contents

Table of figures

Nomenclature

- ADF – Augmented Dickey-Fuller
- AR(1) – Autoregression model of order one
- ARCH – Autoregressive conditional heteroskedasticity
- BG – Breusch-Godfrey
- Btu – British thermal unit
- EMH – Efficient Market Hypothesis
- EIA – U.S. Energy Information Administration
- JB – Jarque-Bera
- LM – Lagrange Multiplier
- MMcf – Millions Cubic Feet
- OECD – Organization for Economic Cooperation and Development
- TS – Time Series
- TSD – Time Series Data
- HPTS – Highly Persistent Time Series

1 Introduction

The natural gas market has been subject to major changes in recent years. The US natural gas market has been affected by several factors, such as economic crisis, increasing demand, and a new worldwide environmental mindset. Put these factors aside, there are maybe an even more crucial factor contributing to these changes, which is the rapid change in technology. The introduction of "fracking" made vast amount of natural gas accessible, and further contributed to changes in correlated markets. One of them, the futures market, is one that may be subject to change. An interesting aspect of futures market is the price discovery role of futures contracts. As futures contracts incorporates different actor's expectations for future prices, it can be an effective mechanism to predict future spot price. If futures contracts have a price discovery role it can be used in estimation of cash flow in projects where commodities play a vital role. However, if futures contracts do not play a significant role in price discovery it should not be used in planning future activities. In addition, the risk managers who's main concerns is to seek the optimal hedging solution in securing the commodity may have to make use of optional hedging alternatives (Switzer & El-Khoury, 2007). Then, the knowledge of the futures market's position as a predicator is of great interest for many shareholders. There are many indications that the market has changed, such as a structural break discovered by Oglend, Lindback, and Osmundsen (Oglend, Lindback, & Osmundsen, 2015) and hedging by futures were improved shown by Switzer and El-Khoury (2007).

The main concern of this study is the effect and impact shale gas revolution and the introduction of modern technology have had on the market. New technology could increase the uncertainty in the market related to commodity prices. This could make forecasting more difficult. The basis for this thesis is the discovery of a structural break found by Oglend et al. (2015). The structural break was described as a decreasing cointegrated relationship between natural gas and oil post 2009. The futures contracts being an unbiased estimator of the future spot price, has previously worked as a predictor, and therefore these contracts are vital tools for both hedgers and speculators. Being able to predict future spot price may provide possibilities for profit and secure owns interests.

One way to measure if the futures contracts are an unbiased estimator of the spot price is to test for market efficiency. If the market is considered efficient and certain restrictions hold, then one may say that the futures contracts are indeed an unbiased estimator of the future spot price. This thesis will provide a precise definition of what an efficient market is, as there is disagreement among experts. The definition given in this thesis are considered sufficient for the statement of problem. To be able to address the efficiency problem, this thesis makes use of an established methodology for how to handle the data, and what analysis to apply.

The periods already established by Oglend et al. (2015), are defined as before shale gas revolution, and after shale gas revolution. Even if the technological breakthrough is dated to 2003, the structural break divides the periods between 2008 and 2009. This means that all data as of Jan 1997 through Dec 2008 are defined to be before shale gas revolution, and all data as of Jan 2009 through Dec 2017 are defined as after shale gas revolution. This limitation provides a simplified way to achieve meaningful results from the tests performed.

By this, the statement of problem was derived, and all the analyses and tests are performed with the purpose to resolve this issue. The statement of problem is defined as:

Test for market efficiency for US natural gas before and after shale gas revolution, to determine whether the futures contracts are an unbiased estimator of future spot price.

The thesis is structured in following way. Chapter 2 provides background of US natural gas, chapter 3 provides the general theory, definitions and explanations of the market to be addressed. Chapter 4 explains the methodology, chapter 5 is about the data used in this thesis. Chapter 6 provides the results from the analyses and tests performed, chapter 0 discusses the results and finally, chapter 8 concludes.

2 Background

The U.S Energy Information Administration (EIA) published a report in 2017 (Administration, 2017), concerning the international energy outlook. In the report, the projection for world energy consumption is estimated to increase by 28% between 2015 and 2040. Figure 1 illustrates the outlook of the energy consumption presented by EIA. The forecast displays an increasing demand for energy. The non-OECD[1] countries is estimated to have the highest increase of demand, as the population rapidly increase. In addition, the economic growth in these countries is projected to be greater than in OECD countries. The economic growth provides more access to the energy market, which ultimately increases the demand. The industrial sector, by far, is the largest participant having projections above 50% of the total energy consumption. Nevertheless, the transport and buildings sectors have higher growth rate. Because of this, the assumption of increasing energy demand should not be overlooked.

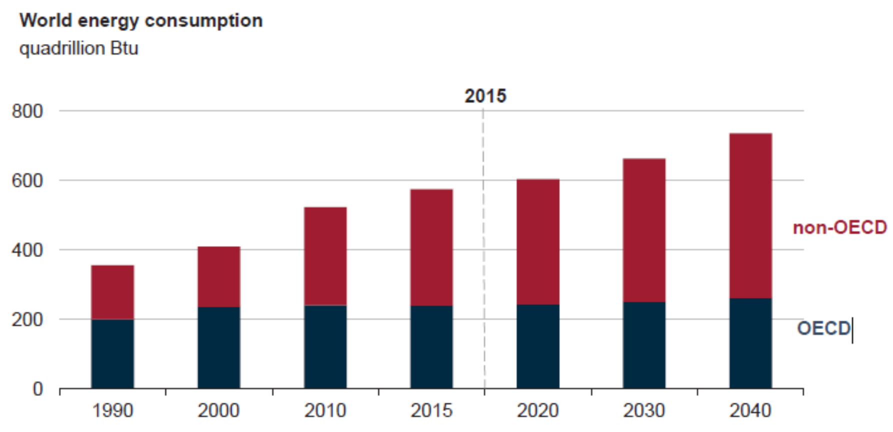

Figure 1: Outlook of world consumption. Downloaded from International Energy outlook 2017 (Administration, 2017). [Date: 12.03.2018]

The use of resources providing the energy, this includes all fuels, renewable resources and nuclear power increase (except coal) (EIA (2017)). Because of environmental focus, it is intuitive that the renewables including nuclear power are the fastest growing energy resources. However, petroleum liquids will be the dominant source of power in the current period, with natural gas as the fastest growing fossil fuel. This is based on the forecast where natural gas increases in all sectors, such as buildings, electric power, industrial and transport. Arguments that supports natural

[1] OECD countries in the journal are defined as Organization for Economic Cooperation and Development. The journal lists all the countries involved in the organization. The non-OECD countries are mainly countries geographically located east of Europe and includes China and India. The completed list is presented in “International Energy Outlook 2017”.

gas as an energy resource are low capital costs for new power plants, favourable heat-rates (which increase the utility of the resource), and low fuel cost. According to Gebre-Mariam, the North American continent consumes over 30% of worldwide natural gas consumption. Annually consumption of natural gas in the US covers almost a quarter of the total energy consumption. 300 000 miles transmission lines and 1,2 million miles of pipeline infrastructure creates a comprehensive network distributing natural gas from production site to consumers (Gebre-Mariam, 2011). By this, natural gas is a vital part and contributor to the US economy and are therefore an interesting subject to investigate.

The motivation for this thesis is based on previous paragraph, where the utility of natural gas obviously is increasing worldwide, and especially in the US. Based on projections from EIA (2017), shale gas has been, and still is, a vital part of the gas market. Figure 2 illustrates projections indicating that shale gas in the US will dominate the natural gas market.

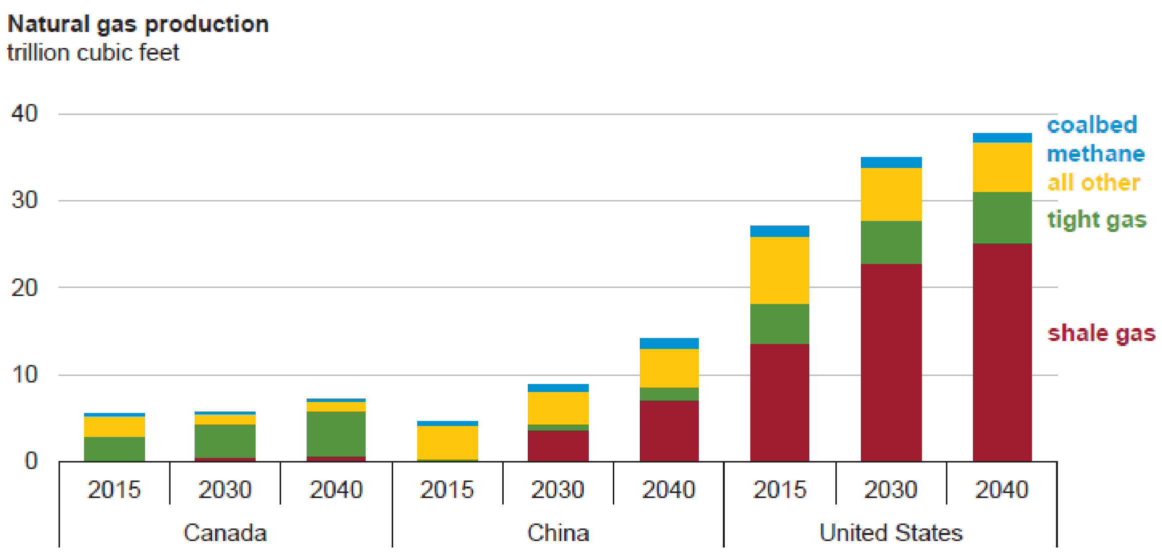

Figure 2: Projections of several types of natural gas. Downloaded from International Energy Outlook 2017. [Date: 12.03.2018]

2.1.1 Recovery of natural gas and shale gas revolution

Natural gas resources are classified according to where it is accumulated and produced from. Gas migrates from a source rock into a reservoir where it accumulates due to a trap. Reservoirs differ greatly, with different physical variations affecting the performance and recovery. Conventional reservoirs are relatively easy to produce from as the pressurized gas will flow through the well to the surface once a conduit is opened.

Unconventional reservoirs have been more difficult to produce from. Unconventional gas is found in “tight” sandstones, absorbed into the matrix of shales and in coal seams. It is the gas found in the matrix of shales that lead to the shale gas revolution. Before 2009 it was not economically profitable to produce from such reservoirs. To be able to produce oil and gas from a low permeable rock such as shale, pathways need to be created. A common method of doing so is hydraulic fracturing, also called “fracking”. The breakthrough could be traced back to 1990’s, because of George P. Michell’s work. The combination of hydraulic fracturing and horizontal drilling introduced in 2003 by Devon Energy provided a substantially reduced production cost, concerning the extraction of shale gas and shale oil trapped in the formation (Oglend et al. (2015)).

A simple explanation of fracking is that fracturing fluid is pumped into the well to create a pressure high enough to fracture the rock. The fluid contains ceramic beads to keep the fractures open when the pressure declines, allowing oil and gas to be produced (Regulator, 2018). Figure 3 illustrates the increasing amount of extracted natural gas in the US. The graph shows changes in production as of 2005, due to the breakthrough of horizontal fracking in 2003. As a result, U.S. has changed from importing natural gas to satisfy the demand, to be self-sufficient and exporting.

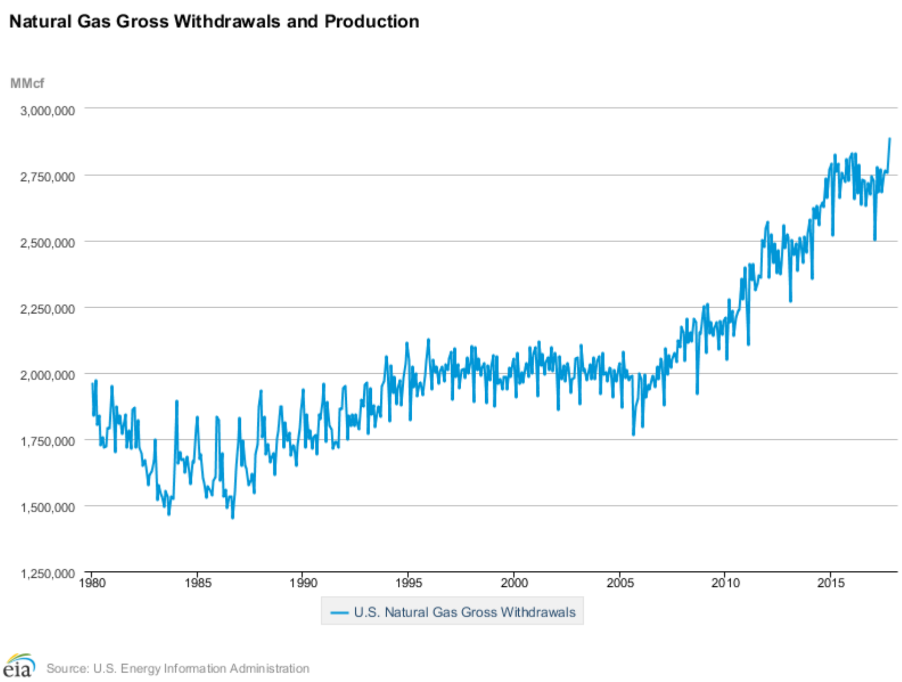

Figure 3: Gross withdrawals and Production of Natural gas in U.S. Downloaded from [*https://www.eia.gov/dnav/ng/ng_prod_sum_a_EPG0_FGW_mmcf_m.htm*](https://www.eia.gov/dnav/ng/ng_prod_sum_a_EPG0_FGW_mmcf_m.htm) *[Date: 12.03.2018]*

.

