BI Norwegian Business School - campus Oslo

GRA 19703

Master Thesis

Thesis Master of Science

An empirical analysis of the risk premium in the crude oil futures market

Navn: Endre Aamodt, John Hill Yip

Start: 15.01.2019 09.00

Finish: 01.07.2019 12.00

Thesis Report
BI Norwegian Business School

An empirical analysis of the risk premium in the crude oil futures market

Date
1st July 2019

Supervisor
Costas Xiouros

Study program
Msc. Business (Finance)

Table of Contents

Table overview

Figure overview

Acknowledgement

We would like to thank our supervisor, Associate Professor in the Department of Finance at BI Norwegian Business School, Costas Xiouros for valuable input and guidance during the process of writing.

Abstract

This paper investigates the unbiasedness of the crude oil futures price in two time periods: 1986-2019 and 2006-2019. The unbiasedness of the futures price is examined using linear regression in an in-sample setting and through assessing the predictive accuracy of alternative forecasting models, with different assumptions concerning the risk premium, in an out-of-sample setting. The results from the full time period (1986-2019) suggest that the futures price is an unbiased predictor of the future spot price of crude oil, indicating that there is no risk premium in the futures price. This finding is consistent in both the in- and out-of-sample analyses. The results from the sub-period (2006-2019) suggest that the futures price is a biased predictor of the future spot price of crude oil for medium-long maturities, indicating a risk premium in the futures price. However, whether the risk premium is constant and/or time-varying is inconclusive.

1. Introduction

The predicted future spot price of oil is one of the key variables in modeling macroeconomic forecasts (Alquist & Kilian, 2010). Such forecasts are used by analysts, central banks and governments in generating expectations about the direction of the economy and hence about the future state of the world (Baumeister & Kilian, 2012). Real-time forecasts of the future price of oil thus largely affect financial markets all over the world. The unstable nature of the price of oil, combined with the important role of predicted oil prices, makes it important to understand oil price determination and to have an ability to accurately forecast the future price of oil.

Extensive research has already been devoted to investigating the unbiasedness of the crude oil futures price in an out-of-sample setting. In practice, this entails examining the forecasting accuracy of a set of models with different assumptions concerning the risk premium in the futures price. However, despite extensive amount of research, there is no consensus on which forecasting method produces the highest predictive accuracy. Consequently, it has proven difficult to construct forecasting models which systematically outperform the naïve no-change forecast over longer horizons (Baumeister & Kilian, 2012).

A common view is that the futures price of oil is an unbiased predictor of the future spot price of oil. This belief implies that financial markets on average generate more accurate predictions of the spot price of oil than econometrical models and survey forecasts. The first objective of this thesis is to challenge this view by examining the unbiasedness of the futures price of crude oil. This is examined using linear regression in an in-sample setting assuming the futures prices are rational. Whether this assumption holds is tested by comparing the results from the in-sample analysis with the results obtained from running the Capital Asset Pricing model (CAPM) (Lintner, 1965) and Fama French three-factor model (Fama & French, 1993).

The second objective of this thesis is to examine the predictive accuracy of a set of alternative forecasting models on the future spot price of crude oil. The alternative forecasting models make different assumptions concerning the unbiasedness of the futures price. This means that these forecasting models examine the existence and potential form of a risk premium in the crude oil futures market. The proposed models are tested pseudo-out-of-sample and are evaluated based on their mean squared prediction error and directional accuracy.

A multivariate model is introduced in order to examine whether there is a time-varying risk premium embodied in the futures price. The multivariate model includes several financial and fundamental variables that are suggested to affect the future spot price of crude oil through the risk premium. The variables that are included in the multivariate model are: US crude oil inventory, US crude oil production, global rig activity, market excess return, treasury yield curve, credit spread, realized volatility and gasoline product spread. The inclusion of these variables is based on theoretical arguments for their predictive power on the spot price of oil through the risk premium and findings in previous literature.

The data set used to conduct the analyses stretches from January 1986 to April 2019. This chosen time period is due to data availability. The type of oil chosen is the WTI Crude, a light, sweet crude oil which is ideal for conversion to gasoline and diesel fuel. Finally, the thesis focuses on nominal prices of crude oil, rather than real prices.

In summary, the objective of this thesis is to assess the unbiasedness of the futures price and its real-time out-of-sample forecasting ability of the forecasting models on the spot price of crude oil. The models are tested upon two time periods and at horizons from one month up to 12 months forward. Finally, the results and their implications are examined along with relevant economic theory.

The objective of the thesis can thus be summarized into the following research question:

"Is the crude oil futures price an unbiased estimator of future spot price? What predictive implications can be drawn from the results?"

Finally, this thesis is structured in the following way: chapter 2 examines literature regarding unbiasedness of the futures price in the oil market, various forecasting methods used in the oil market and the findings in these studies.

Chapter 3 presents the theories aiming at explaining the relationship between the spot- and futures price of a commodity. In addition, this chapter also presents the efficient market hypothesis and the market risk premium in order to assess the findings in this thesis along with economic theory. Chapter 4 derives the regression and forecasting models used in the in- and out-of-sample analyses. Chapter 5 presents the variables suggested to be included in the multivariate forecasting model and the theoretical arguments for including them. Chapter 6 explains the methodology used. Chapter 7 explains the data collection process and examines the data. Chapter 8 discusses the results and the implications that can be drawn from these results. Lastly, chapter 9 concludes based on the results and discussion provided in chapter 8, while chapter 10 outlines the limitations of this analysis with suggestions for further research.

2. Literature Review

This chapter presents and examines the literature available on different forecasting methods used to forecast the future spot price of crude oil. In addition, it also examines the literature concerning efficiency- and unbiasedness in the oil futures market.

2.1 Forecasting based on the futures price

Forecasting methods based on the futures price are often characterized as financial forecasting models. These models use the futures price either directly or indirectly, through the basis[1], in order to predict the future spot price of crude oil. This means that the financial forecasting models also examine the unbiasedness of the futures price, and hence whether a risk premium is reflected in the futures price.

Zeng and Swanson (1998) use the random walk model, vector autoregressive models (VEC) and vector error correction models (VECM) to investigate the forecasting ability of futures prices on the spot price of the underlying. The study is based on the period 1990-1995 and examines, among others, the price of crude oil. The study finds a cointegrating relationship between the futures- and spot price of crude oil. This finding is supported by the fact that the study finds the VECM to possess superior forecasting ability for the spot price of crude oil relative to the other forecasting models tested in the short-run.

Chernenko et al. (2004) examine the unbiasedness and efficiency of futures contracts in the petroleum market. The study uses daily prices of WTI crude oil traded at NYMEX in the period of 1989-2003 to examine the crude oil market. The study finds that the futures price is not an unbiased estimator of the spot price nor is it efficient. In addition, the authors find only suggestive results supporting the existence of risk premium in the crude oil market. Lastly, the authors find that the random walk model outperforms the forecasting model based on the futures price. Another study examining the efficiency and unbiasedness of futures prices is conducted by Abosedra (2005). Abosedra examines the efficiency and

[1] $\text{basis}_t = F^{oil}_{t+h|t} - S^{oil}_t$

unbiasedness of crude oil spot and future prices using a univariate forecasting model. The study is based on data from 1991 to 2001 and uses monthly observations of the WTI spot- and futures price. The author finds evidence suggesting that the futures price is, in contrast to Chernenko et al. (2004), an unbiased estimator of the spot price of crude oil and that futures market forecasts are semi-strongly efficient.

Coppola (2008) investigates whether there are short-run and long-run relationships between the crude oil futures and spot price. The author examines these relationships using a cointegration test and a VECM. The study is based on weekly WTI crude oil prices in the period 1986-2006, and finds that there is a long-run relationship between crude oil futures price and spot price. The author therefore suggests that this relationship can be used to predict the future spot price of crude oil. This is later supported by a finding stating that VECM generates more accurate predictions of the spot price than the random walk model with no drift.

2.1.1 Time-varying risk premium

A time-varying risk premium can be translated into risk premium being predictable by a set of variables. Cochrane (2008) showed that it is possible to forecast stock returns (i.e. the risk premium) based on the dividend-price ratio. He argued that the market return could be predicted to some extent, but only based on factors reflecting the state of the economy.

For returns in the commodity futures market, Fama & French (1987) decompose the basis for a range of commodity futures into two information components; one for the futures return (interpreted as the risk premium) and another for the change in the spot price. Presuming market efficiency, the authors conclude that there is a time-variation in the risk premium for 5 of the 21 commodities tested. However, their commodities do not include energy futures.

Another study examining whether the risk premium in commodity futures market is time-varying is conducted by Deaves & Krinsy (1992a). The authors focus on the oil market and find a significantly positive risk premium for the periods August 1986-July 1987 and December 1988-April 1990, while they find a negative risk premium for the period August 1987-November 1988. However, as suggested by Frankel & Froot (1987), one possible explanation is that asset values may be consistently underpredicted when appreciating (resulting in positive futures returns) and overpredicted when depreciating (negative returns). Moreover, the authors also find that the futures return covaried with recent volatility, meaning risk premium may be related to changing market conditions.

Supporting the findings of Deaves and Krinsky (1992a), empirical tests by Moosa & Al-Loughani (1994) reveal that the futures price of crude oil is not an unbiased predictor of the spot price. They also show the existence of a time-varying risk premium based on a GARCH-M process. Brooks, Prokopczuk & Wu (2013), on the other hand, does not find any evidence of a time-varying risk premium in the crude oil futures market when regressing the realized risk premium against the basis.

Despite the extensive amount of research on forecasting methods using futures prices as a predictor of the spot price, there is no consensus on whether crude oil futures price is an unbiased and/or efficient estimator of the spot price. Some literature finds that there is a time-varying long-run relationship between the futures price and spot price of crude oil, indicating a time-varying risk premium. However, the existing forecasting models have not been able to accurately predict a risk premium.

2.2 Forecasting based on fundamental- and financial factors

Both fundamental and financial factors are used in structural forecasting models in order to predict oil price movements. The fundamental factors cover the supply and demand aspect of the market, while financial variables are argued to contain information about the expected future state and riskiness of the economy. This means that both fundamental and financial variables may contain information about the future spot price of crude oil which is not reflected in the futures price.

2.2.1 Forecasting based on inventory

Ye et al. (2002) use OECD inventory levels to perform short-run forecasts of the spot price of WTI crude oil. The authors argue that OECD petroleum inventory levels serve as a measure of the imbalance between demand and supply in the petroleum market and that this imbalance can be used as a signal to forecast changes in the spot price of crude oil. The authors therefore construct a forecasting model which uses three explanatory variables connected to the OECD inventory levels to forecast the spot price of crude oil. The study is based on a dataset from 1992 to 2001. The study finds that the forecasting model demonstrates a relatively solid forecasting ability of the spot price of crude oil in an in-sample setting. However, the study only focuses on forecasting at short horizons.

Ye et al. (2005) made a modification to the previous study from 2002. The new study only uses the deviation from OECD normal inventory level as the explanatory variable on the spot price of crude oil. In addition, the new study is conducted on an extended dataset from 1992-2003 and controls for outliers in the crude oil price[2]. The objective of the study is to forecast the 1-month ahead spot price of WTI crude oil. The study finds a relatively solid forecasting performance of the inventory model on the spot crude oil price relative to an autoregressive forecasting model serving as benchmark, both in- and out-of-sample.

A study by Merino and Ortiz (2005) builds on the findings in Ye et al. (2005) and extends the forecasting methodology, which can be divided into three steps. The first step is to forecast the spot price of crude oil using the inventory model constructed by Ye et al. (2005), where the deviations between forecasted prices and the actual prices are defined as a “price premium”. The second step is to investigate whether the price premium is Granger caused (Granger, 1969) by any oil market variables and/or financial prices. This step shows that speculation, OPEC spare capacity and the relative U.S. gasoline inventory level Granger cause the price premium. The third step consists of forecasting the spot price of crude oil using the inventory model constructed by Ye et al. (2005) through extending the model with the oil market variables (one-by-one). The study finds that out of the oil variables, only speculation and oil prices have a cointegrated relationship. This means that this is the only extended forecasting model tested. The forecasting

[2] Controls for the period affected by the terrorist attack 11. September 2001 and the period affected by changes in OPEC strategy in 1999.

model is tested on the same dataset as Ye at al. (2005) where the inventory model of Ye et al. (2005) serves as the benchmark model. The study shows that the extended forecasting model outperforms the benchmark model in the periods 1992-2001 and that the models perform similarly in the period 2001-2004.