Figure 4 shows different types of reservoirs. A conventional reservoir can be seen in the top left corner. As shown, the gas is trapped in an anticline due to a seal rock. To the far right is another type of conventional reservoir, with both oil and gas present. In the middle of the figure there is the example of a gas-rich shale and a relatively horizontal well. When observing the figure and the well in the gas-rich shale, one may understand why this method became so profitable. With a horizontal well you get more contact with the reservoir than with a vertical well, which means that less number of wells are needed.

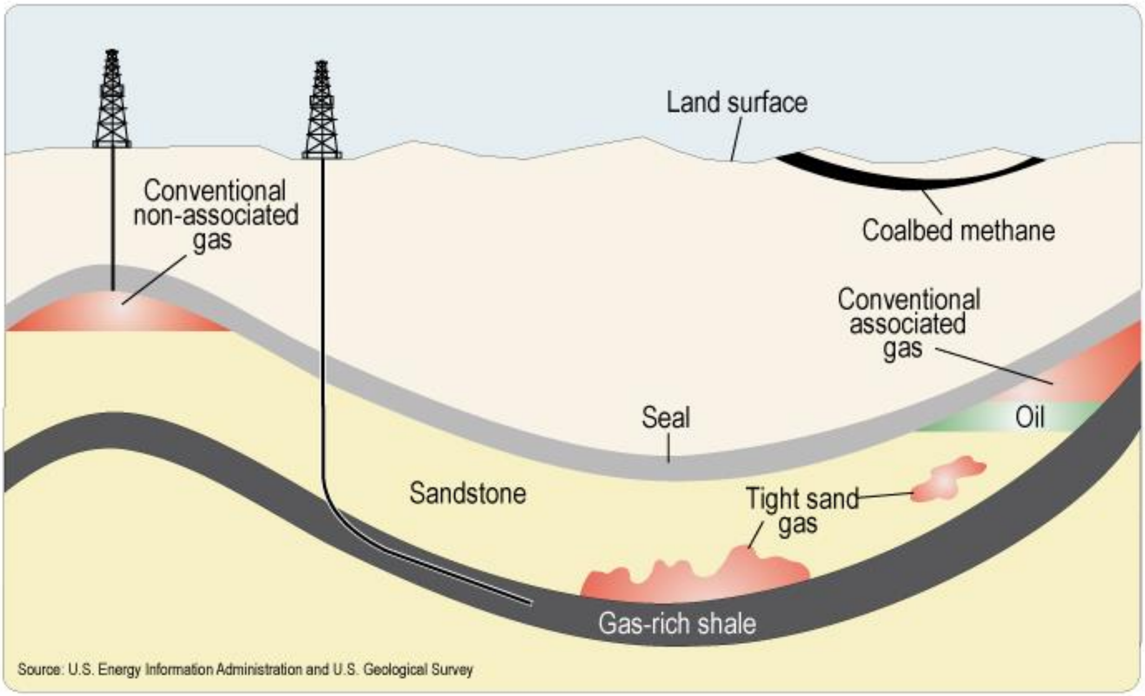

Figure 4: Overview of conventional and unconventional gas reservoirs. Downloaded from https://www.eia.gov/todayinenergy/detail.php?id=110. [Date: 16.01.2018].

2.1.2 Factors affecting the natural gas market

Indeed, the impact of shale gas production provided a significant point in time, where the natural gas market changed. Pricing of natural gas became a hot issue in the field of energy economics after the shale gas revolution. As price in general is the most important market signal and affects investments and demand, it is important to understand the driving mechanisms of natural gas prices. The U.S natural gas market is deregulated and market-oriented and thus, the determinants of natural gas prices are complex and diversified (Ji, Zhang, & Geng, 2018). As mentioned by Ji et al. (2018), factors include macroeconomic situations, speculative activities, seasonality and substitution effects. The research on driving forces of natural gas prices is thin compared to that of oil, but major supply- and demand factors are listed and commented below.

Three major supply-side factors that affect prices are:

- Amount of natural gas production
- Level of natural gas in storage
- Volumes of natural gas imports and exports

And three major demand-side factors are:

- Variations in winter and summer weather
- Level of economic growth
- Availability and prices of competing fuels

2.1.2.1 Amount of natural gas production

In the U.S most of the natural gas consumption comes from domestic production. Dry natural gas production increased from 2005 to 2016 where spot- and consumer prices generally decreased in the same period (Administration, 2017).

2.1.2.2 Level of natural gas in storage

Natural gas storage plays a key role in meeting seasonal peak demand and has a considerable influence in supply and, hence prices. Theses and articles have been written on how storage affect prices, but generally it reduces the volatility of prices. When demand is lower, storage absorbs excess production and when higher, storage helps to keep up the supply. On a general basis, storage has contributed and been a major factor for economic crisis. An example is the crises in 2014, where the market had a surplus of petroleum on the market (Marvik, 2016), which created unbalance between supply and demand.

2.1.2.3 Variations in winter and summer weather

Weather is a factor affecting the demand side of prices, but it can also affect the supply side. As an example, hurricanes along the U.S Gulf Coast lead to a 4% decline in production. Due to natural gas being used for heating, demand increases during cold months. A sudden cold may affect the prices if the supply is not able to meet demand.

Hot summers affects prices both directly and indirectly. High temperature in the summer leads to higher usage of air conditioning which leads to an increase in the power sectors demand for natural gas. Due to the increased demand in the summer, lower levels of natural gas are stored for the winter. This can result in higher prices because of the above mention problem that supply is not able to meet demand.

2.1.2.4 Availability and prices of competing fuels

Large-volume factories or electricity producers that easily can switch between resources may cause fluctuations in gas prices. As a competing source to natural gas, lower coal prices relative to natural gas can make the consumer switch to coal and reduce the demand for natural gas.

2.1.2.5 Summary

This concludes the section of factors affecting natural gas prices. The factors mentioned above are the major factors listed on EIA and these factors all have different impact on natural gas prices. Although production is a key factor, it can be observed from Figure 5 that prices are affected by other influential factors as well. The findings of Ji et al. (2018) have three aspects. First, there is a contemporaneous causal flow from crude oil to natural gas, although it has weakened after the financial crisis. Second, storage and seasonality have a causal relationship on natural gas, which emphasizes weather and inventories as important drivers on natural gas prices. Third, speculation activities have trivial influence on natural gas prices and indicates that financialization has not been a strong driver of natural gas prices.

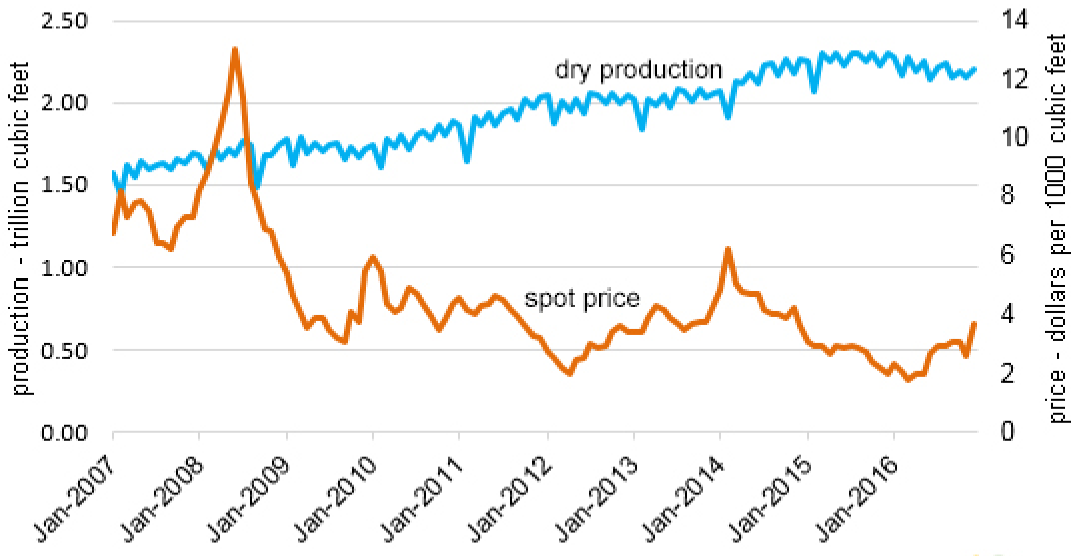

Figure 5: Production of dry natural gas and spot prices. Downloaded from https://www.eia.gov/energyexplained/index.cfm?page=natural_gas_factors_affecting_prices. [Date: 18.01.2018]

3 Theory

In this section, spot market and futures market and the relationship between the two markets will be explained, before presenting market efficiency, which is the foundation that this thesis will be based on. A review of former studies on market efficiency is presented before moving on to the methodology section.

3.1 Commodities

A commodity is a raw material or primary agricultural product that can be bought or sold, such as copper or coffee. Because storability of commodities plays an important role on pricing, they are often classified according to the grade of storability and availability. A commodity is said to be storable if it is non-perishable and have low cost of storage relative to the total cost. Examples of storable commodities are metals, crude oil and natural gas. Non-storable commodities will perish if they are stored over a longer period and prices tend to be more cyclical. Example of non-storable commodities are livestock (Fabozzi, Füss, & Kaiser, 2008).

3.2 Spot market

Spot market or “cash market” is a marketplace for immediate settlement of transactions between buyer and seller. Commodities are bought or sold on the spot at current spot prices for delivery now or in the near future (Investopedia, 2018c).

Spot price is the price of a commodity at the spot market. It reflects the price of a certain amount of a commodity with a certain quality at a specific location. Commodities may have different qualities according to location and producer. For example, the amount of protein in wheat from one producer may differ from another. This difference in quality may result in local differences in spot price relative to market price (Tomek & Kaiser, 2014). This results in the following definition of spot price, where ε_t represent difference in quality;

$$\widetilde{S_t} = S_t + \varepsilon_t \qquad 3.1$$

3.3 Futures market

A futures market is a market where participants buy or sell commodities for delivery at a specified date and location in the future. The futures market provides an opportunity for producers, consumers and inventory-holders of commodities to reduce the risk of rapid price movements. In this market it is not the commodity itself that is bought or sold, but futures contracts. Futures contracts are a legal obligation between participants to either make or accept delivery of a specific amount of a commodity during a specific month at a price established in the market (Tomek & Kaiser, 2014, pp. 246-254). The contracts are standardized and traded on their own exchanges, for example New York Mercantile Exchange. Without going into too much detail about futures contracts trading, it should be noted that in most cases traders do not intend to make or take delivery. Before contract maturity, the initial position is offset by taking the opposite position, but it is the possibility of delivery that makes the futures price converge to spot price.

There are essentially two types of participants in a futures market, hedgers and speculators. Hedgers use futures contracts to shift the price risk. For example, a producer of grain sells futures contracts for delivery in the future. If the spot price declines, the loss in the spot market are offset by the gain in the futures market. The reason being that spot price and future price are positively correlated. Futures contracts are also used by firms buying commodities. If they know they will have to buy grain in the future, they purchase futures contracts to secure themselves against a price increase.

In a market there is always someone trying to make a profit. These are called speculators and they take positions in futures contracts with the purpose of making a profit on price changes. There are different types of speculators that will not be explained, but speculators have an important function in the market as they provide market liquidity and contribute to price determination (Tomek & Kaiser, 2014, pp. 251-254).