The consensus in the studies examining the forecasting ability of inventory models is that these models produce semi-strong predictions of the spot price of crude oil in the short run. However, in the long run the forecasting ability of the models are weak. Including crude oil inventory in a multivariate model is therefore only expected to improve the forecasting ability at short horizons.

2.2.2 Forecasting based on other supply- and demand factors

Tang and Hammoudeh (2002) examine whether OPEC carried out a target zone strategy for the oil price in the period 1988-1999. The authors find evidence supporting this hypothesis and construct a model consisting of production quotas, inventory levels and expectations about the future price of oil to forecast future spot price of oil. The model is tested in an out-of-sample setting and results in a relatively solid forecasting performance.

Yang et al., (2002) construct a model consisting of fundamental variables argued to be determinants of the U.S oil price. The model consists of three factors; U.S. GDP, OPEC production and the demand elasticity for U.S oil. The analysis is conducted on a data set from 1975 to 2000 and uses monthly prices. The authors use a GARCH model and the results indicate that OPEC production and the oil price is negatively correlated. This study therefore finds suggestive evidence that the oil production level contains predictive power on the spot price of crude oil.

There are several other studies using more complex fundamental models in order to investigate determinants of the oil price and to forecast the future spot price of oil (Chevillon & Rifflart, 2009; Kaufmann, Dees, Gasteuil, & Mann, 2008; Mirmirani & Cheng Li, 2004). Overall, the literature on fundamental models indicates that fundamentals do contain some predictive power on the spot price of oil. However, the fundamental models are found to be outperformed by econometrical models based on the futures price. This finding suggests that a multivariate model including both the futures price and fundamental factors may prove to be a useful topic of investigation.

2.2.3 Forecasting based on product spread

In contrast to the models mentioned, only a fraction of previous research focuses on the dynamic relationship between crude oil and product spreads. An even lesser amount focuses on forecasting based on this relationship. Some examples of research on the long-term relationship between crude oil and refined oil products are Paulson (1999) and Gjolberg & Johnsen (1999). The former finds a cointegrating relationship between the crude oil price and different product spreads. The authors also identify arbitrage possibilities when the product spreads are at their extremes. Furthermore, Gjolberg & Johnsen (1999) test the cointegrating relationship between crude oil and six different petroleum products and conclude with stationary spreads for five of them. However, none of the studies mentioned evaluate any forecasting models based on the spread. They however indicate possible forecasting abilities using crude margin deviations.

Only a few studies which use product spreads to forecast the spot price of crude oil were found when researching for this thesis. One of these is a study by Murat & Tokat (2009), which conducts a one-week ahead analysis of the forecasting ability of the 3:2:1 crack spread futures on the WTI crude oil price. The 3:2:1 crack spread is defined as three crude oil futures contracts (long), two gasoline futures contracts (short) and one heating oil futures contract (short). This ratio is meant to replicate a typical refiner’s yield (CME Group, 2017). Further, this study is based on a dataset between 2000 and 2008. The authors find a unidirectional relationship from crack spread futures on the spot oil market both in the long- and short-run based on a VEC framework. The forecasting method outperforms the random walk model and performs nearly as good as the forecasting method based on oil price futures. However, as Baumeister, Kilian & Zhou (2013b) point out, Murat & Tokat (2009) only test the predictive power within-sample, and not out-of-sample.

Baumeister et al. (2013b) derive several alternative forecasting models in a comprehensive study based on product spreads. The study shows that not all models are accurate in an out-of-sample forecasting setting. However, they find a number of models that accurately predict the WTI and Brent oil price between one- and two-year horizons. Among the single-spread forecasting models[3] they find that the gasoline spot spread alone is the most accurate. For the 3:2:1 crack spread model, they find no evidence for forecasting ability.

The consensus in the studies examining the forecasting ability of futures product spread models is that these models generate mediocre predictions of the spot price of crude oil at short horizons. In addition, the predictive accuracy of these models is weaker than in the forecasting models based on futures prices.

2.3 Forecasting using a multivariate forecasting approach

Only a handful of studies look into a multivariate forecasting approach combining the futures price, fundamental and financial factors in order to predict the future spot price of crude oil. One study examining the forecasting ability of a multivariate model is conducted by Baumeister et al., (2014). The study includes the oil futures price, a time-varying parameter of the gasoline- and heating oil product spreads, cumulative changes in the oil inventory and non-oil commodity prices in a multivariate model in order to predict the spot price of oil. The authors find that such pooled forecasts often, but not always, have lower MSPE than the best individual forecast model. One reason may be that a pooled forecast gives insurance against failures of the individual forecasts. (Baumeister et al., 2014).

Baumeister & Kilian (2015) later extend the approach in a new study by weighing six models according to previous accuracy of the models. This study excludes the inventory model and includes a no-change forecast and a model based on the simple product spread between gasoline price and crude oil price. The study finds that this extended approach is less accurate compared to Baumeister et al. (2014).

Westgaard et al., (2017) investigate a multivariate forecasting approach on the spot price of crude oil in the period February 2000 to June 2013. The authors include fundamental, financial, price shock and factors for political proxies to construct three multivariate forecasting models. Further, the authors apply a

[3] Either using gasoline or heating oil

general-to-specific model selection approach in order to specify three models of interest. In an out-of-sample setting the study finds that the most parsimonious model, which includes only financial factors outperforms the other models in terms of MSPE. In addition, all of the multivariate models produce more accurate predictions of the spot price than the benchmark models[4]. The results from this study suggest that multivariate models may contain additional information about the future spot price of crude oil compared to the univariate futures models. However, multivariate models may also introduce an overfitting-problem which should be taken into account.

The literature on multivariate forecasting models documents that these models outperform the univariate futures models in an out-of-sample setting. This finding indicates that a potential risk premium in the crude oil futures market may be time-varying. This finding also suggests that fundamental, financial and other relevant factors may contain information which is not reflected in the futures price and should thus be included in a multivariate forecasting model.

[4] No-change model and ARIMA (2, 0, 2)

3. Theory

This chapter examines the relationship between the spot and futures price of a commodity and presents two theories attempting to explain this relationship. Furthermore, this chapter also examines the market efficiency hypothesis (Fama, 1970; Jensen, 1978) and theory concerning the market risk premium.

Before presenting the formulas and models used in this thesis, let the following denotations apply for the rest of the paper: S_t^{oil} denotes the spot price of crude oil at time *t*. $\hat{S}_{t+h|t}^{oil}$ denotes the predicted spot price of crude oil in *h* periods contingent on information available at time *t*. $F_{t+h|t}^{oil}$ denotes the futures price of crude oil maturing in *h* periods contingent on information available at time *t*. $\mathbb{E}_t[S_{t+h}^{oil}]$ denotes the expected spot price of crude oil in *h* periods contingent on information available at time *t*. $F_{t+h|t}^{refined\ prod}$ denotes the futures price of a refined petroleum product in *h* periods contingent on information available at time *t*.

3.1 Futures contract

In order to investigate the relationship between the spot- and futures price, a definition of a futures contract is required. A futures contract is defined as "*an agreement to buy or sell an asset at a certain time in the future at a certain price*" (Hull, 2017, p. 30). Furthermore, a futures contract is a standardized contract which is normally traded on an exchange (Hull, 2017, p. 30).

The common view in financial markets is that the futures price is the best predictor of the future spot price of crude oil. This implies that a univariate forecasting model, with the futures price as the only predictor of the spot price of crude oil, is the most used forecasting model in practice (Alquist & Kilian, 2010). Furthermore, when using the futures price to predict the spot price of crude oil different restrictions can be made regarding the existence and potential form of the risk premium in the crude oil futures market. Testing the forecasting ability of the futures price in its simplest form implies an assumption of no risk premium and originates from the *unbiased expectations hypothesis* (Fisher, 1896).

3.1.1 The unbiased expectations hypothesis

The unbiased expectations hypothesis states that the futures price, with maturity *h*, is an unbiased estimator of expected future spot price of the underlying at the period of maturity, $t + h$ (Fisher, 1896; Hicks, 1939; Lutz, 1940). This implies that:

$$F^{oil}_{t+h|t} \approx \mathbb{E}S^{oil}_{t+h|t} \tag{3.1}$$

The unbiased expectations hypothesis assumes that investors are risk neutral. This assumption is later challenged as the existence of risk premium in the crude oil futures market is tested. Moreover, in order to introduce the risk premium, the relationship between the spot and futures price of the underlying needs to be examined.

3.2 The basis

The *basis* is defined as the difference between the futures price and the contemporaneous spot price of the same underlying (Fama & French, 1987):

$$basis_t = F^{oil}_{t+h|t} - S^{oil}_t \tag{3.2}$$

There are two recognized theories which break down and attempt to explain the dynamic relationship of the basis: *the theory of storage* (Brennan, 1958; Kaldor, 1939; Telser, 1958; Working, 1948) and *the theory of risk premium* (Breeden, 1980; Cootner, 1960; Dusak, 1973; Keynes, 1930).

3.2.1 The theory of storage

The theory of storage is the first recognized theory which describes the relationship between the futures and the spot price of a commodity (i.e., the basis). The theory states that the basis is determined by interest foregone related to holding the commodity, $r_{t,t+h}$, storage costs related to storing the commodity, $\lambda_{t,t+h}$, and a marginal convenience yield on inventory, $c_{t,t+h}$ (Brennan, 1958; Kaldor, 1939; Telser, 1958; Working, 1948). Further, the theory argues that the basis is determined by a no-arbitrage condition and that the no-arbitrage condition depends on whether the futures price is above or below the contemporaneous spot price. When the futures price is below the contemporaneous spot price (negative basis), the theory of storage states that this difference can be explained by interests foregone by holding the commodity, $r_{t,t+h}$, storage costs by holding the commodity, $\lambda_{t,t+h}$, and a marginal convenience yield, $c_{t,t+h}$. This implies that:

$$F_{t+h|t}^{oil} = S_t^{oil} * e^{r_{t,t+h}+\lambda_{t,t+h}-c_{t,t+h}} \quad (3.3)$$

where the marginal convenience yield attributes the value of having available inventory to meet unexpected demand (Fama & French, 1987). This approach describes the no-arbitrage lower bound for the futures price in relation to the contemporaneous spot price. If the futures price is below this bound, one could lease the commodity and sell it (spot), invest the proceeds risk-free and buy futures contracts at time *t*, and then reverse the positions at time *T*, earning an arbitrage profit.

Conversely, the upper bound for the futures price is derived from the opposite strategy: borrowing risk-free, buying commodity (spot), selling futures contracts and holding the commodity to delivery at time *T*. With this strategy, the holder of the commodity does not miss out on the convenience yield and hence the upper bound is determined by:

$$F_{t+h|t}^{oil} = S_t^{oil} * e^{r_{t,t+h}+\lambda_{t,t+h}} \quad (3.4)$$

Furthermore, numerous studies (Brennan, 1958; Fama & French, 1988; Telser, 1958) have shown an inverse relationship between the convenience yield and inventories (implied by the basis). This finding is consistent with theory. However, the “problem” with the theory of storage and the convenience yield is that the convenience yield is only a theoretical value and it is unobservable. Therefore, the theory of storage and the concept of convenience yield is, though highly accepted in theory, not very useful for forecasting spot prices nor testing market efficiency.

3.2.2 The theory of risk premium

While the theory of storage describes the upper and lower bounds of the basis, the risk premium theory states that the difference between *expected* spot price and the futures price[5] is equal to a risk premium (Cootner, 1960; Keynes, 1930). This implies that:

[5] I.e. the expected return from a long futures position

$$F_{t+h|t}^{oil} = \mathrm{e}^{-\gamma_{t,t+h}} * \mathbb{E}_t[S_{t+h}^{oil}] \tag{3.5}$$

where $\gamma_{t,t+h}$ is the (unobservable) risk premium. The challenge with this theory is that there is no consensus on the size, or even the sign of the risk premium for commodity futures.

The traditional view is that the sign of the risk premium depends on whether there are net short- or net long hedgers in the market. The former describes a situation where most hedgers possess the commodity and wish to protect themselves from price fluctuations by selling futures contracts. The speculators on the other side of the contract demand a reward for taking that risk. Thus, the futures price should be trading below the expected spot price, i.e. at “backwardation” (Keynes, 1930). The latter describes the opposite, where most hedgers are consumers and speculators sell the futures. In this case, the risk premium should be negative and futures thus trades above expected spot price, i.e. at “contango” (Cootner, 1960).