3.3.1 Futures price and price discovery

The futures price reflects the price of the contract for a specific amount of a commodity. There are many buyers and sellers in the futures market, dealing on a homogenous product, and therefore come closely to a perfect competitive market. A lot of public information exists about the U.S commodity markets and are equally available to all traders. It is this information that futures prices are based on. In other words; the current futures price, F_t is the expected price at maturity T, based on current information, I_t, and can be written as equation *3.2*. Because futures price at maturity, in theory, equals spot price, equation *3.2* can be written as in equation *3.3*, where S_T is spot price (Tomek & Kaiser, 2014, pp. 255-256).

$$F_t = E[F_T|I_t] \quad 3.2$$

$$F_t = E[S_T|I_t] \quad 3.3$$

According to the above equations, futures price are potentially unbiased forecasts of future spot prices and thus, plays an important role in the price discovery process. As futures prices should be based on all available current information, futures prices can be used as a reference for establishing future spot prices. However, due to the many different participants in a futures market with a variety of motives, speculation can have detrimental effects on price discovery. There have been questions to whether futures market has beneficial or adverse effects on spot prices. One example, as mentioned in Tomek et al. (2014), the spike in petroleum prices during the Gulf War was alleged to be a result of speculators on NYMEX. Although the logical reasoning was that futures reflected the economic conditions. However, the allegations implied that without futures market, spot prices would have been less volatile. Another interesting thought with respect to price discovery is that futures trading attracts more traders to the market and hence, the amount of available information in the market has increased. Thus, existence of futures market leads to a more precisely determination of the systematic component of price behaviour. This thesis will apply the Granger causality test to show the information flow between spot- and futures market, i.e. which market serves as a price discovery centre.

3.4 Relationship between spot price and futures price

There are two popular views relating spot prices and futures prices. The first and non-controversial view is the *theory of storage (Brennan, 1958; Kaldor, 1939; Working, 1948).* Key concepts in theory of storage are cost of carry, expectations and basis. The second view splits futures prices into a risk premium and the expected future spot price. The existence of risk premium is a more controversial topic and is explained more in section 3.4.2.

The Theory of Storage provides a fundamental pricing condition and the necessary link between futures price and spot price:

$$E_t(S_T) - S_t = m_{t|T} \tag{3.4}$$

$$F_{t|T} - S_t = m_{t|T} \tag{3.5}$$

$$m_{t|T} = r_{t|T}S_t + (1 + r_{t|T})m(I_t) \tag{3.6}$$

Where $m_{t|T}$ is the total cost of storing one unit of a commodity from t to T. For a commodity, cost of carry involves cost of not selling today, financing, storing and insuring the commodity. Futures price must therefore equal the expected spot price plus these costs. Due to these costs, as maturity approaches, the difference between spot- and futures price converges. If this correlation does not hold, there would be possibility of arbitrage and in an efficient market, everyone would exploit this possibility and earn a profit. This would make the arbitrage short-lived due to the market balancing itself. Arbitrage can be explained by following two examples (Banks, 2005):

If $F_T > S_T$

1. Buy the commodity for S_T
2. Sell the futures contract and earn a risk less profit, $F_T - S_T$

If $F_T < S_T$

1. Buy futures contract for F_T
2. Accept delivery
3. Sell the commodity for S_T and earn a risk less profit, $S_T - F_T$

Another example of explaining arbitrage is by the cost-of-carry relationship. A producer holding inventory wants to short[2] futures to reduce price risk. The producer would earn a profit of $F_{t|T}$ – $F_{T|T}$ at maturity. Assuming a futures price of $F_{t|T}$ = $112, a spot price of S_t=$100 and a cost of carry $m_{t|T}$ = $5. The producer's strategy is to short one futures contract and store one unit. From Theory of Storage and equation 3.2, $F_{T|T} = S_T$ and $F_{t|T} = E(F_{T|T})$ is derived.

If S_T = $110 and, the producer earns $110 - $100 - $5 = $5 on the storage position and $112 - $110 =$2 on the short position. Resulting in an economic profit of $7. If cost of carry is now at time *t*, this is a risk less profit.

If S_T = $102, the producer loses $102 - $100 - $5 = -$3 on the storage position and earns $112 - $102 - $5 = $10 on the short position, resulting in an economic profit of $7. Meaning that, if cost of carry is known, there would be a possibility of a risk less profit. Due to such arbitrage opportunities, it must be that $F_{t|T} – S_T = m_{t|T}$.

3.4.1 Basis

The difference between futures price and spot price is called *basis*. It is also known as the cost-of-storage and is defined as in equation *3.7*. Following this definition, a positive basis means a positive incentive to store more and vice versa (Tomek & Kaiser, 2014, pp. 256-265). A market may be classified as being in *contango*[3] or *backwardation*[4], depending on positive or negative basis. Contango typically occurs when investors prefer to pay premium to acquire the commodity in the future rather than acquiring the commodity right away and pay for storage and cost of carry. As storing fossil fuel can be costly, the natural gas market historically has for the most been in contango.

$$B_t = F_{t|T} - S_t \qquad 3.7$$

[2] Shorting means that the producer earn if the future price falls relative to the futures price contracted at time t.
[3] A market where futures price is higher than spot price is classified as being in contango. Spot prices are expected to increase(Investopedia, 2018b).
[4] A market where futures price is below spot price is classified as being in backwardation. Spot prices are expected to decrease (Investopedia, 2018a).

As Tomek and Kaiser (2014) points out, there exists different basis at any point in time. One reason is, as explained in previous section, due to spot price varying by location and quality. It also varies with the length of the contract.

3.4.2 Risk premium

Speculators in a futures market profit from price fluctuations. If a speculator buys a futures contract at time t and sells at time T, profit at time T, equals $F_T – F_t$. As futures price at time t is the expected futures price at time T, defined by equation *3.2*, the assumption is that profit from buying and selling futures contract equals zero.

When producers of commodities seek to reduce price risk, i.e., selling futures contracts, there needs to be someone taking on the risk of buying the contracts. Speculators are willing to take on this risk, but as the expected profit equals zero they need compensation in form of a risk premium (Tomek & Kaiser, 2014, pp. 291-294). Meaning that, for speculators to be willing to take on the risk, they need a positive risk premium. Thus, futures price F_t is set to be lower than the expected futures price $E[F_T]$. Risk premium is defined by $E[F_T] – F_t = r_{T-t}$. This difference is also defined as biasedness of futures prices as forecast for future spot price and dates to Keynes (1930) and Kaldor (1939) as referred in Tomek & Kaiser (2014). Risk premium might also be negative if there are a lot of consumers seeking to secure their position in the market, and speculators are the ones selling contracts. Thus, the magnitude of risk premium depends on whether the pressure comes from buyers or sellers.

Whether or not there actually exists a risk premium in futures markets is a controversial topic and have been studied a lot in the literature. The results from studies on risk premiums in futures markets are mixed. Alquist and Kilian show that futures-based forecasts are biased in the crude oil market (Alquist & Kilian, 2010), while Chinn, LeBlanc, and Coibon find natural gas futures prices to be unbiased predictors of future spot prices, except for 3-month contracts (Chinn, LeBlanc, & Coibion, 2005). Considine and Larson studied risk premium in natural gas and crude oil and found the existence of risk premium and that risk premium rise sharply with higher volatility (Considine & Larson, 2001). They suggested that these findings helped explaining why spot prices often exceed futures prices. Although this thesis will not examine risk premium explicitly, it should be kept in mind that it might be a reason for biasedness of futures-based forecasts.

3.5 Literature review

The efficient market hypothesis is one of the most studied topics in economic literature. Although the literature on efficient market hypothesis related to natural gas is thin, it has been studied and the results are mixed. This section provides a review of literature related to this thesis.

Former studies show that there is no consensus among experts on the definition of efficient market, and further no consensus in the methods applied testing it. Moosa and Al-Loughani addressed this problem in their journal article (Moosa & Al-Loughani, 1994). Many of the studies performed does not provide a precise definition of market efficiency, the methods applied in testing are not used for the right purpose, the testing of restrictions on coefficients are absent, and there are almost no explanation of irrationality or risk premia presence when the hypothesis is rejected. Moosa and Al-Loughani (1994) tested for market efficiency of crude oil futures. They concluded that futures prices are not an unbiased forecaster of future spot price and there is a time varying risk premium.

A similar study of unbiasedness hypothesis testing of natural gas was performed by Movassagh and Modjtahedi, where the unbiasedness hypothesis was rejected, and natural gas futures were considered a biased estimator of future spot price (Modjtahedi & Movassagh, 2005). A similar conclusion is provided by Mishra and Smyth, by implementing another methodology in their studies. The methodology applied is of Mean Square Prediction Error (MSPE) and Mean Absolute Prediction Error (MAPE). The conclusion was that the futures prices provide information for predicting the direction of change of natural gas spot price. The main difference is that the futures does not predict the magnitude of the spot price any better than a random walk model (Mishra & Smyth, 2016). Another study displaying results rejecting the unbiasedness/efficiency hypothesis is the journal article by Lee and Lee. Allowing for structural break, the efficient market hypothesis was rejected for total energy prices, incorporating coal, oil, gas and electricity (Lee & Lee, 2009). Also, Wei and Zhu concluded that the futures price of natural gas was a biased predictor of future spot price (Chiou Wei & Zhu, 2006).

On the other hand, Switzer and El-Khoury (2007) found that crude oil futures prices are unbiased estimators of future spot prices. This is supported by the study of Walls (Walls, 1995) on natural gas futures predicting spot prices on numerous locations. Indicating that in general, the natural gas futures are in fact unbiased estimators of future spot price. The results from this paragraph and the previous paragraph illustrates conflicting results of the hypothesis, which makes this even more interesting to investigate.

Based on the different results presented above, some may argue that if a market is efficient, the futures are also an unbiased estimator. However, Dwyer and Wallace's results shows otherwise. Their research is concerning exchange and interest rates, and spot and forward exchange rates. They criticise the definition of efficiency provided by Fama, stating that *it is hard to see how a market with no expected utility increasing profit opportunities available to agents based on expected utility maximising acquisition of information could be characterized as inefficient in any sense of the word* (Dwyer & Wallace, 1992, p. 319). In addition, the research concluded that there is no necessary connection between market efficiency and forwards being unbiased predictors, but rather that the unbiasedness implies cointegration. This is similar to Malkiel's conclusion, where he states that the stock market is *far more efficient and far less predictable* (Malkiel, 2003) than other studies implies. In other words, this indicate that there can be an efficient market without and unbiased estimator.

Testing for market efficiency demands a structured methodology. There are several models and tests that can provide somewhat same information, and therefore it is appropriate to address the choices made in this thesis. The tests performed in this thesis are to a certain degree dependent on each other, where the results from a previous test could affect the results on the following test. Such as described by Lai and Lai (Lai & Lai, 1991) the cointegration test relies on a certain criteria making the results significant. Cointegration tests are in general test performed to test the long-run equilibrium between two variables.

3.6 Market efficiency

Market efficiency is essential to address the statement of problem for this thesis. There is a general disagreement among experts of the definition and function of market efficiency as Malkiel (2003) discusses. The evidence used to undermine the efficiency hypothesis, is presented by Malkiel as thin, and that some of the predictable patterns presented disappears, as a result of not being statistically large enough (Malkiel, 2003, pp. 61-63).

An efficient market is a market where the prices reflect all relevant available information. Fama divided the efficiency hypothesis into three forms varying by the information taken into consideration (Fama, 1991). The three forms of efficiency are weak, medium and strong. The weak-form efficient market hypothesis says that all past price movements are already reflected in current prices. Medium-form says that all publicly available information is reflected in current prices and strong-form contains all information, even private.

An easy way to define market efficiency, which is a widely applied definition is the definition of Malkiel and Fama (Malkiel & Fama, 1970, p. 387). The conditions for market efficiency are:

- *There are no transaction costs in trading securities*
- *All available information is costless available to all market participants*
- *All agree on the implications of current information for the current price and distributions of future prices of each security.*

Malkiel and Fama (1970) says that these conditions are sufficient but not necessary and despite the criticisms, it is a widely applied definition, and applied in this thesis. This definition shows that efficiency is related to the optimality of forecasting, and thus, a precise and good definition.

As mentioned, in an efficient market, the participants react to and exploit all available information. So, if price changes fluctuate at random and only new information can cause price movements, the information must be reflected in both futures prices and spot prices. Hence, futures prices and spot prices should be correlated or move in the same direction as they react to the same information. A frequently used model for examining futures market is presented in the work of Lai and Lai (Lai & Lai, 1991, p. 567):

$$S_t = \alpha + \beta F_{t-1|t} + \varepsilon_t \qquad 3.8$$

- S_t = spot price at time t
- α = constant coefficient
- β = constant coefficient
- $F_{t-1|t}$ = futures price at time t – 1, at maturity t
- ε_t = error term, with mean zero and finite variance

The definition above is applicable, and in addition, Lai and Lai (1991) states that under the hypothesis of market efficiency, there are no *strategy from which traders can profit consistently by speculating in the [...] futures market on future levels of the spot price* (Lai & Lai, 1991, p. 567). Assigning restrictions on the constant coefficients in equation *3.8* provides an opportunity testing for market efficiency. The restrictions are α=0 and β=1. Lai and Lai (1991) calls this for the unbiasedness hypothesis. This will be discussed in section 4.4.1. By running the Johansen cointegration test and applying restrictions on α and β, it is possible to detect if the market is efficient and hence if futures prices is an unbiased estimator of future spot price.

To summarize, this thesis uses the market efficiency definition of Malkiel and Fama (1970). The market is weak-form efficient if there are cointegration, and at the same time the joint hypothesis of α=0 and β=1 in equation 3.8 cannot be rejected.