Furthermore, as stated in the literature review there is no clear consensus on the form of the risk premium in the crude oil futures market. A constant risk premium in the crude oil futures market entails that the futures price is a biased predictor of the underlying. However, the futures price may still be rational and should thus be the only predictor necessary to predict the spot price of the underlying. This motivates the use of a univariate model to examine the crude oil market. However, some literature suggests that the risk premium in the oil futures market is time-varying (Deaves & Krinsky, 1992a; Moosa & Al-Loughani, 1994). If the risk premium in the crude oil futures market is in fact time-varying, other variables are expected to influence the spot price of crude oil through the risk premium. This motivates the use of a multivariate model in order to examine the crude oil futures market.

3.3 Market efficiency

The efficient market hypothesis (EMH) states that investors are not able to systematically earn abnormal returns[6] (Baumeister & Kilian, 2012; Merino &

[6] Abnormal returns are returns that are above what is justified by the risk of the asset

Ortiz, 2005; Ye et al., 2002). According to Jensen, "*a market is efficient with respect to information set* θ_t *if it is impossible to make economic profits by trading on the basis of information set* θ_t" (1978). The EMH therefore implies that all relevant information is immediately reflected in the asset prices and stable forecasting patterns should not exist for longer periods when discovered by investors (Fama, 1970).

Testing market efficiency for commodity futures prices is not unproblematic as it requires a joint hypothesis test. This means that one cannot distinguish whether realized futures return ($S_{t+h}^{oil} - F_{t+h|t}^{oil}$) is a result of a risk premium, market inefficiency or a combination of both. The implication of this is that one can only examine whether the EMH holds under the assumption of no risk premium or by correctly predicting the risk premium. Furthermore, it is generally insufficient to assume no risk premium when examining whether the EMH holds (Timmermann & Granger, 2004). Therefore, in order to correctly conclude on market efficiency or inefficiency, one must compare the forecasting results against models that are correctly specified in relation to the assumption of risk premium (i.e. how large it is, and whether it is constant or not). In this thesis, two asset pricing models are used to complement the estimation of the risk premium conducted with the commodity pricing models. These two asset pricing models are; the Capital Asset Pricing Model (CAPM) (Lintner, 1965; Sharpe, 1964) and the Fama-French Three-Factor model (Fama & French, 1993).

Another relevant factor that affects the test of market efficiency is transaction costs. High transactions costs limit the possibility to exploit arbitrage opportunities as such opportunities may be costly. Predictability should therefore be seen in relation to the transaction costs. For commodity futures, the transaction costs consist mainly of the bid-ask spread and exchange fees (CME Group, 2018). However, for oil futures, these costs are argued to be small, especially for shorter maturities with high liquidity. This statement is supported by Deaves & Krinsky (1992b) who estimate the transaction costs to represent approximately 1 % of the futures price. Due to lack of available data on transaction costs, and the fact that transaction costs are relatively small for crude oil futures, this thesis does not include transaction costs in the analysis. Furthermore, excluding transaction costs has been the most common approach in previous empirical studies.

3.4 Market Risk Premium

Cortazar et al. (2015) argue that commodity pricing models are not able to provide reliable estimates of the risk premium due to the fact that the distribution of spot price is unreliable. The authors therefore suggest that asset pricing models could be used to estimate the risk premium. Following the argumentation of Cortazar et al. (2015) the CAPM and the Fama-French three-factor model is used to complement the estimation of the risk premium in this thesis. These two models capture the sources for systematic risk and assume that the unsystematic risk is zero due to diversification. This implies that the futures risk premium will emerge in the context of a well-diversified portfolio.

3.4.1 CAPM

The CAPM model distinguishes between diversifiable and non-diversifiable risk, where the latter is reflected in the beta (price of risk). The expected return on a long futures position can therefore be specified as follows (Deaves & Krinsky, 1992a):

$$\mathbb{E}S_{t+h|t}^{oil} - F_{t+h|t}^{oil} = \beta(\mathbb{E}RM_{t+h|t} - RF_t) \tag{3.6}$$

where $\mathbb{E}RM_{t+h|t}$ is the expected return at t of the market portfolio and RF_t is the risk-free rate. β is the coefficient between the market risk premium and expected futures return measuring how sensitive the futures return is relative to the market risk premium. Therefore, the right-hand side is simply another specification of the risk premium.

An article which undertakes whether speculators in a commodity futures market earn a risk premium is written by Dusak (1973). Dusak examines whether there is a risk premium in the futures price of wheat, corn and soybeans using the capital asset pricing model. The author finds that the returns of investments in the futures contracts are very close to zero, consistent with the CAPM, as the commodity futures prices have no systematic risk. However, a study by Bodie & Rosansky (1980) finds negative betas for the commodity futures, while the returns are similar to the returns on common stocks. One can conclude that these results are not consistent with the CAPM.

3.4.2 Fama-French Three-Factor Model

The Fama-French three-factor model is an extension of the CAPM model. The multifactor model includes two additional factors which are considered to be relevant sources of systematic risk, SMB and HML respectively (Bodie, 2014; Fama & French, 1993). The first additional factor, SMB, stands for Small-Minus-Big and measures the excess return of a portfolio consisting of small stocks relative to a portfolio of large stocks. This factor is included because small stocks indicate a higher systematic risk and hence should result in higher expected returns. The second additional factor, HML, stands for High-Minus-Low and measures the excess return of a portfolio consisting of high book-to-market ratio stocks relative to a portfolio of low book-to-market ratio stocks. This factor is included due to the fact that companies with high book-to-market ratios are considered more likely to experience financial distress and hence possess a higher systematic risk resulting in higher expected return. (Bodie, 2014, p. 340).

The Fama-French three-factor model in terms of a long futures position can therefore be specified as follows:

$$\mathbb{E}S_{t+h|t}^{oil} - F_{t+h|t}^{oil} = \beta_1(\mathbb{E}RM_{t+h|t} - RF_t) + \beta_2\text{SMB} + \beta_3\text{HML} \qquad (3.7)$$

where the β-values measure the excess return's sensitivity towards the different factors. Furthermore, both the CAPM model and the Fama-French three-factor model are computed in order to examine whether the estimation of the risk premium using commodity pricing models is reliable and to assess whether the crude oil futures price is rational. The results from the CAPM and the Fama-French three-factor model computations are presented and discussed in chapter 9.

4. Model specification

This chapter derives and presents the regression models used to examine the unbiasedness of the futures price in the in-sample analysis. This chapter also presents the forecasting models used in the out-of-sample analysis.

4.1 In-sample analysis

As stated in the theory section, there are two popular theories attempting to explain the relationship between the futures and spot price of a commodity: the *Theory of Storage* and the *Theory of Risk Premium*. The regression models used in the in- and out-of-sample analyses are based on the theory of risk premium, as the objective of the thesis is to examine the unbiasedness of the futures price. Furthermore, this involves examining the existence and potential form of a risk premium in the crude oil futures market.

The in-sample analysis is conducted by running both a univariate- and a multivariate regression model. The univariate model examines the unbiasedness of the futures price, while the multivariate model examines whether the variables included in the model do contain predictive power on the spot price of oil through the risk premium.

4.1.1 Univariate model

Equation (4.1) presents the theory of risk premium in discrete terms. This is the starting point of deriving the regression model used in the in-sample analysis:

$$F_{t+h|t}^{oil} = \mathbb{E}_t\left[S_{t+h}^{oil}\right] - \gamma S_t^{oil} \tag{4.1}$$

Dividing both sides with the contemporaneous spot price[7] and substituting the expected spot price with realized next-period spot price results in the following regression model:

$$\frac{S_{t+h}^{oil}}{S_t^{oil}} = \gamma + \beta \frac{F_{t+h|t}^{oil}}{S_t^{oil}} + \varepsilon_{t+h} \tag{R.1}$$

where γ measures a constant risk premium and β measure a time-varying component of the risk premium.

[7] In order to normalize the futures price and next period spot price

4.1.2 Multivariate model

A time-varying risk premium implies that the risk premium may depend on a set of variables in addition to the futures price. This means that the time-varying risk premium can be derived from the following model:

$$\begin{aligned} \gamma_{t+h} &= \gamma_0 + \gamma_1 \frac{F^{oil}_{t+h|t}}{S^{oil}_t} + \gamma_2 x_{2,t} + \cdots + \gamma_n x_{n,t} + v_{t+h} \\ &= \gamma_0 + \gamma_1 \frac{F^{oil}_{t+h|t}}{S^{oil}_t} + \sum_{k=2}^{n} \gamma_k x_{k,t} + v_{t+h} \end{aligned} \tag{R.2}$$

where all variables influencing the risk premium are included. Substituting γ in (R.1) with the time-varying risk premium, γ_{t+h}, in (R.2), results in the following regression model:

$$\frac{S^{oil}_{t+h}}{S^{oil}_t} = \gamma_0 + \beta_1 \frac{F^{oil}_{t+h|t}}{S^{oil}_t} + \sum_{k=2}^{n} \gamma_k x_{k,t} + u_{t+h} \tag{R.3}$$

where $\beta_1 = 1 + \gamma_1$ and $u_{t+h} = v_{t+h} + \varepsilon_{t+h}$. n is the number of variables and γ_k is the time-varying component of the risk premium related to a set of variables. The multivariate model is regressed in order to examine whether the explanatory variables included in the model are significantly different from zero, and thus whether the variables influence the spot price of crude oil through the risk premium. Furthermore, chapter 5 presents and examines the variables that are suggested to be included in the multivariate model in this thesis.

4.2 Out-of-sample analysis

The in-sample analysis assesses the unbiasedness and forecasting efficiency of the crude oil futures market and provides some expectations regarding which forecasting model can possess the highest predictive accuracy. However, the fact that a risk premium might be predictable in an in-sample setting does not imply that a risk premium is predictable in an out-of-sample setting (Welch & Goyal, 2008). This means that the in- and out-of-sample analyses might give conflicting results.

4.2.1 Benchmark model

A natural benchmark model for the forecasting models on the spot price of crude oil is the *random walk model with no drift* (Alquist & Kilian, 2010; Baumeister, Kilian, & Zhou, 2013a; Murat & Tokat, 2009). This model implies that it is impossible to predict changes in the spot price of the crude oil, meaning that the best predictor of the future spot price of crude oil is the current price of crude oil:

$$\hat{S}^{oil}_{t+h|t} = S^{oil}_{t} \tag{F.1}$$

The random walk model is used as a benchmark model in order to decide whether alternative forecasting methods produce sufficient accuracy. Accordingly, the proposed forecasting models should at least produce as accurate predictions as the random walk model in order to be of interest.

4.2.2 Simple futures model

The unbiased expectations hypothesis states that the futures price is an unbiased estimator of the expected future spot price of the underlying and that all relevant information is reflected in the futures price. This translates into the following forecasting model:

$$\hat{S}^{oil}_{t+h|t} = F^{oil}_{t+h|t} \tag{F.2}$$

Consequently, this model assumes no risk premium.

4.2.3 Univariate forecasting model

This forecasting model assumes a risk premium that is only reflected through the futures price without additional variables. This is tested through running the following model;

$$\hat{S}^{oil}_{t+h|t} = S^{oil}_{t}\left(\hat{\gamma} + \hat{\beta}\frac{F^{oil}_{t+h|t}}{S^{oil}_{t}}\right) \tag{F.3}$$

where $\hat{\gamma}$ is the estimated constant risk premium and $\hat{\beta}$ is the estimated time-varying component of the risk premium. In order to conduct forecasts using forecasting model (F.3), the $\hat{\gamma}$ and $\hat{\beta}$ coefficients are obtained by running regression model (R.1).

4.2.4 Multivariate forecasting model

The following forecasting model assumes a time-varying risk premium and includes n variables assumed to be related to it. This results in the following forecasting model:

$$\hat{S}_{t+h|t}^{oil} = S_t^{oil}\left(\hat{\gamma}_0 + \hat{\beta}_1 \frac{F_{t+h|t}^{oil}}{S_t^{oil}} + \sum_{k=2}^{n} \hat{\gamma}_k x_{k,t}\right) \tag{F.4}$$

where $\hat{\gamma}_0$ is the estimated constant risk premium, $\hat{\beta}_1$ and $\hat{\gamma}_k$ are estimated time-varying components of the risk premium related to the futures price and the proposed set of variables, respectively. These coefficient estimates are obtained by running regression model (R.3).

5. Factor Identification

This chapter presents the variables that are suggested to be included in the multivariate model and the theoretical arguments for including them.

5.1 Fundamental factors

Fundamental variables have been used in numerous studies in order to model oil price movements and to make predictions about the future price of oil (Merino & Ortiz, 2005; Westgaard et al., 2017; Ye et al., 2002, 2005). The fundamental factors that are suggested to be included in the multivariate model in this thesis are: crude oil inventory, crude oil production and rig activity. These variables are argued to contain predictive power on the spot price of crude oil due to the fact that these variables contain information about the supply and demand aspect of the market which may not be reflected in the futures price.

5.1.1 Inventory

There are mainly two theoretical arguments for using inventory levels to model oil price movements and to predict the spot price of crude oil. The first argument builds on the theory of storage and the negative relationship between inventories and convenience yield[8]. This argument is explained in detail in the following paragraph. The second argument builds on the idea that changes in crude oil inventories have a signaling effect of the supply and demand imbalance in the oil market. This signaling effect is argued to be an indicator of market pressure on price changes and is thus argued to contain predictive power on the spot price of crude oil. As this thesis rather focuses on the variables' theoretical relationship with the risk premium, this argument is left in exhibit 1.

According to the theory of storage, the spot price has a negative relationship with inventories through the convenience yield. This implies that as inventories increase, the convenience yield should decline as stock-out is less likely (Brennan, 1958). Subsequently, this reduces (in absolute terms) the lower bound of the basis[9]. This should require either an increase in the futures price or a decrease in the spot price. For example, refineries could increase their inventories because

[8] Thus, a negative relationship between inventory and the risk premium.
[9] Decreasing the possible size of backwardation

they expect the oil price to increase in the future, which is reflected through increased futures prices and reduced risk premium.

Both arguments suggest that crude oil inventory levels contain predictive power on the spot price of crude oil. Inventory level is therefore included in the multivariate model. However, inventory levels are only expected to contain predictive power on the spot price of crude oil in the short-run as inventory and production is expected to be more adjustable in the long-run.

5.1.2 Oil production and rig count

Supply variables such as oil production and rig activity are suggested to possess predictive power on the spot price of crude oil. These two variables are proposed to have a negative relationship with the crude oil price due to the fact that increased supply puts downward pressure on the crude oil price as shortage becomes less likely (Möbert, 2007). Subsequently, this reduces the convenience yield. By including these two supply variables in the multivariate model, one might be able to capture information that affects the oil price which is not already reflected in the inventory (assuming some lag between changes in production and changes in inventory) and in the futures price. However, the effects of production and rig activity can be ambiguous as it may also positively react to oil price increase (Ringlund, Rosendahl, & Skjerpen, 2008) or increased demand.

Similar to the inventory factor, these variables are regarded only as a short-term indicator. This implies that these variables are only expected to have predictive power on the spot price of crude oil in the short-run.

5.2 Financial factors

Several financial- and economic factors have been suggested to influence the spot price of crude oil through the risk premium in the futures market (Westgaard et al., 2017). Financial and economic factors are argued to possess information about the expected state of the economy and market risk which might influence the spot price of crude oil. The financial- and economic factors included in the multivariate model are presented below.

5.2.1 Market Risk Premium

The excess return of a world market portfolio is by definition the market risk premium according to CAPM (Lintner, 1965; Sharpe, 1964). Moreover, the risk premium of the stock market may affect the risk premium for commodities, by inducing investors to require higher reward for taking risk in the oil futures market. This implies that the market risk premium might contain predictive power of the spot price of crude oil, and this variable is therefore included in the multivariate model. As a proxy for the world market excess return, the return from the S&P 500 index minus the return of the 3-month US Treasury bill is used.

5.2.2 Change in slope of the yield curve

Another financial factor that might be related to the risk premium and risk aversion of the economy is the yield curve of government bonds. If the yield curve is positive (increasing yield as bond horizon increases), it may indicate an expansion in economic activity. If the yield curve is negative, it may indicate a declining interest rate, which in turn is often interpreted as a signal for a coming recession (Bodie, 2014, p. 503). However, one should note that this depends on the assumptions of liquidity premium as the yield for bonds with longer horizons may include a premium required by investors for holding longer-term bonds (positive premium), or the other way around for short term bonds (negative premium).

In general, a downward sloping yield curve is taken as a strong signal that the interest rate is more likely than not to fall, due to the yield curve being normally upward sloping. Thus, a decrease in the slope of the yield curve may also increase the riskiness of oil, and vice versa. To indicate the steepness of the yield curve, the yield for a 10-year US government treasury bond minus the yield of a 3-month US Treasury bill is used.

5.2.3 Change in the credit spread

The difference between the yield on a corporate bond and a government bond is called the credit spread. The credit spread reflects the premium that investors demand for taking credit risk, which is driven by the risk of default and its related loss for the investor (Collin - Dufresn, Goldstein, & Martin, 2001). A change in this spread should reflect the change in the risk aversion of the economy or changes of the economic environment (Nai-Fu, Roll, & Ross, 1986). A decrease in the credit spread could be a result of reduced risk aversion or that investors perceive the general economy to be stable, and vice versa. If these changes also affect the risk aversion towards the oil price (or the riskiness of it), this factor should be related to the risk premium reflected in the oil futures. For the corporate bond yield, Moody's Baa-rated corporate bond yield is used as it should be more sensitive to the economic environment than high-grade bonds.

5.2.4 Volatility of the oil price

As the economic variables may not capture the entire risk of the oil price, the volatility of the oil price itself is included in the multivariate model. A higher volatility is expected to increase the risk premium, as investors may find the oil price being riskier, thus requiring a higher compensation for the risk. Few studies have investigated the relationship between realized volatility and the oil price. However, this has been widely studied in the stock market, where several studies provide some evidence of a positive relationship between expected risk premiums and volatility (French, Schwert, & Stambaugh, 1987; Suss, 2009). Bollerslev, Tauchen & Zhou (2009) find evidence implying that the difference between implied and realized variances can be used to predict future stock returns. For this thesis, the realized variance of daily oil futures prices (with 1-month maturity) is used as a measure for oil price volatility due its high trading volume.

5.2.5 Futures product spread

Futures product spread is defined as the difference between the futures price of a refined petroleum product and the contemporaneous spot price of crude oil:

$$futures\ product\ spread_t = F_{t+h|t}^{refined\ prod} - S_t^{oil} \tag{5.1}$$

The theoretical arguments for the predictive power of futures product spread on the spot price of crude oil originates from two theories; the *Verleger hypothesis* (P. K. Verleger, 1982) and a proposition stating that *convenience yield is increasing in marginal production costs* (Heinkel, Howe, & Hughes, 1990). The argument based on the Verleger hypothesis is explained in exhibit 2.

The second argument builds on a proposition by Heinkel, Howe & Hughes (1990). The authors propose a positive relation between convenience yield and marginal production costs. According to this proposition, oil refineries will respond to unexpected demand in one of the two following ways: (1) when the marginal production costs are low, the refineries will respond by increasing their production, consuming more crude oil (reflecting low convenience yield) or (2) when the marginal production costs are high, the refineries will respond by selling their inventories of finished refined petroleum products, consuming less crude oil (reflecting high convenience yield). They thus argue that increased marginal production costs cause the convenience yield to increase as it reduces the advantage of holding the commodity.

Moreover, presuming that variations in marginal production costs are mainly attributed to variations in the cost factors (Edwards, 1992), there may be a negative relationship between the convenience yield and the product spread. An empirical study conducted by Kocagil (2004) supports this suggested negative relationship. As the negative basis (backwardation) should increase as the convenience yield increases[10], the spot price should increase. Therefore, this proposition implies that the product spread is related to the risk premium. In line with this rationale, there is evidence of a long-run equilibrium relationship between the product spread and spot price of oil (Gjolberg & Johnsen, 1999; Hankyeung, David, & Kunlapath, 2015; Murat & Tokat, 2009). This long-run relationship suggests that futures product spread might have predictive power on the spot price of crude oil and therefore should be included in the multivariate model.

5.3 Overview of all variables included in the multivariate model

Table 1 provides an overview of the explanatory variables suggested to be included in the multivariate model. In order for suggested variables to be included in the multivariate model, the variables have to satisfy the Ordinary Least Square (OLS) assumptions and the stationarity condition. This is explained in detail in the next chapter.

[10] Following the rationale of the theory of storage

Variable	**Formula**	**H.**	**Source**
Fut	$F_{t+h\|t}^{oil}/S_t^{oil}$	+	EIKON
dInv	inv_t/inv_{t-h}	−	EIKON
dProd	$prod_t/prod_{t-h}$	−	EIKON
dRig	$rigs_t/rigs_{t-h}$	−	Baker Hughes
Mkt	$R_{t-h,t}^{S\&P} - R_{t-h,t}^{Tbill}$	+	EIKON / Fed.
dSlope	$\left(R_t^{Tbond} - R_t^{Tbill}\right) - \left(R_{t-h}^{Tbond} - R_{t-h}^{Tbill}\right)$	−	Federal Reserve
dCred	$\left(R_t^{corp} - R_t^{Tbond}\right) - \left(R_{t-h}^{corp} - R_{t-h}^{Tbond}\right)$	+	Federal Reserve
RV	$\frac{\sum_{i=1}^{n}\left(F_i^{oil} - \bar{F}^{oil}\right)^2}{n-1}$	+	EIKON
Gas	$F_{t+h\|t}^{gasoline}/S_t^{oil}$	+	EIKON
Heat	$F_{t+h\|t}^{heating}/S_t^{oil}$	+	EIKON
Crack	$\left(\frac{1}{3}F_{t+h\|t}^{heating} + \frac{2}{3}F_{t+h\|t}^{gasoline}\right)/S_t^{oil}$	+	EIKON

Table 1: overview of potential regression variables for h-months horizon. RV is the realized variance for the period from $t-h$ to t using daily futures prices of WTI (one-month contracts). Second last column shows hypothesized relation with the risk premium, and thus the spot price.

6. Methodology

This chapter presents the methodology used when examining the data, conducting the in- and out-of-sample analyses and assessing the accuracy of the forecasting models.

6.1 OLS

In order to conduct both the in- and out-of-sample analyses, the Ordinary Least Squares (OLS) method is used to run the necessary regressions. The OLS method builds on a set of assumptions which must be satisfied in order to obtain reliable coefficient estimates (Wooldridge, 2016, p. 40). All these assumptions are presented in exhibit 3.

6.1.1 Correlation matrix

One of the OLS assumptions state that the explanatory variables should have "*no linear dependence*" (Wooldridge, 2016, p. 74)[11]. This assumption means that an explanatory variable should not be a perfect linear function of other explanatory variables. If this is the case, there is *perfect multicollinearity* in the variables and the OLS method cannot estimate the model. Moreover, if any explanatory variables are highly correlated there is a case of *strong multicollinearity*. Strong multicollinearity does not entail that the OLS method cannot estimate the model. However, it means that the OLS method produces estimates that are sensitive to changes in the model and estimates with high variance. This makes the coefficient estimates unstable and difficult to interpret (Wooldridge, 2016, p. 84). A correlation matrix (for each horizon) is therefore computed in order to examine whether the suggested explanatory variables are perfectly- or highly correlated. If any of the explanatory variables are perfectly- or highly correlated the variable is excluded from the multivariate model.

6.2 Testing for stationarity

A special assumption for time series regressions is that the data has to be stationary. If the data is non-stationary, one might end up with a spurious

[11] No perfect multicollinearity

regression, i.e. inflated R^2 and t-ratios, which will influence the coefficient estimates. Therefore, in order to examine the data for stationarity, an augmented Dickey-Fuller test (ADF)[12] (Dickey & Fuller, 1979) is conducted. In the ADF-test the null- and alternative hypotheses are given by:

$$H_0: y_t \sim I(1) \qquad\qquad H_1: y_t \sim I(0)$$

where if y_t is integrated of order one, $I(1)$, the data is non-stationary. If the null hypothesis is rejected, the timeseries is stationary. Further, the Bayesian Information Criteria (Schwarz, 1978) is used to determine the optimal lag length used in the ADF-test. However, the ADF-test is criticized for failing to reject the null when the unit root is close to being non-stationary (Brooks, 2014, p. 364). Therefore, a visual inspection of the time series is also conducted.