4 Methodology

This section provides the tools and definitions used to address the problem. It includes critical aspects, which is of great importance for this thesis. The tests performed are presented in a chronological order of execution as some of the test are dependent on past result.

4.1 Time series data (TSD)

The underlying subject in this thesis incorporate the aspect of time. Obviously, this is a major factor, as price change through time. Time series data could be defined as several observations of a variable through time. Considering that a future outcome can be influenced by a past event, it is important to acknowledge time as an important dimension. Wooldridge claims in his work that a chronological arrangement of observed data may provide valuable information (Wooldridge, 2013).

Taking a closer look at what it means handling observations through time there are some central features that must be considered. In economics, assuming observed data is independent of its history is a strong assumption. Because of most time series are correlated to its historical data in various degree. There are different techniques where modifications can be incorporated into a model, handling the trends displayed in a time series. These modifications can exploit the dependencies and trends into an advantage.

Data frequency is another feature that needs attention. The observations of some variables have low volatility, and therefore it may be more appropriate using frequencies covering larger time periods. Vice versa for high volatility variable, it is more appropriate using a high frequency, collecting more observations for a period. It is intuitive, but according to Wooldridge (2013) the most common frequencies is daily, weekly, monthly, quarterly, and annually. As mentioned in the previous section, trends can be exploited, and these trends may be detected using different frequencies, such as seasonality.

4.2 Aspect of stationary and nonstationary TSD

There are two main groups for TSD. These are called stationary and non-stationary. There are certain criteria that defines a TSD as stationary. If the mean and the variance are constant over a time-period, it is characterized as stationary. In addition, the covariance is not dependent on observed time, but rather the length of time, separating the two values (Hill, Lim, & Griffiths, 2012, pp. 476-477). Hill et al. (2012) present these equations as an argument:

$$E(y_t) = \mu \tag{4.1}$$

$$Var(y_t) = \sigma^2 \tag{4.2}$$

$$Cov(y_t, y_{t+h}) = Cov(y_t, y_{t-h}) = \gamma_h \tag{4.3}$$

The equations 4.1, and 4.2, says that a TSD is stationary if the mean of y_t is constant, and the variance of y_t is constant. Equation 4.3, tells us that the for any t, h ≥ 1, the equation is only dependent on h, which in this case is referred to as the length of time. In other words, the covariance of the function does not depend on time, but rather the difference between the two terms (y_t and $y_{t\pm h}$). This means that the correlation between the two terms is dependent on h. As a comment, this is what Wooldridge (2013) defines as a covariance stationary process. The covariance stationary process also assumes that the correlation goes to zero as time passes, reducing the correlation over time.

A nonstationary TSD occurs when these criterions are violated. It can be illustrated in a graph as a line fluctuating away from a constant. Observing such a graph is not a test for stationary/nonstationary, and therefore is only an indication. When dealing with TSD it is critical to test for stationarity. Neglecting a test for stationary/nonstationary may result in false conclusion, reason being that two independent TSD with character of non-stationarity may give significant regression results. This means that using nonstationary data may spuriously indicate a significant relationship when there are none. This is called *spurious regression* (Hill et al., 2012, p. 482). How to test for stationarity and how to conduct regression analysis with nonstationary data will be explained in the following sections.

4.3 Dickey-Fuller (DF) test

The Dickey-Fuller (DF) test is a hypothesis test, performed to test whether a TSD has a unit root, and if so, is nonstationary. As a starter, one introduce a univariate model, also known as an autoregression model of order one, AR(1). This model does not include any explanatory variables.

$$y_t = \rho y_{t-1} + v_t, \quad |\rho| < 1 \qquad 4.4$$

Under the condition of $|\rho|<1$, the model is stationary, which implies that a proportion ρ of last period y_{t-1} plus an error term equals y at a given time t. The error term is independent, has zero mean and constant variance, as seen in section 4.2 (Hill et al., 2012, pp. 477-478). To fully understand the test and its purpose, it is included a section describing the basis of what is tested and the final test through definitions and explanations.

Hill et al. (2012) describes three different variations of the DF test, where there are only two terms being considered. There is a final test called the Augmented Dickey-Fuller (ADF), which is an expansion of the basic tests. The ADF accounts for autocorrelation, by including lags into the test model. Hill et al. (2012) states that *in practice, we always use the Augmented Dickey-Fuller test to ensure the errors are uncorrelated* (Hill et al., 2012, p. 486).

4.3.1 Basis for unit root and random walk.

The basis is considered through the following underlying definitions. Below is a list of definitions with further explanation:

- ***Highly persistent:*** *A time series process where outcomes in the distant future are highly correlated with current outcomes* (Wooldridge, 2013, p. 850)
- ***Unit root process:*** *A highly persistent time series process where the current value equals last period's value, plus a weakly dependent disturbance* (Wooldridge, 2013, p. 860)
- ***Random walk:*** *A time series process where next period's value is obtained as this period's value, plus an independent (or at least an uncorrelated) error term* (Wooldridge, 2013, p. 856)

Starting with highly persistent, consider a time series that has a higher degree of dependence. The problem with such time series, is the fact that the classical assumptions[5] is more frequently exposed to be violated, than other data sets. Further, as seen from the definitions, unit roots process seems very similar to random walk. Actually, the random walk is a special case of unit root process. The principle is the same for both of them, as they both has a high correlation between the value of a variable today and the future value of the same variable. The mathematical derivation, explaining the difference between random walk and unit root are considered to be excessive for this thesis. But the main difference is that the variance of a random walk process increases as a linear function of time, where the expected value of the random walk is independent of time (Wooldridge, 2013, p. 392).

For the unit root process, which is a more general case of a HPTS, there is a weakly dependent error term. The reason for this is that the more general case consider a weak correlation between the explained variable, and past values of the explained variable (Wooldridge, 2013, p. 393).

Handling HPTS could provide wrong conclusions, as HPTS and trends is somewhat different and could be misinterpreted. Examples as interest rates, is considered to be HPTS, but they have no typicall trend. On the other hand, as Wooldridge states, a HPTS often contains a clear trend. There are several models of this, such as *random walk with drift* (Wooldridge, 2013, p. 394). The drift, ultimately creating the trend, is added as a constant in the equation.

4.3.2 Augmented Dickey-Fuller (ADF)

Hill et al. (2012) walks through the basic behind the ADF by dividing steps of the test into different sections. Through this section the basics of the DF are integrated into the ADF, for the sake of simplicity.

There are different terms that are used in the test model, where it is possible to view the data from different angles. As mentioned, there are three variations of the test models. These are listed below:

- No constant, and no trend
- With constant, and no trend

[5] The classical assumptions refer to the ordinary least squares (OLS) assumption in simple and multiple linear regression for time series.

- With constant, and with trend

The test is performed on an AR(1) model, such as equation 4.4. Hill et al. (2012) rewrites the model to a more convenient form, which is also done in the final ADF model. The data is considered stationary for any |ρ|<1, and nonstationary for ρ=1. To determine the state of the TSD, the model is tested with a one-tail hypothesis, where the initial assumption is that the TSD is nonstationary. If H_0 is rejected, the TSD is stationary, and if H_0 cannot be rejected, it is nonstationary (Hill et al., 2012, p. 484). The hypothesis tested is shown in equations 4.5 and 4.6. The hypothesises are equal for all ADF tests.

$$H_0: \rho = 1 \Leftrightarrow \; H_1: \gamma = 0 \Leftrightarrow \; \tau > \tau_c \qquad 4.5$$

$$H_1: |\rho| < 1 \; \Leftrightarrow H_1: \gamma < 0 \Leftrightarrow \tau \leq \tau_c \qquad 4.6$$

The standard model for ADF test is shown in equation 4.7[6]. The nomenclature is listed below. The model addresses a TSD with a constant term, fraction of last period's value, the sum of lags and an error term. By using variants of this model, the data can be tested with and without constant term, and with and without a trend term. See equations 4.8,

4.*9*, and

4.*10*.

$$\Delta y_t = \alpha + \gamma y_{t-1} + \textstyle\sum_{s=1}^{m} a_s \, \Delta y_{t-s} + v_t \qquad 4.7$$

Where,

α = constant term (intersect)

$\Delta y_t = y_t - y_{t-1}$

$\gamma = \rho - 1$, and ρ is the proportion of last period's value y_{t-1}

a_s = estimated lag coefficients

$\Delta y_{t-1} = (y_{t-1} - y_{t-2})$, $\Delta y_{t-2} = (y_{t-1} - y_{t-2})$, ….. , Δy_{t-s}

v_t = error term

λ = trend

[6] The terms α, γ, and λ are rewritten terms, from deriving the model. In this thesis this is considered comprehensive and has neglected the derivation. In addition, the term ρ is included, as this was the basic in the AR(1) model. For further derivation see Hill et al. (Hill et al., 2012).

$$\Delta y_t = \gamma y_{t-1} + \sum_{s=1}^{m} a_s \, \Delta y_{t-s} + v_t \qquad 4.8$$

$$\Delta y_t = \alpha + \gamma y_{t-1} + \sum_{s=1}^{m} a_s \, \Delta y_{t-s} + v_t \qquad 4.9$$

$$\Delta y_t = \alpha + \lambda t + \gamma y_{t-1} + \sum_{s=1}^{m} a_s \, \Delta y_{t-s} + v_t \qquad 4.10$$

- Equation 4.8 represents a model with no constant term and no trend, as α=0 and there is no trend term included.
- Equation
- 4.*9* represents a model with a constant term, α, but has no trend.
- Equation
- 4.*10*, represents a model with both a constant term, α, and a trend λt[7].

When running the hypothesis test, there are some aspects that must be enlightened. The hypothesis is tested regarding the γ value, which expresses the stationarity/non-stationarity. The null-hypothesis (4.5) states that initial assumption is that the data are nonstationary, and therefore has a unit root (Hill et al., 2012, p. 486). The test is done by estimating the least squares, and then examine the t-statistics for the H_0. If H_0 is not rejected, the statistics needs to be changed. The reason is that if the TSD are nonstationary, the variance will increase with the sample size. Make use of a tau(τ)-statistic, and a new set of critical values, solves the problem with not rejecting H_0. According to Wooldridge (2013), after providing the new critical values, the t-statistics can be implemented.

4.3.3 Order of integration

To take the Dickey-Fuller test one step further one looks at the order of integration. If y_t is nonstationary, then γ=0. This results in the first difference becoming, $\Delta y_t = (y_t - y_{t-1}) = v$. *v* being an independent random variable $(0, \sigma^2_v)$, cause the difference to be stationary. The order of integration is the number of times the series must be differenced to become stationary (Hill et al., 2012, pp. 284-288). Hence, a series that has to be differenced once, is integrated of order one,

[7] In equation
4.*10*, the λ term is included. Looking at 4.7, this is not included in the equation for simplicity, to illustrate the difference.

I(1). This result is important to know before moving on with cointegration. To be able to test for cointegration, time series must be integrated of the same order.

4.4 Cointegration

Regression analysis is used for analysing the relationship between variables, but when regressing two nonstationary variables, it has its flaws. It might conclude that two independent, nonstationary variables with no correlation are, in fact correlated, and cause spurious regression. Cointegration is a method to investigate the true relationship between nonstationary variables. If two variables are cointegrated, the difference between the variables is stationary and therefore will exhibit a long-run equilibrium. Two cointegrated variables will have constant mean, constant variance and autocorrelation depending only on the time distance between any two variables (Wooldridge, 2013, pp. 646-652). Meaning that if they are cointegrated, they share the same stochastic characteristics. In other words, testing for cointegration, one can check if the variables are truly correlated.

4.4.1 Johansen test of cointegration

The Johansen cointegration test (Johansen, 1988, 1991; Johansen & Juselius, 1990) uses maximum likelihood estimation and makes it possible to estimate all cointegration vectors for two or more variables. If there are for example three variables and each have unit roots, there are at most two cointegrating vectors. If there are *n* variables, each with unit roots, there are *n*-1 cointegrating vectors. A vector autoregressive (VAR) model of order *k* >1, can be written as

$$\Delta \boldsymbol{x}_t = \boldsymbol{\Pi} \boldsymbol{x}_{t-1} + \sum_{i=1}^{k-1} \boldsymbol{\Pi}_i \Delta \boldsymbol{x}_{t-1} + \boldsymbol{\mu}_t \qquad 4.11$$

Π can be written as

$$\boldsymbol{\Pi} = \boldsymbol{\alpha\beta}' \qquad 4.12$$

Where **α** and **β** are matrixes of adjustment parameters and cointegration vectors, respectively. If there are two variables and one cointegration vector, **β** is a 2x1 vector and **α** is a 2x1 vector. The coefficients in **β**\` multiply the variables and deliver a linear combination of variables that does not have unit root. The adjustment coefficients in **α** deliver the response of the variables to deviations of the cointegration relationship. If **Π** is a matrix of zeroes, the variables are not cointegrated. A way to test if **Π = 0**, is to test if the rank is zero,

$$\text{rank}(\boldsymbol{\Pi}) = 0 \tag{4.13}$$

The variables are cointegrated if rank(**Π**) ≠ 0, and the rank = number of cointegration vectors. As mentioned, the number of cointegration vectors are always less than number of variables and at most *n-1*. If rank(**Π**) is less than *n*, the determinant is zero. In the Johansen test, eigenvalues are used to cope with this problem. The eigenvalues are ordered by size $\lambda_1 < \lambda_2 < \ldots. < \lambda_n$ and $\lambda_i \geq 0$. If $\lambda_1 = 0$, rank(**Π**) = 0 and hence, no cointegration vectors. If $\lambda_1 \neq 0$ there are at least one cointegration vector and testing continues.