6.3 In-sample analysis

The in-sample analysis is conducted during two different time periods. The first time period starts in January 1986 and ends in April 2019, while the second time period starts in January 2006 and ends in April 2019. The reason for conducting the in-sample analysis during two different time periods is to examine whether the findings in the time periods are consistent or whether the findings vary across sub-periods.

The unbiasedness of the futures price is tested by running regression model (R.1) using OLS and testing the joint null- and alternative hypotheses given by:

$$H_0:\ \gamma = 0 \ \&\ \beta = 1 \qquad\qquad H_1:\ \gamma \neq 0 \text{ or } \beta \neq 1$$

The null hypothesis states that there is no risk premium in the futures price and that the futures price is an unbiased predictor of the spot price of crude oil. The alternative hypothesis states that the futures price is a biased estimator of the future spot price. Under the assumption that futures prices are rational, this also

[12] The ADF-test tests for unit roots

indicates a constant and/or time-varying risk premium in the futures prices. This joint hypothesis is tested using the Wald test (described in exhibit 4).

In addition, the risk premium estimated with the commodity pricing model is compared with the expected futures return calculated by the CAPM model and the Fama-French three-factor model. Comparing the results of the models enables one to examine whether the futures return is justified and whether the futures price is rational according to asset pricing theory.

6.4 Out-of-sample analysis

The out-of-sample analysis is structured similarly to the in-sample analysis using the same two time periods. When the forecasting models are tested during the first time period, the univariate and multivariate forecasting models use the first 300 observations as a burn-in period. This is in order to generate solid coefficient estimates. Moreover, these coefficient estimates are estimated using recursive regression and results in approximately 100 forecasts which the models are assessed upon[13].

In the second time period the forecasting model uses the first 60 observations in the initial burn-in period to generate coefficient estimates, while the out-of-sample period is kept unchanged at 100 forecasts. A larger burn-in period is sacrificed to treat the out-of-sample period consistently. This means that the forecasting performance of the models in the two chosen time periods are assessed upon the same data (as the out-of-sample periods are identical). The only difference is that the models in the restricted data set uses less historical data to generate coefficient estimates compared to the models tested on the full data set.

6.5 Model evaluation

The mean squared prediction error (MSPE) is calculated over out-of-sample data to assess how each forecasting model performs. This measures the expected squared distance between the predicted oil price and the actual oil price:

[13] Less observations for the longer maturities

$$MSPE = \mathbb{E}\left[\left(S_{t+h}^{oil} - \hat{S}_{t+h|t}^{oil}\right)^2\right]$$

The statistical significance of the MSPE reductions relative to the random walk model is also tested, where the null hypothesis states that the forecasting model in question does not reduce the MSPE. This is done through the DM-test by Diebold & Mariano (1995). This test is described in exhibit 5.

Another way of assessing the forecasting models is to compute the directional accuracy, i.e. how often they correctly predict an increase or decrease of the oil price. An accurate model should at least have a directional accuracy above 0.5 as anticipated by the random walk model. Finally, all results and its implications are presented in chapter 8.

7. Data and descriptive statistics

Daily spot and futures prices are collected using EIKON for the following commodities:

- Crude oil: WTI (Light-Sweet, Cushing, Oklahoma)
- Heating Oil: No. 2 Heating Oil (New York Harbor)
- Gasoline: RBOB Regular Gasoline (New York Harbor)

where the futures data consist of 1, 2, 3, 4, 6 and 12 months to maturity. Inventory data, US crude oil production, rig count and S&P 500 index are also collected from EIKON, while US Treasury bond and Treasury bill yields are collected from the Federal Reserve website. Daily data is averaged to monthly data in order to remove any noise in the price fluctuation (Alquist & Kilian, 2010; Baumeister & Kilian, 2012). All computations and regressions are conducted in MATLAB.

Figure 1 plots the daily prices of WTI for spot, 1-month and 1-year maturities. The figure documents increased volatility from 2005 onwards, with visually little deviation between the different contract prices. The figure also illustrates the volatile oil price movements during the past five years. The price of WTI Crude oil tumbled down from approximately 107$ per barrel in 2014 to around 30$ throughout 2015. It then steadily climbed up to around 75$ per barrel, before dropping to approximately 45$ at the end of 2018. The cause of the oil price reduction in 2014 has been highly debated, where some points to the oversupply from U.S. shale oil extraction, while others argue that a weakening demand of oil played a larger role (Prest, 2018).

Figure 2 shows monthly trading volume on NYMEX (New York Mercantile Exchange) for WTI futures contracts with different maturities. The figure shows that the market for longer contracts is less liquid compared to the shorter contracts (note the different scales on the axes). This weakens the reliability of the futures prices with longer maturities in this analysis (whether it reflects the contemporaneous market price or not).

Exhibit 6 documents descriptive statistics for the raw data in this analysis as well as Fama and French's three factors. Both spot and futures prices of all the commodities seem to be highly volatile with large standard deviation compared to

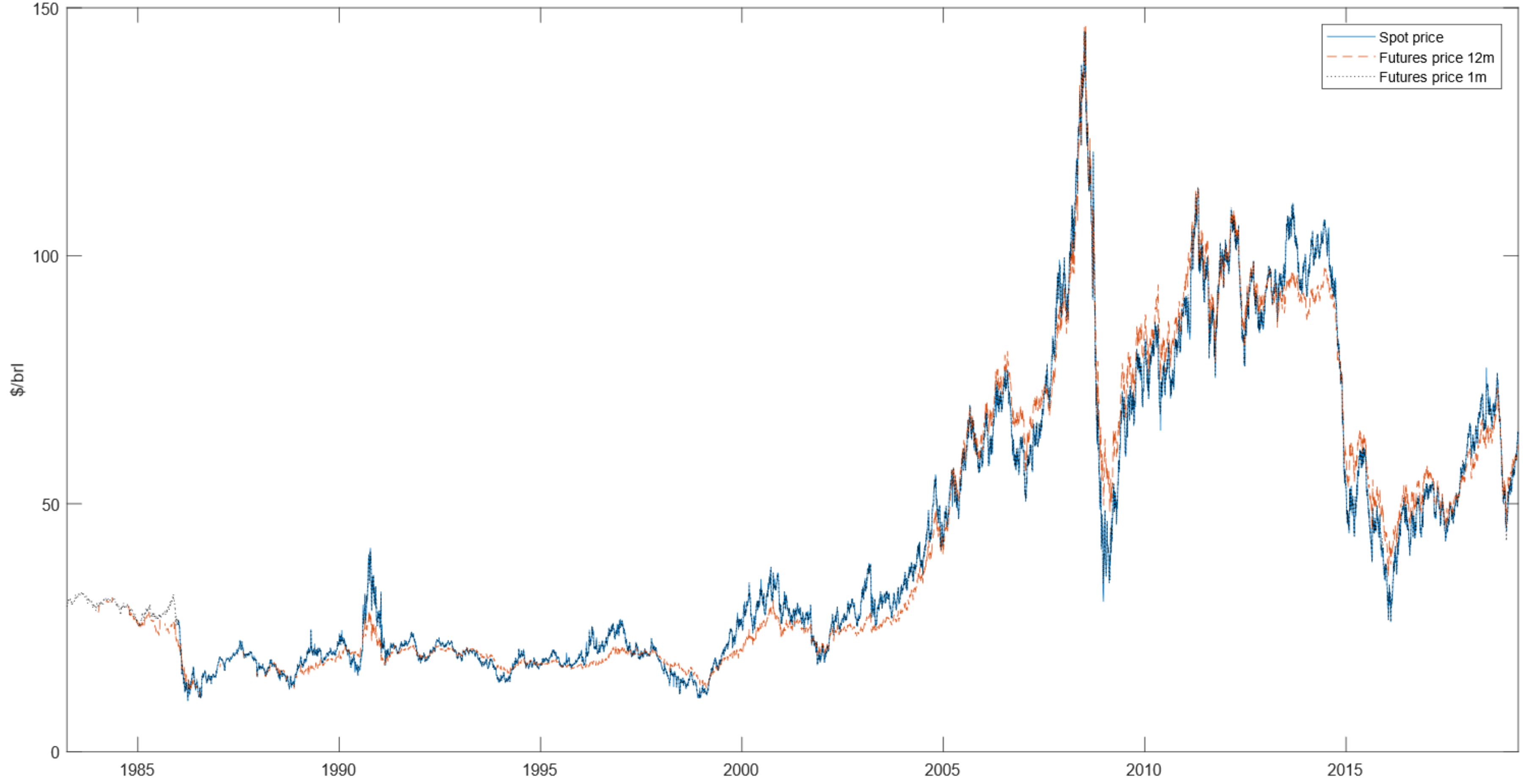


Figure 1: Historical prices of the WTI spot and futures contracts over the period March 1983 to April 2019

their mean values for all maturities. Several of the longer maturing contracts for gasoline did not start trading at the same period as the shorter kinds, as shown in the second last column, meaning the statistics are not fully comparable across variables.

Exhibit 7 shows descriptive statistics of the actual regression variables, i.e. the gross changes in the spot oil price, and the explanatory variables presented in table 1. The values for crude oil futures prices (normalized by the spot price) are on average near one for shorter maturities and decreasing, meaning a larger backwardation, for longer contracts. The standard deviation also shows that the normalized futures prices have relatively low variation compared to the actual changes in the spot price, which may lower the precision of the regression parameters. This can also be seen by plotting these variables against each other. This is illustrated in exhibit 8.

Further, both gasoline- and heating oil product spreads have been mostly positive (crude oil price below product price), which is necessary in order to induce production of these products. The futures price of crude oil also appears to be less variable around its spot price compared to gasoline and heating oil futures, and even less variable than actual spot price changes.

The variability of all the fundamental variables seem to be fairly small relative to the other variables, except for changes in rig count which is almost as varying as the changes of crude oil spot price. The variability of S&P 500 excess return seems, not unexpectedly, quite large comparing to its mean values, and increases with longer horizon. Not unexpectedly, the mean of excess return is positive as anticipated by a positive market risk premium.

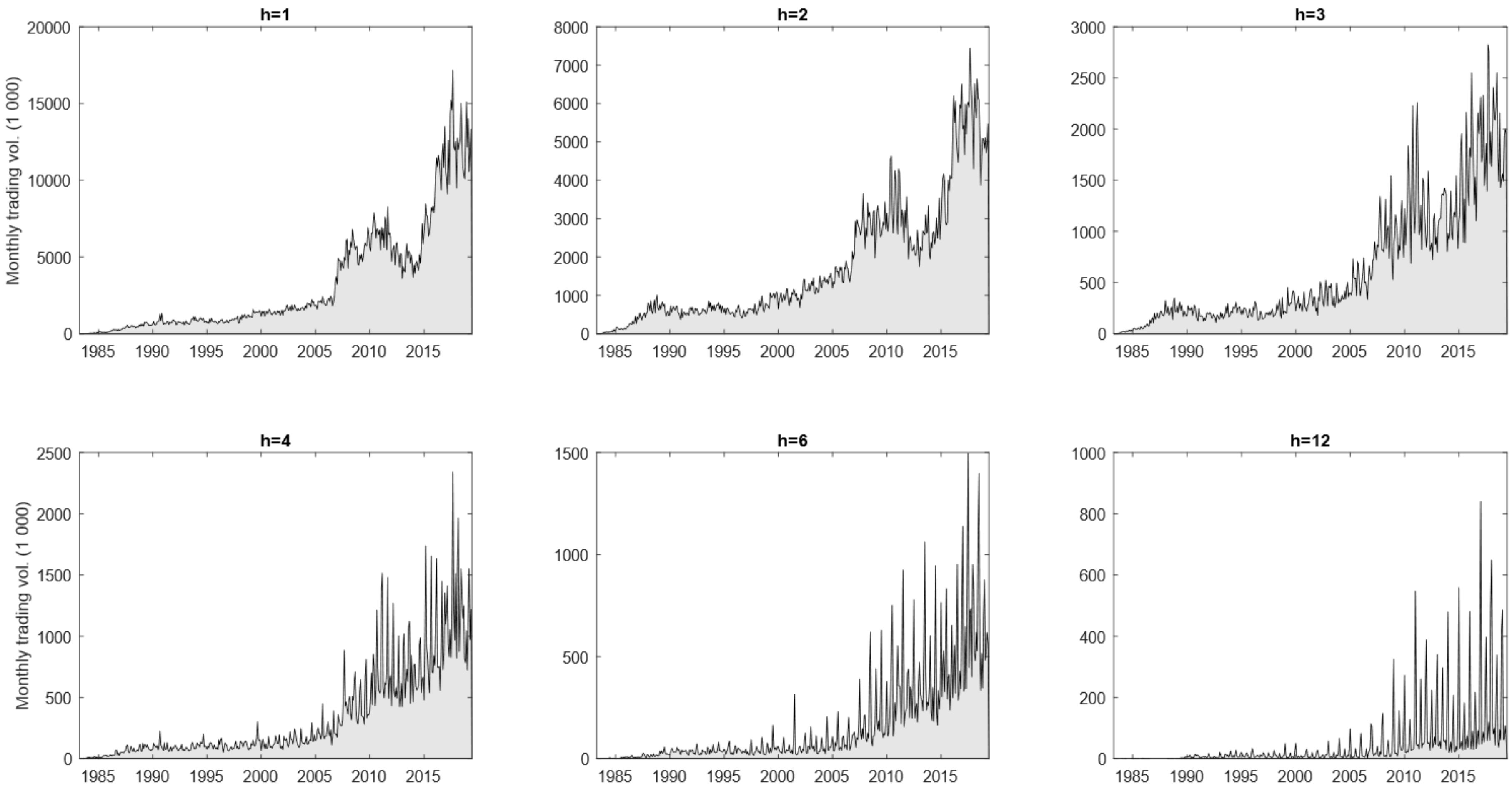


Figure 2: Monthly trading volume of WTI crude oil futures contracts (at NYMEX) for different maturities.

8. Results and discussion

This chapter presents and discusses the results related to the tests and analyses conducted.

8.1 Stationarity

Exhibit 9 documents the results from the ADF-test. The ADF-test rejects that the regression variables have a unit root, indicating that they are stationary. This is also demonstrated in the plots of the variables in exhibit 9. This means that regression models (R.1) and (R.3) can be regressed without the concern of ending up with a spurious regression. This also implies that the coefficients extracted from these regressions are appropriate to use in the forecasting models.

8.2 Correlation matrix

A correlation matrix (for each forecasting horizon) is computed in order to examine whether the suggested explanatory variables are strongly correlated. The results of the correlation matrix are presented in exhibit 10. The tables document a high correlation between the crack spread, gasoline product spread, and heating oil product spread for most horizons. The fact that the crack spread is highly correlated with the single product spreads is not surprising, as the crack spread is a weighted product of the two. Furthermore, the heating oil spread is also highly correlated with the gasoline product spread and the futures spread for longer horizons (0,85 for 12-months horizon). In order to avoid the problem of strong multicollinearity, the crack spread and heating oil product spread are excluded from the multivariate model.

8.3 Full data set

The following section presents the results of the in- and out-of-sample analyses when the analyses are conducted on the full data set (January 1986 - April 2019).

8.3.1 In-sample analysis

Table 2 documents the constant and beta coefficients estimated in the in-sample analysis. In the univariate model, the constants are found to be moderately close to

In-sample analysis		**Sample: 1986/1 - 2019/04, n = 400**					
Variable	Coefficient	$h = 1$	$h = 2$	$h = 3$	$h = 4$	$h = 6$	$h = 12$
Constant	γ	0,2525	-0,1403	-0,1402	-0,1329	0,1076	0,0613
Fut	β	0,7538	1,1553**	1,1649***	1,1676***	0,9448***	1,0475***
R^2 (adjusted)		0,0357	0,0622	0,0805	0,0718	0,0632	0,1112
p-value: $H_0: \gamma = 0$		0,86011	0,76339	0,75027	0,75568	0,7616	0,85523
p-value: $H_0: \beta = 1$		0,86363	0,73896	0,7079	0,69512	0,87601	0,88313
p-value: $H_0: \gamma = 0, \beta = 1$		0,4294	0,2706	0,1799	0,1451	0,0900	0,0632
n		399	398	397	396	394	273

Variable		$h = 1$	$h = 2$	$h = 3$	$h = 4$	$h = 6$	$h = 12$
Constant	γ_0	0,2153	-0,0975	-0,335	0,5762	-0,5755	-2,7103***
Fut	β_1	0,7523	0,8134*	0,9457*	0,3188	2,2954**	2,6273***
Gas	γ_2	0,0668	0,1174	0,2765**	0,4297**	0,1406	0,7465***
dInv	γ_3	-0,0999	-0,0431	0,0003	-0,3566	-1,0335***	-0,6241
dProd	γ_4	0,1644	0,2407	0,1136	0,1741	0,2991	0,4486
dRigs	γ_5	-0,1034	-0,035	-0,0328	-0,1993	-0,201	0,3266
SP	γ_6	-0,1298	0,025	0,0203	-0,1164	0,0889	-0,5451
dSlope	γ_7	-0,0182	-0,0105	0,0094	0,0101	-0,0362	-0,0021
dCred	γ_8	-0,1032***	-0,0706	-0,0446	-0,0321	0,0261	-0,0922
RV	γ_9	-0,0011*	-0,0008	-0,0004	0	0	-0,0001
R^2 (adjusted)		0,0564	0,0910	0,1124	0,0973	0,2421	0,4467
p-value: $H_0: \gamma_0 = 0$		0,88953	0,84125	0,61891	0,34933	0,64137	0,00294
p-value: $H_0: \beta_1 = 1$		0,87390	0,65372	0,91651	0,10489	0,16915	0,010532
p-value: $H_0: \gamma_0 = 0, \beta_1 = 1$		0,9824	0,6613	0,7342	0,2659	0,1750	0,00088
n		399	302	397	300	154	148

Table 2: Coefficients from in-sample regression using full dataset (*, **, *** denote statistical difference from zero at 10 %, 5 % and 1 % levels, respectively). All t- and Wald-tests are computed based on HAC standard errors.

zero and the beta coefficients are found to be relatively close to one for all horizons. In addition, the difference between the coefficient estimates and the values under the null hypothesis (zero and one respectively) decreases as time to maturity increases. The constant values are not significantly different from zero and the beta values are not significantly different from one at any conventional significance level. This is interpreted as suggestive evidence of no risk premium in the futures price and that the futures price being an unbiased estimator of the spot price of crude oil. Moreover, this is examined further through the joint hypothesis test.

The joint null hypothesis of no risk premium and unbiased futures prices is not rejected at the conventional 5 % significance level at any horizon. This is interpreted as an indication that there is no risk premium in the futures oil market, and that the futures price is an unbiased predictor of the spot price of crude oil. However, this only applies under the assumption that the EMH holds (Chernenko et al., 2004; Fama & French, 1987). Furthermore, one should also bear in mind that not rejecting the null hypothesis does not mean that the model under the null hypothesis is true (Neyman & Pearson, 1933).

Relaxing the significance level to 10 % entails that the null hypothesis of unbiased futures price is rejected at the 6- and 12-month horizon. Rejecting the joint null hypothesis suggests that the futures price is not unbiased and that there may be a risk premium in the oil futures market for longer horizons. This finding is also consistent with previous studies such as Alquist & Kilian (2010).

The results of running the multivariate model show that the gasoline product spread is significantly different from zero at a 5% significance level at 3- and 4-month horizon, and significantly different from zero at a 1% significance level at 12-month horizon. Furthermore, if one believes that the risk premium is time-varying, this may suggest that the gasoline product futures are positively related with the risk premium, as the coefficients are positive on all horizons. Put differently, this could indicate that the gasoline product spread provides some predictive information (additionally to the crude oil futures) through the convenience yield when marginal production cost is changing, as explained in chapter 5.2.5. Nevertheless, these results could also be due to pure randomness and should be interpreted carefully.

8.3.2 Comparison with CAPM and Fama-French three-factor model

The expected futures return estimated using the CAPM and Fama-French three-factor model is presented in table 3. The table shows that the beta-coefficient between the expected futures return (i.e., futures risk premium) and the market risk premium is close to zero for both tests. Further, the table also documents that these beta coefficients are not significantly different from zero. This indicates that there should be no risk premium in the future price of crude oil according to the asset pricing models. This also suggests that the return from holding a long position in the crude oil futures market does not vary with the economy.

Together with the results from the in-sample analysis indicating that the futures price is unbiased, the results from the asset pricing models further strengthens the indication of no risk premium and rational futures prices.

CAPM	**Estimate**	**SE**	**t-stat**	**p-value**
Intercept	0,00606	0,004384	1,3834	0,1673
MKT	0,03989	0,098787	0,4038	0,6866
Number of observations: 399				
F-statistic vs. constant model: 0,163, p-value = 0,687				

Three-factor	**Estimate**	**SE**	**t-stat**	**p-value**
Intercept	0,00560	0,00440	1,2728	0,2038
MKT	0,04175	0,10250	0,4073	0,6840
SMB	0,18681	0,14659	1,2744	0,2033
HML	0,21812	0,15782	1,3821	0,1677
Number of observations: 399				
F-statistic vs. constant model: 1,02, p-value = 0,386				

Table 3: In-sample analysis of the risk premium against the CAPM and Fama & French three-factor model.

8.3.3 Out-of-sample analysis

Table 4a shows that the simple futures model (following the rationale of the unbiased expectations hypothesis) outperforms the other models in terms of MSPE at all horizons above 1 month. The MSPE improvements of the simple futures model relative to the random walk model is significantly different from zero at a 10 % significance level at the 1-, 2-, 3- and 12-months horizon. The fact that the simple futures model has the least MSPE is in line with the results from the in-sample analysis, as there was little indication of a risk premium reflected in the futures. Further, the univariate- and multivariate forecasting models are outperformed by the random walk model at almost all horizons in terms of MSPE.

Table 4b reports the directional accuracy of the alternative forecasting models. The simple futures model and the multivariate forecasting model have solid directional accuracy, while the univariate forecasting model has a directional accuracy fluctuating around 0,5. The fact that the directional accuracy of the simple futures model is systematically above 0,5 supports the findings in the in- and out-of-sample analyses indicating no risk premium.

Forecasting model	**MSPE (sample: 1986/1 - 2010/12, out-of-sample: 2011/1 - 2019/4)**					
Model	$h = 1$	$h = 2$	$h = 3$	$h = 4$	$h = 6$	$h = 12$
Random walk	31,742	83,33	132,16	167,78	236,16	438,3
Simple futures	31,206*	**78,83***	**124,12***	**157,16**	**217,35**	**350,48***
Univariate	32,962	87,042	146,67	201,7	311,5	652,37
Multivariate	**30,944**	86,488	143,71	206,23	290,65	487,18

Table 4a: MSPE of spot price forecasts for different models and maturities. Bold values indicate the model with lowest MSPE for each maturity.
Asterisks indicate significant improvement from the no-change (*, **, *** denote statistical significance at 10 %, 5 % and 1 % levels, respectively).

Forecasting model	**Directional accuracy (sample: 1986/1 - 2010/12, out-of-sample: 2011/1 - 2019/4)**					
Model	$h = 1$	$h = 2$	$h = 3$	$h = 4$	$h = 6$	$h = 12$
Simple futures	**0,6053**	0,5263	**0,5526**	0,5263	**0,6184**	0,6711
Univariate	0,5263	0,5132	0,4737	0,4868	0,4868	0,5
Multivariate	0,4886	**0,5682**	0,5455	**0,5682**	0,6136	**0,7273**

Table 4b: Directional accuracy for different models and maturities. Values represent the fraction of correctly forecasted directions of price changes. Bold values indicate the model with highest value for each maturity.

In summary, the results from the in- and out-of-sample analyses indicate that the futures price is an unbiased predictor of the future spot price of crude oil, and thus the futures price does not include a risk premium. However, when relaxing the significance level to a 10% level, there is some suggestive evidence of a risk premium in the crude oil futures price in the in-sample analysis.

8.4 Restricted data set

The following section presents the results of the in- and out-of-sample analyses when the analyses is conducted on the restricted data set (January 2006 - April 2019).