There are two types of Johansen tests; maximum eigenvalue test and trace test. For this thesis, only the trace test is performed, as this *is considered more robust to skewness and excess kurtosis* (Yin-Wong & Lai, 1993, p. 324). The formula for the trace test is:

$$\lambda_{trace}(r,n) = -T \sum_{i=r+1}^{n} \ln(1-\lambda_i) \tag{4.14}$$

And the hypothesis to be tested is the null hypothesis, *H_0: rank(**Π**) is less or equal to r* against the alternative hypothesis, *H_1: rank(**Π**)* > *r*. So, for the first test $H_0=0$ and $H_1 \geq 1$. If $H_0=0$, there are no cointegration and testing ends, however, if the null is rejected the variables are cointegrated. The test statistics are an asymptotically distribution and compared to critical values following a chi squared distribution. The null is rejected if the calculated value is above the critical value.

Regarding the efficient market hypothesis, restrictions on α and β are imposed in the Johansen cointegration test. The market is efficient if there is cointegration and the joint test, β = (1, -1, 0) cannot be rejected. Separate test with restrictions on only α and only β are also performed to better see where possible bias occurs.

4.5 Granger causality test

As the price discovery role of futures prices is of interest for participants in a futures market, Granger causality test is performed. Granger causality test examines the lead-lag relationship between two cointegrated time series. In other words, the test examines whether spot price move after the futures price or futures prices move after spot price. It might also be a bidirectional relationship. If there exists a bidirectional relationship, one provides no information for characterizing the other. The following simple causal model is defined according to (Granger, 1969):

$$S_t = \sum_{i=1}^{m} \alpha_{1i} S_{t-i} + \sum_{j=1}^{m} \beta_{1i} F_{t-j} + e_{1t} \qquad 4.15$$

$$F_t = \sum_{i=1}^{m} \alpha_{2i} S_{t-i} + \sum_{j=1}^{m} \beta_{2i} F_{t-j} + e_{2t} \qquad 4.16$$

Where S_t and F_t denote spot and futures prices, respectively and e_t are uncorrelated white noise. If some of β_{1i}'s are not zero, futures prices are said to Granger cause spot prices. If some of α_{2i}'s are not zero, spot prices are said to Granger cause futures prices. If both are to occur, there exists a bidirectional relationship between futures prices and spot prices.

A F-test is used to test the null hypothesis that F_t does not Granger cause S_t, $\beta_{1i} = 0$ for all i's, or S_t does not Granger cause F_t, $\alpha_{2i} = 0$ for all i's. As the Granger causality test is based on linear relationships the test is only performed for the contract pairs which are cointegrated.

4.6 Diagnostic tests of the data

This section will present a set of tests performed to identify autocorrelation, normality, and heteroskedasticity. The tests conducted are, Breusch-Godfrey (BG) test for autocorrelation, Jarque-Bera (JB) test for normality, and ARCH test for heteroskedasticity. In addition, this section will provide definitions of the conditions, and the reason for testing. The results are discussed in the last paragraph of this section.

The characteristics of the BG test is of a Lagrange Multiplier (LM) method, which is *applied to test for autoregressive disturbance in a linear model where some of the regressors are lagged dependent variables* (Breusch, 1978, p. 342; Godfrey, 1978). The null hypothesis is given as H_0: there are no autocorrelation, while the alternative hypothesis, H_1, is that there is autocorrelation, providing a lagged dependency. In the sense of market efficiency, the futures price should reflect all available information, this is also enlightened in Movassagh and Modjtahedi (2005). Further, if the futures price fully reflects all available information, the error term should be independent, meaning autocorrelation is absent.

Skewness and kurtosis are the main preferences tested in the JB test for normality. Hill et al. (2012) portrays skewness as the symmetry around zero in a normal distribution. In addition, they portray the kurtosis as the "*peakedness*" (Hill et al., 2012, p. 148). The result of normality of the TSD, is independence of large sample approximations. The null hypothesis is that the data is normal distributed.

Heteroskedasticity is the violation of homoscedasticity. In other words, when the variance for all observation is not consistent, there is a condition of heteroskedasticity (Hill et al., 2012). There are ultimately two problems in the presence of heteroskedasticity. First, Hill et al. (2012) describes that one still has ordinary least squares (OLS) estimators that are unbiased, but they are no longer best estimates, meaning there are other estimates with smaller variance. Second, the standard errors are wrong, resulting in misleading confidence interval and hypothesis testing. Therefore, tests for heteroskedasticity is performed by using the ARCH model test.

The ARCH model (Autoregressive conditional heteroskedasticity) model is used in the case of working with data where variance varying in time (heteroskedasticity), and that depend (conditional) on lagged effects (autocorrelation) (Hill et al., 2012). The ARCH test is performed to check for ARCH effects. This means that one has a null hypothesis stating that there are no ARCH effects in the data set, and in the case of rejecting the null hypothesis, the conclusion is that there are ARCH effects in the data set.

The GB test were initially based on the information provided by the Johansen cointegration test, where the AIC criterion was default. The AIC criterion suggested for some of the outputs, that a lag of size twelve should be used. As Lade (Lade, 2016) enlightened, this is considered to reduce the significance of the test. This was solved by iteration, starting at two and increasing number of lag by one, until the H_0 was satisfied. The highest numbers of lags were detected in future 2 before shale gas revolution, being three (Table 11 in Appendix A). For all the others, both before and after, the lag number was two. The lag number were applied in the other tests, resulting in rejection of the null hypothesis for all JB test, concluding that the data are not normally distributed. For the ARCH test, the results showed signs to ARCH effects in contract 3 and contract 4 before shale gas, and contract 1 and contract 2 after shale gas. The test results are displayed in Table 12 in Appendix B.

5 Data

5.1 Description of data and time frame

Monthly historically spot and futures prices was downloaded for free at U.S Energy Information Administration (EIA). The durations of futures contracts chosen for this thesis were 1-, 2-, 3- and 4-months, further abbreviated as F1, F2, F3 and F4. All prices are based on delivery at the Henry Hub in Louisiana. US Natural gas futures contracts expire three business days prior to the first calendar day of the delivery month. Hence, the delivery month for F1, is the calendar month following the trade date. Prices are official daily closing prices from the trading floor of the New York Mercantile Exchange (NYMEX) for a specific delivery month. The amount of historically data available for spot- and futures prices were different. For simplicity considering comparisons, this thesis has used prices that coincide with the earliest historically available price for spot prices and as close as possible to today, which is December 2017. In other words, this thesis is based on historically prices from January 1997 until December 2017. For convenience, this thesis is based on average daily prices for a specific month. Meaning that, spot price for January 1997 is the average spot price recorded in January 1997. To make the prices more streamlined and adequate for analysis they are all log-transformed prior to analyzing.

5.2 Preliminary look at data

In this section a preliminary look at the data is conducted to see if there are any takeaways from descriptive statistics. For both periods the mean increases with maturity and could indicate the market of being in contango. For period one, the standard deviation increases with maturity. For the second period, standard deviation decreases with maturity. It should be noted that the volatility for period one is more than twice the volatility for period two. It could give implications of the market being more robust after the shale gas revolution. However, the overall price level has declined after the shale gas revolution. After a quick look at Figure 6 one can observe that for period one, spot- and futures prices follow each other fairly well, but as expected, the basis increases with maturity. Thus, spot price can be expected to be cointegrated with all futures contracts. In Figure 7 one can observe the relationship between futures prices and spot prices after 2009. Right after 2009 one can see that the basis increases a lot with maturity and is quite large for F1. However, the relationship seems to better the further away in time from 2009, and it could be expected that spot price is cointegrated with F1, F2 and F3. F4 might not be cointegrated with the spot price. It should be noted that this is just what is expected from pure observations of Figure 6 and Figure 7, and not a statement.

The last row in Table 1 and Table 2, is reserved for a calculation of the naive forecast (expectations). Tomek and Kaiser (2014) describes that the forecast of price is described only by the observed price. The two periods outline two different forecasts. In the period before 2009, the naive forecast expects a higher price than last observed price. In the period after 2009 the forecast is opposite and expects a decreasing price. Explained from Figure 6 and Figure 7, displaying increasing trend and decreasing trend, respectively. This is based on the calculated mean.

Table 1: Descriptive statistics of prices and basis before 2009.

Variable	Mean	Std.dev	Coefficient of variation	Min.	Median	Max
Spot	1,4925	0,5383	0,3607	0,5423	1,6302	2,5967
Contract 1	1,5088	0,5396	0,3576	0,5659	1,6465	2,5993
Contract 2	1,5311	0,5406	0,3531	0,5811	1,6605	2,6266
Contract 3	1,5448	0,5425	0,3512	0,5988	1,6659	2,6517
Contract 4	1,5511	0,5425	0,3498	0,6152	1,6705	2,6398
Basis 1	0,01632	0,044837	2,7467	-0,14685	0,009552	0,2095
Basis 2	0,03855	0,076682	1,9889	-0,25694	0,023582	0,32592
Basis 3	0,05225	0,107485	2,0573	-0,36583	0,03735	0,53099
Basis 4	0,05858	0,133486	2,2787	-0,44453	0,04546	0,59535
Naive forecast	0,00366	0,145349	39,7474	-0,47291	0,00172	0,47767

Table 2: Descriptive statistics of prices and basis after 2009.

Variable	Mean	Std.dev	Coefficient of variation	Min.	Median	Max
Spot	1,2152	0,2590	0,2131	0,5481	1,2326	1,7918
Contract 1	1,2269	0,2481	0,2022	0,5944	1,2682	1,7226
Contract 2	1,2513	0,2417	0,1932	0,6450	1,3056	1,7158
Contract 3	1,2739	0,2379	0,1868	0,6986	1,3219	1,7630
Contract 4	1,2930	0,2364	0,1828	0,7448	1,3263	1,7712
Basis 1	0,011719	0,045459	3,8790	-0,15024	0,004007	0,234662
Basis 2	0,03611	0,076735	2,1250	-0,25899	0,016086	0,396056
Basis 3	0,058693	0,095982	1,6353	-0,27531	0,039855	0,538467
Basis 4	0,077789	0,108133	1,3901	-0,27115	0,058868	0,586485
Naive forecast	-0,01534	0,150128	-9,7891	-1,03318	-0,02587	0,379633

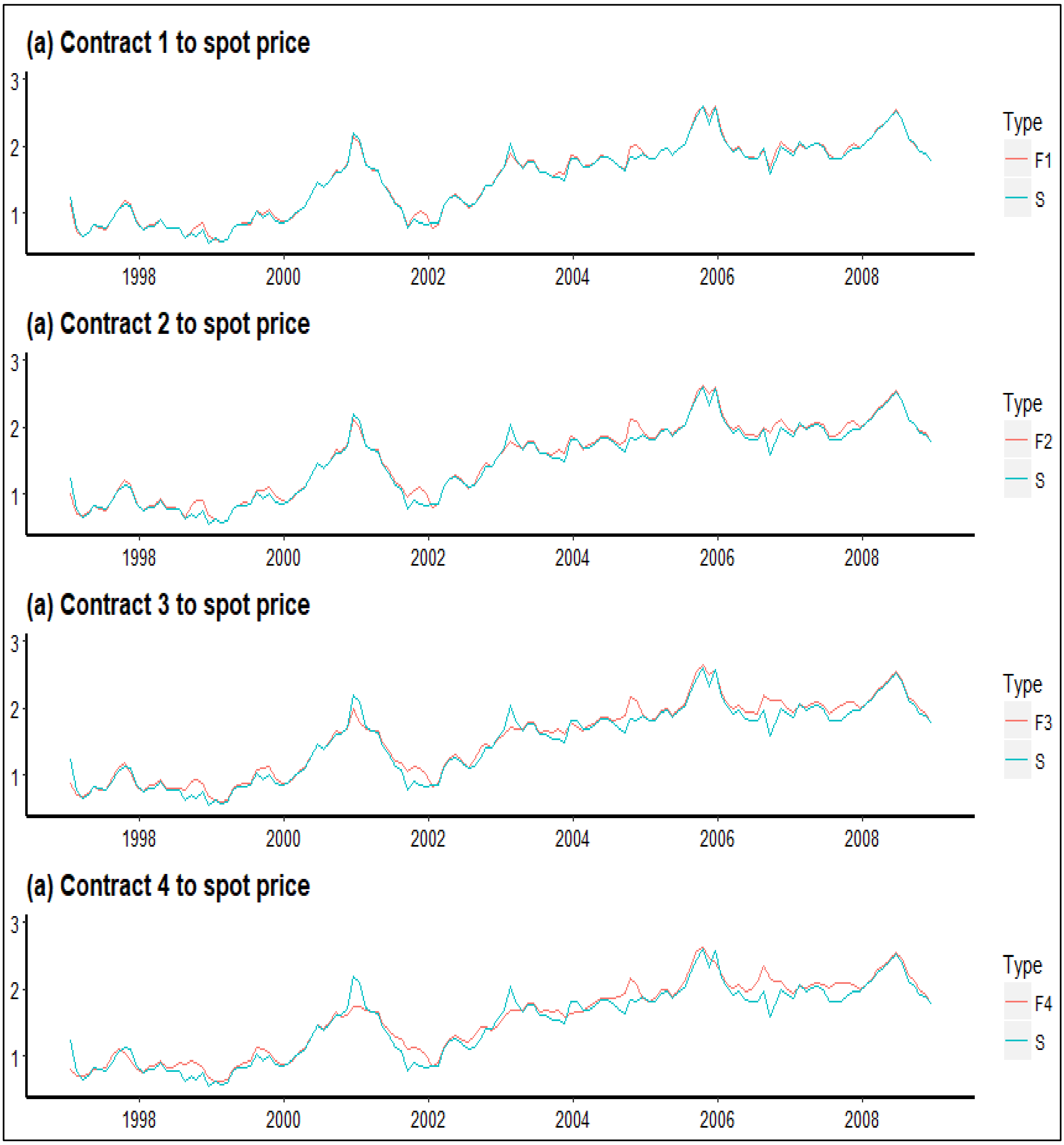

Figure 6: Futures prices and spot prices before 2009. All prices are in log-prices.

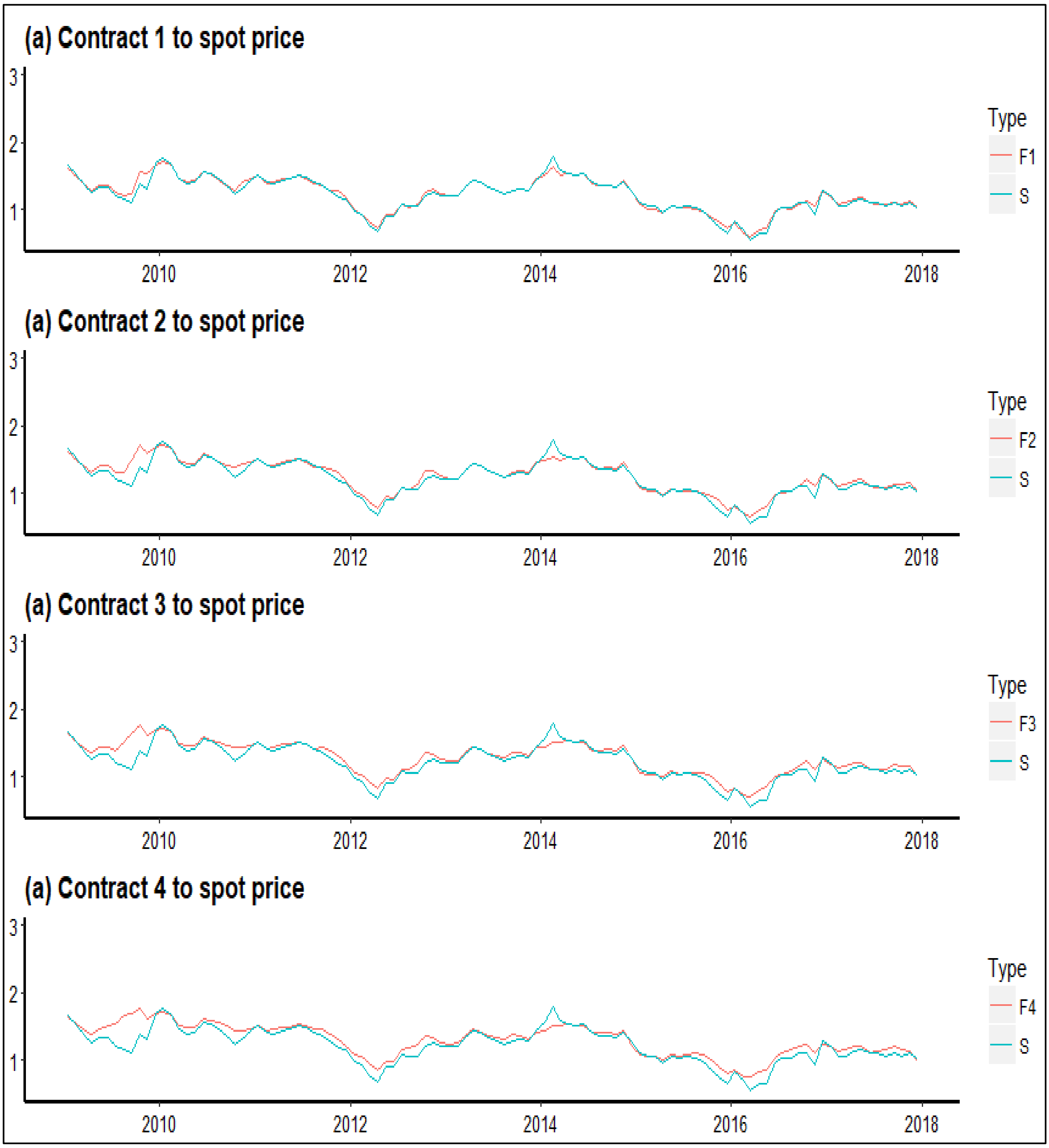

Figure 7: Futures prices and spot prices after 2009. All prices are in log-prices.

5.3 Contango/backwardation analysis

Simple calculations of basis have been performed to check for contango/backwardation characteristics. Basis has been calculated as defined by equation 3.7 were average values have been used as inputs. The results can be seen in Table 1 and Table 2, basis being positive and increasing with maturity, indicating characteristics of contango for both periods. Having compared the basis values against the standard deviation values, it is obvious that the uncertainty of the basis is large, as the standard deviations are large relative to the basis value. There is a belief that futures prices can be used to predict spot prices and hence, knowing if a market is in contango or backwardation can be useful for hedgers and speculators. A causality test will be performed to see if this common belief really is true.

5.4 ADF-results

The results from the ADF tests are listed in Table 3, Table 4, Table 5, and Table 6. The tables are categorized as before and after and tested for levels and for difference. The reason for this is to detect whether one can use the data sets in further analysis of cointegration. In addition to the test-statistics, critical values are also provided. Further discussion will be done below each table.

Table 3: ADF test for stationarity in levels, period before Shale gas revolution.

	Constant	Constant and trend
Contract 1	-2,100947	-3,313037
Contract 2	-2,182751	-3,477576
Contract 3	-2,04971	-3,267826
Contract 4	-2,053851	-3,350182
Spot price	-1,886892	-2,991573
	Critical values	
1 pct	-3,43	-3,99
5 pct	-2,88	-3,43
10 pct	-2,57	-3,13

Table 3 displays the test results of stationarity in levels of the period before introduction of shale gas. According to the hypothesis criteria described in section 4.3.2, equations *4.5* and *4.6*, it is observed that H_0 cannot be rejected for any of the no constant and no trend tests, at any of the critical values. Likewise, for the test with a constant. The final test provides a possibility for misinterpretation. The test displays a variety of rejecting and not rejecting H_0. By this, there is a possibility that H_0 could easily be rejected, when H_0 is true. Keeping in mind that one is dealing with negative numbers, it is observed that none of the test-statistics are smaller than critical value at 1 percent. Using the 1 percent criteria provides a stricter criterion for rejecting the null hypothesis. This makes it harder to reject the null hypothesis and reduces the risk of rejecting when H_0 is true. Summarized, all contracts and spot price are considered nonstationary in levels before shale gas revolution.

Table 4 ADF test for stationarity at levels, period after Shale gas

	Constant	Constant and trend
Contract 1	-2,27196	-2,779428
Contract 2	-2,119903	-2,758537
Contract 3	-1,91622	-2,663909
Contract 4	-1,783912	-2,695523
Spot price	-2,529314	-2,894917
	Critical values	
1 pct	-3,46	-3,99
5pct	-2,88	-3,43
10 pct	-2,57	-3,13

To avoid the confusion, which occurred in Table 3, it is desided to mainly focus on the 1 percent critical value in the testing. This means that the values displayed in Table 4 concludes that all contracts and spot price are nonstationary in levels after shale gas revolution for all tests.

Table 5 ADF test for stationarity at difference, period before Shale gas

	No constant no trend	Constant	Constant and trend
Contract 1	-8,69032	-8,694557	-8,68754
Contract 2	-8,649428	-8,653511	-8,653027
Contract 3	-7,938345	-7,947082	-7,958853
Contract 4	-7,554437	-7,568979	-7,593692
Spot price	-8,54703	-8,554466	-8,541145
	Critical values		
1 pct	-2,58	-3,46	-3,99
5 pct	-1,95	-2,88	-3,43
10 pct	-1,62	-2,57	-3,13

Table 5 and Table 6 are easier to interpret, as the test values are large negative numbers, which exceeds the critical values by a large margin[8], and it is observed that H_0 is rejected for all contracts and spot price, for all the tests. One can conclude that the data are stationary in differences.

Table 6 ADF test for stationarity at difference, period after Shale gas

	Constant	Constant and trend
Contract 1	-6,756201	-6,712521
Contract 2	-6,97292	-6,925225
Contract 3	-6,432733	-6,383931
Contract 4	-5,760675	-5,71434
Spot price	-7,158452	-7,114568
	Critical values	
1 pct	-3,46	-3,99
5 pct	-2,88	-3,43
10 pct	-2,57	-3,13

To summarize, the conclution of the ADF test is that all the provided data are characterised as nonstationary in levels, and stationary in difference. This is very important for further analysis such as the Johansson Cointegration test, as this is some of the criteria for the test. In other words one may say that the data is integrated of first order I(1) and applicable for cointegration test.

[8] As we talk of negative values it may be somewhat misleading saying large with negative numbers, as a large negative value means a very small value.

6 Results

This section provides the empirical results from tests on the efficient market hypothesis and Granger causality test.

6.1 Results from Johansen cointegration test

In this section the results from Johansen cointegration test are presented. Table 7 shows the test statistics for period 1 and Table 8 shows the test statistics for period 2. Column 1 shows the cointegration pair, columns 2 lists the null hypothesis with corresponding trace statistics in column 3. The estimated coefficients α and β for the unrestricted cointegration test is listed in column 4. The joint test, test with restriction on β and test with restriction on α are listed in column 5, 6 and 7, respectively. The lags chosen can be seen in Appendix A.

As mentioned, column 2 displays the null hypothesis. If the null, $r=0$, is rejected there are cointegration between the pairs and then they are further tested to find the number of cointegration vectors. If also $r\leq 1$ is rejected, there is only one cointegration vector. Column 5 shows the test statistics of the restricted joint hypothesis $\alpha=0$ and $\beta=1$, with p-values in brackets. The significance level was set to 5%, and a p-value less than 0,05 results in rejection of the joint hypothesis. With a p-value above or equal to the selected significance level it can be concluded as no rejection of the joint hypothesis, and hence, market efficiency. Separate tests for restrictions on only α and only β are also performed.

6.1.1 Period 1

Spot price was cointegrated with F1, F2, F3 and F4 at 1% significance level. All cointegration pairs had only one cointegration vector at a 1% significance level. From column 5 it can be observed that the joint hypothesis test was rejected for all cointegrated pairs and thus, the market has not been efficient for any of the futures contracts. When a separate test was performed with restriction on β only, the null was not rejected for any contract pair. For restrictions on α only, the null was rejected for every contract pair.