8.4.1 In-sample analysis

The result of the in-sample analysis conducted on the restricted data set is presented in table 5. For the univariate model, the constant is estimated to be systematically negative and significantly different from zero at a 5% significance level at the horizons: 3-, 4-, 6- and 12-months. This is interpreted as an indication of a constant risk premium in the futures price. Furthermore, the beta coefficient is consistently above one and significantly different from one at a 1% significance level at the horizons: 2-, 3-, 4-, 6- and 12-months. This is interpreted as an indication of a time-varying risk premium in the futures price.

In addition, the joint null hypothesis of unbiased futures prices is rejected at a 5% significance level for 3-, 4- and 12-months horizons. The rejection of the null hypothesis indicates that the futures price is a biased estimator of the future spot price of crude oil and the existence of a risk premium in the futures market (at least for medium to long horizons).

The results for the multivariate model show that two fundamental variables (inventory and production) are significantly different from zero at a 10% significance level (or lower) at medium horizons (3-6 months). This indicates that inventory and production may be related to the risk premium at these horizons. Moreover, the negative coefficients for change in inventory is supported by theory as the convenience yield is expected to fall when stock-out is less likely. However, this effect was rather expected for the short horizons. The coefficients for changes in production are positive for all horizons. This indicates that an increase in production predicts an increase in the risk premium, thus increasing the spot price of oil. This contradicts with the theory stating that a decrease in risk premium, through decreased convenience yield, is expected when production increases. However, as explained in chapter 5.1.2, the effect of production can be

In-sample analysis		**Sample: 2006/1 - 2019/04, n = 145**					
Variable	Coefficient	$h = 1$	$h = 2$	$h = 3$	$h = 4$	$h = 6$	$h = 12$
Constant	γ	-3,1485	-1,1359*	-1,1011**	-1,1734**	-1,0512**	-1,414***
Fut	β	4,1471**	2,1222***	2,0791***	2,1432***	2,0159***	2,3619***
R^2 (adjusted)		0,0187	0,0832	0,1289	0,1694	0,1970	0,3584
p-value: $H_0: \gamma = 0$		0,1305	0,0739	0,0439	0,0290	0,0333	0,0006
p-value: $H_0: \beta = 1$		0,1303	0,0712	0,0389	0,0244	0,0297	0,0009
p-value: $H_0: \gamma = 0, \beta = 1$		0,3132	0,1442	0,0259	0,0137	0,0747	0,0016
n		159	158	157	156	154	148

Variable		$h = 1$	$h = 2$	$h = 3$	$h = 4$	$h = 6$	$h = 12$
Constant	γ_0	-2,496	-1,4355*	-1,75**	-1,9038**	-0,5755	-2,7103***
Fut	β_1	3,5551*	2,2433***	2,3567***	2,6934***	2,2954**	2,6273***
Gas	γ_2	0,051	0,1296	0,2057	0,2101	0,1406	0,7465***
dInv	γ_3	-0,1826	-0,3626	-0,6209*	-0,8292**	-1,0335***	-0,6241
dProd	γ_4	0,2289	0,5405	1,0006**	1,0473**	0,2991	0,4486
dRigs	γ_5	-0,1612	-0,1443	-0,2408	-0,2957	-0,201	0,3266
SP	γ_6	0,1035	-0,0546	-0,489	-0,0981	0,0889	-0,5451
dSlope	γ_7	0,0081	0,0237	0,0538	0,0405	-0,0362	-0,0021
dCred	γ_8	-0,0796	-0,0359	-0,0223	0,0382	0,0261	-0,0922
RV	γ_9	-0,0011*	-0,0011*	-0,0011	-0,0007	0	-0,0001
R^2 (adjusted)		0,1085	0,1752	0,2624	0,2663	0,2421	0,4467
p-value: $H_0: \gamma_0 = 0$		0,2627	0,0529	0,0309	0,0338	0,64137	0,00294
p-value: $H_0: \beta_1 = 1$		0,2288	0,0449	0,0351	0,0215	0,16915	0,010532
p-value: $H_0: \gamma_0 = 0, \beta_1 = 1$		0,4602	0,1039	0,0477	0,0496	0,1750	0,00088
n		159	158	157	156	154	148

Table 5: Coefficients from in-sample regression using full dataset (*, **, *** denote statistical difference from zero at 10 %, 5 % and 1 % levels, respectively). All t- and Wald-tests are computed based on HAC standard errors.

ambiguous as increased production may simply be interpreted as a result of increased exogenous demand. Furthermore, gasoline product spread is not significant for medium horizons, as was found in the analysis of the full data set.

8.4.2 Comparison with CAPM and Fama-French three-factor model

The expected futures return estimated using CAPM and Fama-French three-factor model is presented in table 6. The two asset pricing models estimate a market beta coefficient of 0,5 and 0,59 respectively. Both values are significantly different form zero at a 1% significance level. This suggests that the return in holding a long crude oil futures position is (to a certain degree) positively related with the economy. Since the in-sample analysis indicates that futures prices are biased, these asset pricing models supports the indication of a risk premium and rational futures prices.

CAPM	**Estimate**	**SE**	**t-stat**	**p-value**
Intercept	-0,00167	0,0068	-0,2478	0,80456
MKT	0,59066	0,1581	3,7360	0,00026

Number of observations: 159
F-statistic vs. constant model: 14, p-value = 0,000261

Three-factor	**Estimate**	**SE**	**t-stat**	**p-value**
Intercept	-0,00163	0,00677	-0,2413	0,80968
MKT	0,50275	0,17265	2,9120	0,00412
SMB	0,53715	0,30519	1,7600	0,08037
HML	-0,11048	0,26407	-0,4184	0,67625

Number of observations: 159
F-statistic vs. constant model: 5,76, p-value = 0,000926

Table 6: Regression results between crude oil risk premium and market excess return, and between the risk premium and Fama-French three factors.

8.4.3 Out-of-sample analysis

Table 7a documents that the univariate forecasting model outperforms the other models in terms of MSPE at all horizons. However, the MSPE improvement of the univariate model relative to the random walk model is only significantly different from zero at a 10% significance level at the 2-month horizon. Moreover, the multivariate model performs the worst at all horizons. This result suggests either that the risk premium is not time-varying or that the multivariate model does not include the relevant predictive variables for the risk premium.

Table 7b reports the directional accuracy the forecasting models. The simple futures- and multivariate model possesses solid directional accuracy which is systematically above 0,5. The univariate model has a directional accuracy fluctuating around 0,5. This finding is somewhat surprising as the univariate model has the lowest MSPE.

Forecasting model	**MSPE (sample: 2006/1 – 2010/12, out-of-sample: 2011/1 – 2019/4)**					
Model	$h = 1$	$h = 2$	$h = 3$	$h = 4$	$h = 6$	$h = 12$
Random walk	28,666	74,94	119,84	153,93	221,51	413,32
Simple futures	28,179*	70,876*	112,59*	144,26	204,59	331,71*
Univariate	**27,903**	**69,128***	**109,11**	**138,97**	**198,64**	**284,65**
Multivariate	35,68	90,612	143,32	185,12	290,65	487,18

Table 7a: Bold values indicate the model with lowest MSPE for each maturity. Asterisks indicate significant improvement from the no-change (*, **, *** denote statistical significance at 10 %, 5 % and 1 % levels, respectively).

Forecasting model	**Directional accuracy (sample: 2006/1 – 2010/12, out-of-sample: 2011/1 – 2019/4)**					
Model	$h = 1$	$h = 2$	$h = 3$	$h = 4$	$h = 6$	$h = 12$
Simple futures	**0,6364**	0,5568	0,5568	0,5455	**0,6136**	0,7045
Univariate	0,5568	0,5455	0,4886	0,4546	0,5341	0,6818
Multivariate	0,5000	**0,5682**	**0,6023**	**0,6477**	**0,6136**	**0,7273**

Table 7b: Values represent the fraction of correctly forecasted directions of price changes. Bold values indicate the model with highest value for each maturity.

In summary, the in-sample analysis indicates that the futures price is a biased estimator of the spot prices and documents some indications of a constant- and a time-varying risk premium. Furthermore, the out-of-sample analysis demonstrates that the univariate forecasting model marginally outperforms the simple futures model and the random walk model, while the multivariate forecasting model clearly performs worse in terms of MSPE. This finding is interpreted as indications of a risk premium in the futures price. However, whether the risk premium is time-varying is not clear as we have not been able to fully identify relevant predictive factors for the risk premium.

9. Conclusion

This thesis investigates the unbiasedness of the crude oil futures price. In an in-sample setting, this involves examining the existence and potential form of a risk premium in the futures price. In an out-of-sample setting, this involves examining the forecasting accuracy of a set of forecasting models with different assumptions on the risk premium.

The results of the in-sample analysis, based on a time series of monthly oil prices between 1986 and 2019, documents no clear evidence of either a constant or time-varying risk premium. This implies that the unbiased expectations hypothesis cannot be rejected. Furthermore, the analysis documents no significant relationship between futures return (i.e., the realized risk premium) and the market portfolio calculated using CAPM and Fama-French three-factor model. This result supports the findings in the in-sample analysis and indicates that an investor obtains no reward for holding a long futures position (i.e., for risk-taking). Finally, the out-of-sample analysis supports this result as the simple futures model (assuming unbiased futures price, and thus no risk premium) outperforms the alternative forecasting models.

Performing the same analysis on a restricted data set between 2006 and 2019 results in a rejection of the unbiased expectations hypothesis. Rejecting the unbiased expectations hypothesis indicates that the futures price is a biased estimator of the future spot price of crude oil, and thus that there exists a risk premium in the futures price. This finding is supported in the out-of-sample analysis documenting that the univariate forecasting model outperforms the other models in terms of MSPE. However, whether the risk premium is constant and/or time-varying is not clear as we have not been able to fully identify relevant predictive factors for the risk premium.

Finally, the fact that the analyses provides contradicting results depending on the time period, shows that these models may not be robust across time. Alternatively, this could suggest that oil futures prices only reflected a risk premium in recent history, but such interpretations should be made carefully when no further research has been conducted.

10. Limitations and further research

The forecasting performance of the multivariate forecasting model builds on the assumption that all relevant variables which affect the risk premium in the crude oil futures market are included. However, the fact that the multivariate forecasting models are outperformed when the in-sample analysis indicates a time-varying risk premium (in the restricted set) could indicate that not all variables related to the risk premium are included. Examples of such variables could be the net exposure of the hedgers, political factors, VIX and stock dividend yield, which could be an interesting extension for further research. In addition, testing for structural changes in the market and isolating this effect could also be a possibility to further locate where and whether the characteristics of the risk premium have changed throughout the history.

Another relevant extension of this thesis could be assessing the performance of the forecasting models relative to forecasts made by U.S. Energy Information Administration (EIA) in addition to the no-change forecast. EIA produces monthly and quarterly oil price forecasts, which are widely used to guide natural resource development and investments in infrastructure (Bank of England, 2004, p. 31; Federal Reserve, 2018). Therefore, by assessing the forecasting abilities of the forecasting models in question relative to the EIA forecasts, one could obtain a deeper view of the performance of the forecasting models relative to another forecasting model used in practice.

Finally, it could be argued that the tested variables in this thesis affect the price and risk premium much more frequently than on a monthly basis. In the current age of high frequency data, it is expected that all new information is immediately reflected in the prices. Thus, whether arbitrage opportunities (or anomalously high excess returns) have existed for shorter periods would be an interesting question to address for future research.

References

Abosedra, S. (2005). Futures versus univariate forecast of crude oil prices. *OPEC Review*, *29*(4), 231-241. doi:10.1111/j.1468-0076.2005.00153.x

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

Bank of England. (2004). *Inflation Report*.

Baumeister, C., & Kilian, L. (2012). Real-Time Forecasts of the Real Price of Oil. *Journal of Business & Economic Statistics*, *30*(2), 326-336. doi:10.1080/07350015.2011.648859

Baumeister, C., & Kilian, L. (2015). Forecasting the Real Price of Oil in a Changing World: A Forecast Combination Approach. *Journal of Business & Economic Statistics*, *33*(3), 338. doi:10.1080/07350015.2014.949342

Baumeister, C., Kilian, L., & Lee, T. K. (2014). Are there gains from pooling real-time oil price forecasts? *Energy Economics*, *46*(S1), S33-S43. doi:10.1016/j.eneco.2014.08.008

Baumeister, C., Kilian, L., & Zhou, X. (2013a). Are Product Spreads Useful for Forecasting? An Empirical Evaluation of the Verleger Hypothesis. *IDEAS Working Paper Series from RePEc*.