Table 7: Johansen cointegration test statistics for period 1

Contract	H0	Trace statistics	Estimated coefficients α, β	Joint test H0: α=0, β=1	Restriction on β H0: β=1	Restriction on α H0: α=0
Spot	r=0	60,22***	-0,0239 1,0039	12,51 (0,00)**	1,92 (0,38)	18,83 (0,00)**
Contract 1	r≤1	3,91				
Spot	r=0	35,87***	-0,0368 0,9965	12,37 (0,00)**	0,36 (0,84)	10,32 (0,01)**
Contract 2	r≤1	4,02				
Spot	r=0	41,59***	-0,0542 0,9983	11,06 (0,00)**	0,04 (0,98)	10,35 (0,01)**
Contract 3	r≤1	3,69				
Spot	r=0	35,37***	-0,05901 0,9969	7,61 (0,02)**	0,06 (0,97)	6,74 (0,03)**
Contract 4	r≤1	3,88				
Critical values						
10%	r=0	17,85				
	r≤1	7,52				
5%	r=0	19,96				
	r≤1	9,24				
1%	r=0	24,6				
	r≤1	12,97				

*Note: *, ** and *** denotes rejection of the null for 10%, 5% and 1% respectively.*

6.1.2 Period 2

For period 2, spot price was cointegrated with F1, F2, F3 and F4 at 1% significance level. All the pairs that was cointegrated only had one cointegration vector, and at 1% significance level. Although all the contract pairs were cointegrated, the joint hypothesis test was rejected for F2, F3 and F4, implying market efficiency only for the futures contract with shortest time to maturity, F1. With restriction on β, the null hypothesis was rejected for F1 and F2, but not rejected for F3 and F4. When imposing restrictions on α, the null was rejected for all the contract pairs.

Table 8: Johansen cointegration test statistics for period 2.

Contract	H0	Trace statistics	Estimated coefficients α, β	Joint test H0: α=0, β=1	Restriction on β H0: β=1	Restriction on α H0: α=0
Spot	r=0	41,23***	-0,0395 1,0144	4,79 (0,09)	7,15 (0,03)**	15,87 (0,00)**
Contract 1	r≤1	5,07				
Spot	r=0	37,83***	-0,0697 1,0253	9,1 (0,01)**	6,68 (0,04)**	18,23 (0,00)**
Contract 2	r≤1	4,59				
Spot	r=0	35,90***	-0,0847 1,0188	11,26 (0,00)**	2,53 (0,28)	16,26 (0,00)**
Contract 3	r≤1	4,01				
Spot	r=0	29,42***	-0,0745 0,9952	10,53 (0,01)**	0,11 (0,95)	9,57 (0,01)**
Contract 4	r≤1	3,77				
Critical values						
10%	r=0	17,85				
	r≤1	7,52				
5%	r=0	19,96				
	r≤1	9,24				
1%	r=0	24,6				
	r≤1	12,97				

*Note: *, ** and *** denotes rejection of the null for 10%, 5% and 1% respectively.*

6.2 Results from Granger causality test

The results from Granger causality test are listed in Table 9 for period 1 and Table 10 for period 2. The test were performed for every pair that was cointegrated. Column 2 shows the tested null hypothesis with the corresponding contract pair in column 1. F-test statistics and p-values are listed in column 3 and 4, respectively. The significance level was set to 5%, hence, a p-value less than 0,05 results in rejection of the null.

In period 1, there are bidirectional relationships between every contract pair except for F4. Meaning that for F1, F2 and F3 none of the prices are following the other or containing more information than the other. For the contract pair with the longest maturity, spot price Granger cause futures price.

In period 2, there is bidirectional relationship between spot price and contract 1. F2 and F3 seems to Granger cause spot price. For contract 4 it seems to be a bidirectional relationship, although the p-value for the null that spot price does not Granger cause futures price is very high relative to the p-value that futures price does not Granger cause spot price.

Table 9: Granger causality test statistics for period 1.

Contract pair	Ho	F-test statistics	p-value
Spot	S_t does not Granger cause F_{t-1}	6,2195	0,002285**
Contract 1	F_{t-1} does not Granger cause S_t	6,7635	0,001359**
Spot	S_t does not Granger cause F_{t-2}	7,1254	0,000128**
Contract 2	F_{t-2} does not Granger cause S_t	5,544	0,00105**
Spot	S_t does not Granger cause F_{t-3}	5,8393	0,003288**
Contract 3	F_{t-3} does not Granger cause S_t	4,3516	0,0138**
Spot	S_t does not Granger cause F_{t-4}	3,8335	0,02281**
Contract 4	F_{t-4} does not Granger cause S_t	2,9252	0,05534

*Note: ** denotes rejection of the null at chosen significance level 0,05.*

Table 10: Granger causality test statistics for period 2.

Contract pair	Ho	F-test statistics	p-value
Spot	S_t does not Granger cause F_{t-1}	4,7397	0,00975**
Contract 1	F_{t-1} does not Granger cause S_t	7,99	0,00046**
Spot	S_t does not Granger cause F_{t-2}	2,6683	0,07184
Contract 2	F_{t-2} does not Granger cause S_t	3,4904	0,03236**
Spot	S_t does not Granger cause F_{t-3}	2,0208	0,1352
Contract 3	F_{t-3} does not Granger cause S_t	3,168	0,04421**
Spot	S_t does not Granger cause F_{t-4}	0,9071	0,4053
Contract 4	F_{t-4} does not Granger cause S_t	2,9658	0,05379

*Note: ** denotes rejection of the null at chosen significance level 0,05.*

7 Discussion

In this section a summary of key findings will be presented and a discussion on what the results may imply and possible explanations. First, a discussion of volatility of natural gas prices and basis and how this has changed after the shale gas revolution. Second, a discussion of cointegration analysis. Third, naive forecast versus futures-based forecast are discussed. Following will be a discussion of granger causality test and finally contango/backwardation will is discussed.

7.1 Discussion of volatility

The discussion of volatility is based on calculations displayed in Table 1 and Table 2, for period 1 and 2, respectively. The US natural gas market suffered a massive increase in production and higher availability due to increased storage, which resulted in decreasing prices after the shale gas revolution. Not only did prices decrease, the volatility of prices also decreased. These findings imply that it was more difficult to forecast natural gas prices in the period before 2009. With the massive increase in storage after the shale gas revolution, suppliers can adapt faster to changes in demand, which results in a decrease in volatility. These results could imply that with more stable prices, it might not be as interesting for speculators to invest in natural gas futures contracts.

Another aspect is that with a lower price volatility, the risk premium will be lower. With a lower risk premium, the market might exhibit characteristics of an efficient market as risk premium often is a popular explanation for the efficient market hypothesis to fail.

7.2 Discussion of cointegration analysis results

This thesis defines weak-form efficiency when there exists cointegration and the joint hypothesis cannot be rejected. For period 1 there was significant cointegration between all contract pairs, however the joint hypothesis was rejected and hence, the natural gas futures market was not efficient. These findings are consisten with Movassagh and Modjtahedi (2005), Wei and Zhu (2006), Mazighi (2003), and Moosa and Al-Loughani (1994). When performing a separate test with restriction on β only, the hypothesis was not rejected for any of the contracts. On the other hand, performing a test with restriction on α only, the hypothesis was rejected for every contract pair. This implies that α is the reason for futures contracts being a biased estimator for future spot price. By rejecting the α hypothesis, it is reasonable to say that there are costs in the form of risk premium. "*A common explanation for the rejection of the simple efficiency hypothesis has been* *the existence of a risk premium. Such a risk premium can account for the existence of non-zero speculative returns in the futures market. This does not imply markets are inefficient, only that investors require compensation for the risk they undertake"* (Crowder & Hamed, 1993).

For period 2, the results was more mixed than for period 1. There was significant cointegration between all contract pairs, and for contract 1 the market was efficient as the joint hypothesis was not rejected. For F2, F3, and F4 the market is inefficient. For F3 and F4 the principle of co-movement is respected, and the bias seem to occur in the α coefficient. For F2 the bias occur in both α and β. Since longer time to maturity increases the price risk in the sence of unpredictable events, it is reasonable to believe that time to maturity could be an explanation.

Having results indicating that the marked changes from unefficient to efficient after 2009 could be very important to shareholders in the natural gas market. Tomek and Kaiser (2014) argues that the determined price based on the available information is wrong due to the information being wrong. This does not make a market inefficient, but the price is determined on wrong information. Further they states that market efficiency means that *the price correclty reflects what is known at the current time even if the information is flawed* (Tomek & Kaiser, 2014, p. 295). For period 1 the tests displays that the market is inefficient. The data used are historical, but this does not mean that the price at that time correclty reflects the information at the same time. If the market in period 1 was indeed inefficient, the futures contracts are in fact considered to be biased estimators. This means that investments and speculations was previously based on a forecast model with biased estimator. Meaning that there existed possibilities of arbitrage before the shale gas revolution.

Further, the tests indicate that in period 2 the market is efficient for F1. This is based on futures contract 1 and spot price are cointegrated, and no rejection of unbiasedness hypothesis. It is considered to be sufficient to say that the price correctly reflects the information as mentioned above. One may ask if there has been a change in how to interpret information. However, based on these results one may say that the changes from period 1 to period 2 has had an effect in favor of futures contracts as a forecast model. A question for further studies could be if this model now indeed is more reliable in forecasting. Therefore it is still reasons to question both the definition of market efficiency, and the futures contracts as an unbiased estimator.

In regards to over- and underestimation, Lai and Lai (1991) states that futures price and spot price needs to be cointegrated to achieve market efficiency. Further, if the futures contract is an unbiased estimator, the futures price does not consistently over- or underestimate spot prices. Based on the findings of this thesis one may expect no consistent over- and underestimation for futures contract 1 in period 2, as it complies both conditions of cointegration and joint test making it market efficient and unbiased. All other contracts in period 1 and 2 does not satisfy the market efficicency and unbiasedness criteria. This leads to the assumption of over- and undersestimation. From the results of the cointegration test displayed in Table 7 and Table 8 it is possible to evaluate over-/underestimation. If the spot price is on the lefthandside in the equation of regression and the β coefficient equals 1, the focus should be directed to the α coefficient. All the α coefficients displayed in column "estimated coefficients" are negative for both periods. Meaning that the futures price consitently overestimates the spot price. The difference between future price and spot price is α. As already mentioned it is reasonable to assume that α here represents the risk premium, and for US natural gas, by going short in a futures contract. Risk premium is often related to systematic risk and hedging pressure (De Roon, Nijman, & Veld, 2000). Hedging pressure means that if net positions of hedging exceedes net positions of speculations, speculators needs compensation in form of risk premium. De Roon et al. (2000) finds that hedging pressure has a significant affect on futures returns.

7.3 Naive forecast versus futures contract forecast

As previously mentioned, according to naive expectations, the price tomorrow is only dependent on price today. There is a question as to whether simple naive forecasts can predict prices better than futures contract forecasts. For the first period, naive forecasts are on average expecting higher prices. Comparing with basis which has a higher mean, one could say that naive expectations are a better forecast than futures contract forecast. However, according to coefficient of variation, the volatility of naive forecasts was much higher than for futures contracts. For the second period, naive forecast expects decreasing prices on average. Comparing with basis, which has higher mean, one could again say that naive forecasts are better than futures contract forecasts. But again, coefficient of variation is close to 3-4 times the CV of basis.

To summarize, futures contract forecasts are better than naive forecasts in the sense that the volatility is a lot higher for naive forecasts. This is also intuitive as futures contract contains all relevant available information.

7.4 Granger causality test

There is a general belief that futures price can exhibit a price discovery role for future spot price. Results from Granger causality test shows a change between the two periods regarding causality between futures- and spot price. Results implies that there was a bidirectional relationship between futures price and spot price before the shale gas revolution, contradicting the belief that futures prices lead spot price. However, as the two prices converge as maturity date approaches, people tend to believe that futures prices can provide information for predicting future spot price.

After the shale gas revolution, the relationship seems to be different. For the contract with shortest time to maturity (F1) and the only contract which exhibit market efficiency, there is a bidirectional relationship. This result is also intuitive as an unbiased futures contract indicates that the futures contract does not on average over- or under-estimated future spot price. The only futures contracts that Granger cause spot price is F2 and F3 after the shale gas revolution. Hence, F2 and F3 can have a price discovery role for future spot price. These findings imply that information flows from futures price to spot price.

Another study which tested causality on natural gas futures found that futures prices Granger caused spot price using a linear model (Zhang & Liu, 2018). However, using a nonlinear model Zhang and Liu (2018) found a bidirectional relationship.

7.5 Discussion of contango/backwardation relationship

As already mentioned, the only futures contract being efficient is the one-month contract in period 2. All other futures prices are a biased predictor of future spot price, meaning that there are other factors than futures prices affecting spot price. A simple contango/backwardation analysis might be of interest with respect to the market dynamics before and after the shale gas revolution. From Table 1 and Table 2 it can be observed an average basis being positive, indicating a market which exhibits characteristics of contango. Comparing the basis of both periods they seem to be relatively equal. Basis is increasing with increasing maturity in both periods and with a relatively equal volatility.

The relationship between futures prices and spot prices can be explained by either the theory of storage, where basis is related to interest rates, storage cost and convenience yield. Contango is then a result of increasing storage cost and interest rates and lower convenience yield. However, explaining contango using risk premium where futures market serves as a risk management tool for risk averse producers is also possible. In addition, this could be supported by the previous section discussing the results from cointegration test (7.2). Sellers of futures contracts are often seen as risk averse short traders, often producers, seeking to hedge for unwanted price changes. As the net positions in a futures market needs to be zero, there needs to be sufficient number of traders (speculators) taking long positions. For a speculator to take a long position he need to be compensated in the form of a risk premium. Contango is typically a result due to investors preferring to pay a premium to have the commodity in the future rather than paying storage and carry cost for acquiring the commodity in the present. One reason for natural gas market being in contango may be because of the high storage cost related to fossil fuels. For both periods natural gas futures prices has on average over-estimated future spot price.

It should be noted that the contango/backwardation analysis in this thesis is quite simple and a more thorough study performed by Movassagh and Modjtahedi (2005) found natural gas futures generally under-predict spot prices which is evidence on a market being in backwardation in the period January 1991 to November 2003. This leads to an inconsistency in the market, as the results of this thesis indicates both contango and an overestimation of spot price when the market is not efficient.

8 Conclusion

This thesis performed empirical analysis on the US natural gas futures market to test the efficient market hypothesis. The structural change in 2009 found by Oglend et al. (2015) founded the basis for this thesis as to whether there has been a change in the US natural gas futures market after the shale gas revolution. The problem was: "*Test for market efficiency for US natural gas before and after shale gas revolution, to determine whether the futures contracts are unbiased estimator of the future spot price*". Following will be summary of the most important findings and conclusions drawn from the empirical analysis.

- US natural gas prices suffered a decline after the shale gas revolution. However, the volatility of spot and futures prices declined indicating that predicting prices before 2009 was more difficult.
- Before the shale gas revolution, futures prices were a biased estimator of future spot price and the market was not efficient. The principle of co-movement was respected, hence the bias occurred in the α coefficient. The α coefficient was negative, hence futures prices have overestimated spot prices before the shale gas revolution.
- After the shale gas revolution, only the contract with the shortest time to maturity was efficient and an unbiased estimator of future spot price. These results imply that futures contract with one month to maturity has neither over- or underestimated spot price. For the contract with 2-, 3- and 4-months to maturity the market has not been efficient. As the α coefficient was negative, futures price has overestimated spot price.
- It is fair to assume that bias occurs due to a risk premium of going short in US natural gas futures as futures has in general been overestimating spot price. An increasing α is intuitive with higher risk as time to maturity increases. The fact that volatility decreased after the shale gas revolution can be an explanation for futures contract with 1-month to maturity has been an unbiased estimator of future spot price.
- Comparing futures-based forecast to naive forecasts, naive forecasts has on average been closer to spot price. However, the volatility is of naive forecasts is extreme compared to futures-based forecasts.
- The simple contango/backwardation analysis showed that for both periods, the US natural gas market has been in contango. These findings are also consistent with the results from cointegration analysis.

- Results from granger causality test showed that there was a bidirectional relationship between futures prices and spot prices before 2009. After 2009 there was a bidirectional relationship between spot price and contract 1. However, contracts with 2- and 3-month to maturity has lead the spot price.

9 References

Administration, E. I. (2017, August 23, 2017). Factors Affecting Natural Gas Prices. Retrieved from https://www.eia.gov/energyexplained/index.cfm?page=natural_gas_factors_affecting_prices

Alquist, R., & Kilian, L. (2010). What do we learn from the price of crude oil futures? *Journal of Applied Econometrics*, *25*(4), 539-573. doi:10.1002/jae.1159

Banks, E. (2005). *Exchange-Traded Derivatives*: John Wiley & Sons.

Brennan, M. J. (1958). The supply of storage. *American Economic Review, 48*, 50-72.

Breusch, T. S. (1978). TESTING FOR AUTOCORRELATION IN DYNAMIC LINEAR MODELS*. *Australian Economic Papers, 17*(31), 334-355. doi:10.1111/j.1467-8454.1978.tb00635.x

Chinn, M. D., LeBlanc, M., & Coibion, O. (2005). *The predictive content of energy futures: an update on petroleum, natural gas, heating oil and gasoline*. Retrieved from

Chiou Wei, S. Z., & Zhu, Z. (2006). Commodity convenience yield and risk premium determination: The case of the U.S. natural gas market. *Energy Economics, 28*(4), 523-534. doi:10.1016/j.eneco.2005.10.002

Considine, T. J., & Larson, D. F. (2001). Risk premiums on inventory assets: the case of crude oil and natural gas. *Journal of Futures Markets*, *21*(2), 109-126.

Crowder, W. J., & Hamed, A. (1993). A Cointegration Test for Oil Futures Market Efficiency. *Journal of Futures Markets, 13*(8), 933-941. doi:http://onlinelibrary.wiley.com/journal/10.1002/%28ISSN%291096-9934/issues

De Roon, F. A., Nijman, T. E., & Veld, C. (2000). Hedging Pressure Effects in Futures Markets. *Journal of Finance, 55*(3), 1437-1456. doi:10.1111/0022-1082.00253

Dwyer, G. P., & Wallace, M. S. (1992). Cointegration and market efficiency. *Journal of International Money and Finance*, *11*(4), 318-327. doi:10.1016/0261-5606(92)90027-U

Fabozzi, F. J., Füss, R., & Kaiser, D. G. (2008). *The handbook of commodity investing*. In Frank J. Fabozzi Series.

Fama, E. F. (1991). Efficient Capital Markets: II. *Journal of Finance*, *46*(5), 1575-1617. doi:http://onlinelibrary.wiley.com/journal/10.1111/%28ISSN%291540-6261/issues

Gebre-Mariam, Y. K. (2011). Testing for unit roots, causality, cointegration, and efficiency: The case of the northwest US natural gas market. *Energy*, *36*(5), 3489-3500. doi:10.1016/j.energy.2011.03.055

Godfrey, L. G. (1978). Testing Against General Autoregressive and Moving Average Error Models when the Regressors Include Lagged Dependent Variables. *Econometrica*, *46*(6), 1293-1301. doi:10.2307/1913829

Granger, C. W. J. (1969). INVESTIGATING CAUSAL RELATIONS BY ECONOMETRIC MODELS AND CROSS-SPECTRAL METHODS. *Econometrica*, *37*(3), 424-438.

Hill, R. C., Lim, G. C., & Griffiths, W. E. (2012). *Principles of econometrics* (4th ed. ed.). Hoboken, N.J: Wiley.

Investopedia. (2018a). Backwardation. Retrieved from https://www.investopedia.com/terms/b/backwardation.asp

Investopedia. (2018b). Contango. Retrieved from https://www.investopedia.com/terms/c/contango.asp

Investopedia. (2018c). Spot Market. Retrieved from https://www.investopedia.com/terms/s/spotmarket.asp

Ji, Q., Zhang, H.-Y., & Geng, J.-B. (2018). What drives natural gas prices in the United States? – A directed acyclic graph approach. *Energy Economics*, *69*, 79-88. doi:10.1016/j.eneco.2017.11.002

Johansen, S. (1988). Statistical Analysis of Cointegration Vectors. *Journal of Economic Dynamics and Control, 12*(2/3), 231-254. doi:http://www.sciencedirect.com/science/journal/01651889

Johansen, S. (1991). Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models. *Econometrica, 59*(6), 1551-1580. doi:http://www.econometricsociety.org/tocs.asp

Johansen, S., & Juselius, K. (1990). Maximum Likelihood Estimation and Inference on Cointegration--With Applications to the Demand for Money. *Oxford Bulletin of Economics and Statistics, 52*(2), 169-210. doi:http://onlinelibrary.wiley.com/journal/10.1111/%28ISSN%291468-0084/issues

Kaldor, N. (1939). Speculation and economic stability. *Review of Economic Studies, 7*, 1-27.

Lade, R. M. (2016). Market Efficiency in the Crude Oil Futures Market - an Empirical Study after the Shale Oil Revolution. In: University of Stavanger, Norway.

Lai, K. S., & Lai, M. (1991). A Cointegration Test for Market Efficiency. *Journal of Futures Markets, 11*(5), 567-575.

Lee, C.-C., & Lee, J.-D. (2009). Energy prices, multiple structural breaks, and efficient market hypothesis. *Applied Energy, 86*(4), 466-479. doi:10.1016/j.apenergy.2008.10.006

Malkiel, B. G. (2003). The Efficient Market Hypothesis and Its Critics. *Journal of Economic Perspectives, 17*(1), 59-82. doi:10.1257/089533003321164958

Malkiel, B. G., & Fama, E. F. (1970). EFFICIENT CAPITAL MARKETS: A REVIEW OF THEORY AND EMPIRICAL WORK*. *Journal of Finance, 25*(2), 383-417. doi:10.1111/j.1540-6261.1970.tb00518.x

Marvik, E. (2016). Sammenligning av forholdet mellom pris og lager av naturgass før og etter skifergassrevolusjonen. In: University of Stavanger, Norway.

Mishra, V., & Smyth, R. (2016). Are natural gas spot and futures prices predictable? *Economic Modelling, 54*, 178-186. doi:10.1016/j.econmod.2015.12.034

Modjtahedi, B., & Movassagh, N. (2005). Natural-gas futures: Bias, predictive performance, and the theory of storage. *Energy Economics, 27*(4), 617-637. doi:10.1016/j.eneco.2005.04.005

Moosa, I. A., & Al-Loughani, N. E. (1994). Unbiasedness and time varying risk premia in the crude oil futures market. *Energy Economics, 16*(2), 99-105.

Oglend, A., Lindback, M. E., & Osmundsen, P. (2015). Shale Gas Boom Affecting the Relationship between LPG and Oil Prices. *Energy Journal, 36*(4), 265-286. doi:http://www.iaee.org/en/publications/journal.aspx

Regulator, A. E. (2018). What is Hydraulic Fracturing? Retrieved from https://www.aer.ca/about-aer/spotlight-on/unconventional-regulatory-framework/what-is-hydraulic-fracturing

Switzer, L. N., & El-Khoury, M. (2007). Extreme volatility, speculative efficiency, and the hedging effectiveness of the oil futures markets. *Journal of Futures Markets, 27*(1), 61-84. doi:10.1002/fut.20235

Tomek, W. G., & Kaiser, H. M. (2014). *Agricultural product prices* (Fifth ed. ed.). Ithaca: Cornell University Press.

Walls, W. D. (1995). Econometric analysis of the market for natural gas futures. *Energy Journal, 16*(1), 71-83. doi:10.5547/ISSN0195-6574-EJ-Vol16-No1-5

Wooldridge, J. M. (2013). *Introductory econometrics : a modern approach* (5th ed., international ed. ed.). S.l.: South-Western, Cengage Learning.

Working, H. (1948). Theory of the inverse carrying charge in futures markets. *Journal of Farm Economics, 30*, 1-28.

Yin-Wong, C., & Lai, K. S. (1993). Finite-sample sizes of Johansen's likelihood ratio tests for cointegration. *Oxford Bulletin of Economics & Statistics, 55*(3), 313. doi:10.1111/j.1468-0084.1993.mp55003003.x

Zhang, Y., & Liu, L. (2018). The lead-lag relationships between spot and futures prices of natural gas. *Physica A, 490*, 203-211. doi:10.1016/j.physa.2017.08.018

Appendix A Lag selection for Johansen cointegration test

Table 11: Lag selection for Johansen cointegration test

	Period 1			
	Spot price	Spot price	Spot price	Spot price
Lags	Contract 1	Contract 2	Contract 3	Contract 4
AIC	2	12	12	13
Lags chosen	2	3	2	2
	Period 2			
	Spot price	Spot price	Spot price	Spot price
Lags	Contract 1	Contract 2	Contract 3	Contract 4
AIC	2	2	2	2
Lags chosen	2	2	2	2

Note: All the lags listed in Table 11 are lags chosen according to Akaike Information Criterion (AIC). Table 11 list the lags suggested for a VAR model according to "RStudio"-software and by using the "VARselect"-function in the "vars"- package.

Appendix B P-values for diagnostic tests

Table 12: P-values for the diagnostic tests

Diagnostic test in period befor shale gas revolution				
	Spot price F1	Spot price F2	Spot price F3	Spot price F4
Breauch-Godfrey	0,4033	0,0896	0,2306	0,0500
Jarque-Bera	0,0000	0,0000	0,0000	0,0000
ARCH	0,4355	0,9291	0,0054	0,0044
Diagnostic test in period after shale gas revolution				
	Spot price F1	Spot price F2	Spot price F3	Spot price F4
Breauch-Godfrey	0,3098	0,4521	0,6032	0,7427
Jarque-Bera	0,0000	0,0000	0,0000	0,0000
ARCH	0,0000	0,0246	0,0617	0,1460

Note: The table is restricts the values to only include the p-values from the tests. The p-values are used in the hypothesis testing (the hypothesises are given in section 0) and are considered more critical than the test statistics for our thesis. In addition, the conclusions drawn from these results are based on a significance level of 5% (0,05).