Baumeister, C., Kilian, L., & Zhou, X. (2013b). Are product spreads useful for forecasting? An empirical evaluation of the Verleger hypothesis.

Bodie, Z. (2014). *Investments* (10th global ed. ed.). Berkshire: McGraw-Hill Education.

Bollerslev, T., Tauchen, G., & Zhou, H. (2009). Expected Stock Returns and Variance Risk Premia. *The Review of Financial Studies*, *22*(11), 4463-4492. doi:10.1093/rfs/hhp008

Breeden, D. T. (1980). Consumption riks in futures markets. In. *The Journal of Finance* , *35* (2), 503-520.

Brennan, M. (1958). The Supply of Storage. *The American Economic Review*, *48*(1), 50.

Brooks, C. (2014). *Introductory econometrics for finance* (3rd ed. ed.). Cambridge: Cambridge University Press.

Brooks, C., Prokopczuk, M., & Wu, Y. (2013). Commodity futures prices: More evidence on forecast power, risk premia and the theory of storage. *Quarterly Review of Economics and Finance*, *53*(1), 73. doi:10.1016/j.qref.2013.01.003

Chernenko, S., Schwarz, K., & Wright, J. H. (2004). The information content of forward and futures prices: Market expectations and the price of risk.

Chevillon, G., & Rifflart, C. (2009). Physical market determinants of the price of crude oil and the market premium. *Energy Economics, 31*(4), 537-549. doi:10.1016/j.eneco.2009.01.002

CME Group. (2017). Introduction to Crack Spreads. Retrieved from https://www.cmegroup.com/education/articles-and-reports/introduction-to-crack-spreads.html

CME Group. (2018). Futures v Forwards: Cost Implications. Retrieved from https://www.cmegroup.com/education/articles-and-reports/futures-v-forwards-cost-implications.html

Cochrane, J. H. (2008). The Dog That Did Not Bark: A Defense of Return Predictability. *The Review of Financial Studies*, *21*(4), 1533-1575. doi:10.1093/rfs/hhm046

Collin - Dufresn, P., Goldstein, R. S., & Martin, J. S. (2001). The Determinants of Credit Spread Changes. *Journal of Finance, 56*(6), 2177-2207. doi:10.1111/0022-1082.00402

Cootner, P. H. (1960). Returns to Speculators: Telser versus Keynes. *Journal of Political Economy, 68*(4), 396-404. doi:10.1086/258347

Coppola, A. (2008). Forecasting oil price movements: Exploiting the information in the futures market. *Journal of Futures Markets: Futures, Options, and Other Derivative Products, 28*(1), 34-56.

Cortazar, G., Kovacevic, I., & Schwartz, E. S. (2015). Expected commodity returns and pricing models. *Energy Economics, 49*(C), 60-71. doi:10.1016/j.eneco.2015.01.015

Deaves, R., & Krinsky, I. (1992a). Risk Premiums and Efficiency in the Market for Crude Oil Futures. *The Energy Journal, 13*(2), 93. doi:10.5547/ISSN0195-6574-EJ-Vol13-No2-5

Deaves, R., & Krinsky, I. (1992b). The behavior of oil futures returns around OPEC conferences. *Journal of Futures Markets, 12*(5), 563-574. doi:10.1002/fut.3990120507

Dickey, D. A., & Fuller, W. A. (1979). Distribution of the Estimators for Autoregressive Time Series With a Unit Root. *Journal of the American Statistical Association, 74*(366), 427-431. doi:10.2307/2286348

Diebold, F. X., & Mariano, R. S. (1995). Comparing Predictive Accuracy. *Journal of Business & Economic Statistics, 13*(3), 253-263. doi:10.1080/07350015.1995.10524599

Dusak, K. (1973). Futures Trading and Investor Returns: An Investigation of Commodity Market Risk Premiums. *Journal of Political Economy, 81*(6), 1387-1406. doi:10.1086/260133

Edwards, F. R. (1992). *Futures and options*. New York: McGraw-Hill.

Fama, E. F. (1970). Efficient Capital Markets: A Review of Theory and Empirical Work. *The Journal of Finance, 25*(2), 383-417. doi:10.2307/2325486

Fama, E. F., & French, K. R. (1987). Commodity Futures Prices: Some Evidence on Forecast Power, Premiums, and the Theory of Storage. *The Journal of Business, 60*(1), 55-73.

Fama, E. F., & French, K. R. (1988). Business Cycles and the Behavior of Metals Prices. *Journal of Finance, 43*(5), 1075-1093. doi:10.1111/j.1540-6261.1988.tb03957.x

Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics, 33*(1), 3-56. doi:10.1016/0304-405X(93)90023-5

Federal Reserve. (2018). *Unraveling the Oil Conundrum: Productivity Improvements and Cost Declines in the U.S. Shale Oil Industry*.

Fisher, I. (1896). What is Capital? *The Economic Journal, 6*(24), 509-534. doi:10.2307/2957184

Frankel, J. A., & Froot, K. A. (1987). Short-term and long-term expectations of the yen/dollar exchange rate: Evidence from survey data. *Journal of The Japanese and International Economies, 1*(3), 249-274. doi:10.1016/0889-1583(87)90011-6

French, K. R., Schwert, G. W., & Stambaugh, R. F. (1987). Expected stock returns and volatility. *Journal of Financial Economics, 19*(1), 3-29. doi:10.1016/0304-405X(87)90026-2

Gjolberg, O., & Johnsen, T. (1999). Risk management in the oil industry: can information on long-run equilibrium prices be utilized? *Energy Economics, 21*(6), 517-527. doi:10.1016/S0140-9883(99)00023-7

Granger, C. W. J. (1969). Investigating Causal Relations by Econometric Models and Cross-spectral Methods. *Econometrica, 37*(3), 424-438. doi:10.2307/1912791

Hankyeung, C., David, J. L., & Kunlapath, S. (2015). Oil Price Forecasting Using Crack Spread Futures and Oil Exchange Traded Funds. *Contemporary Economics, 9*(1), 1-108. doi:10.5709/ce.1897-9254.158

Heinkel, R., Howe, M. E., & Hughes, J. S. (1990). Commodity convenience yields as an option profit. *Journal of Futures Markets, 10*(5), 519-533. doi:10.1002/fut.3990100508

Hicks, J. R. (1939). The Foundations of Welfare Economics. *The Economic Journal, 49*(196), 696-712. doi:10.2307/2225023

Hull, J. C. (2017). *Options, Futures, and Other Derivatives, Global Edition* (9 ed.): United Kingdom: Pearson Education M.U.A.

Jensen, M. C. (1978). Some anomalous evidence regarding market efficiency. *Journal of Financial Economics, 6*(2), 95-101. doi:10.1016/0304-405X(78)90025-9

Kaldor, N. (1939). Speculation and Economic Stability. *The Review of Economic Studies, 7*(1), 1-27. doi:10.2307/2967593

Kaufmann, R. K., Dees, S., Gasteuil, A., & Mann, M. (2008). Oil prices: The role of refinery utilization, futures markets and non-linearities. *Energy Economics, 30*(5), 2609-2622. doi:10.1016/j.eneco.2008.04.010

Keynes, J. M. (1930). Treatise on money.

Kocagil, A. E. (2004). Optionality and Daily Dynamics of Convenience Yield Behavior: An Empirical Analysis. *Journal of Financial Research, 27*(1), 143-158. doi:10.1111/j.1475-6803.2004.00082.x

Lintner, J. (1965). The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets. *The Review of Economics and Statistics, 47*(1), 13-37. doi:10.2307/1924119

Lutz, F. A. (1940). The Structure of Interest Rates. *The Quarterly Journal of Economics, 55*(1), 36-63. doi:10.2307/1881665

Merino, A., & Ortiz, A. (2005). Explaining the so-called "price premium" in oil markets. *OPEC Review, 29*(2), 133. doi:10.1111/j.0277-0180.2005.00148.x

Mirmirani, S., & Cheng Li, H. (2004). A comparison of VAR and neural networks with genetic algorithm in forecasting price of oil. In *Applications of Artificial Intelligence in Finance and Economics* (pp. 203-223): Emerald Group Publishing Limited.

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. doi:10.1016/0140-9883(94)90003-5

Murat, A., & Tokat, E. (2009). Forecasting oil price movements with crack spread futures. *Energy Economics, 31*(1), 85-90. doi:10.1016/j.eneco.2008.07.008

Möbert, J. (2007). Crude oil price determinants. In (Vol. 186): Darmstadt University of Technology, Department of Law and Economics.

Nai-Fu, C., Roll, R., & Ross, S. (1986). Economic Forces and the Stock Market. *The Journal of Business (1986-1998), 59*(3), 383.

Newey, W. K., & West, K. D. (1987). A Simple, Positive Semi- Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. *Econometrica, 55*(3), 703-708. doi:10.2307/1913610

Neyman, J., & Pearson, E. S. (1933). *The testing of statistical hypotheses in relation to probabilities a priori.* Paper presented at the Mathematical Proceedings of the Cambridge Philosophical Society.

Paulson, A. (1999). Risk arbitrage opportunities in petroleum futures spreads. *The Journal of Futures Markets*, *19*(8), 931. doi:10.1002/(SICI)1096-9934(199912)19:83.0.CO;2-L

Prest, B. C. (2018). Explanations for the 2014 oil price decline: Supply or demand? *Energy Economics, 74*, 63-75. doi:10.1016/j.eneco.2018.05.029

Ringlund, G. B., Rosendahl, K. E., & Skjerpen, T. (2008). Does oilrig activity react to oil price changes? An empirical investigation. *Energy Economics, 30*(2), 371-396. doi:10.1016/j.eneco.2007.06.002

Schwarz, G. (1978). Estimating the Dimension of a Model. *The Annals of Statistics*, *6*(2), 461-464. doi:10.1214/aos/1176344136

Sharpe, W. F. (1964). CAPITAL ASSET PRICES: A THEORY OF MARKET EQUILIBRIUM UNDER CONDITIONS OF RISK*. *Journal of Finance, 19*(3), 425-442. doi:10.1111/j.1540-6261.1964.tb02865.x

Suss, S. (2009). Do implied volatilities predict stock returns. *Journal of Asset Management*, *10*(4), 222-234. doi:10.1057/jam.2009.14

Tang, L., & Hammoudeh, S. (2002). An empirical exploration of the world oil price under the target zone model. *Energy Economics*, *24*(6), 577-596. doi:10.1016/S0140-9883(02)00057-9

Telser, L. G. (1958). Futures Trading and the Storage of Cotton and Wheat. *Journal of Political Economy*, *66*(3), 233-255. doi:10.1086/258036

Timmermann, A., & Granger, C. W. J. (2004). Efficient market hypothesis and forecasting. *International Journal of Forecasting*, *20*(1), 15-27. doi:10.1016/S0169-2070(03)00012-8

Verleger, P. (2011). The Margin, Currency, and the Price of Oil. *Business Economics*, *46*(2), 71-82. doi:10.1057/be.2011.3

Verleger, P. K. (1982). The Determinants of Official OPEC Crude Prices. *The Review of Economics and Statistics*, *64*(2), 177-183. doi:10.2307/1924296

Welch, I., & Goyal, A. (2008). A Comprehensive Look at The Empirical Performance of Equity Premium Prediction. *The Review of Financial Studies*, *21*(4), 1455-1508. doi:10.1093/rfs/hhm014

Westgaard, S., Osmundsen, P., Stenslet, D., & Ringheim, J. (2017). Modeling superior predictors for crude oil prices.

Wooldridge, J. M. (2010). *Econometric Analysis of Cross Section and Panel Data*. Cambridge: Cambridge: MIT Press.

Wooldridge, J. M. (2016). *Introductory econometrics : a modern approach* (6th ed. ed.). Australia: Cengage Learning.

Working, H. (1948). Theory of the Inverse Carrying Charge in Futures Markets. *Journal of Farm Economics*, *30*(1), 1-28.

Yang, C. W., Hwang, M. J., & Huang, B. N. (2002). An analysis of factors affecting price volatility of the US oil market. *Energy Economics, 24*(2), 107-119. doi:10.1016/S0140-9883(01)00092-5

Ye, M., Zyren, J., & Shore, J. (2002). Forecasting crude oil spot price using OECD petroleum inventory levels. *International Advances in Economic Research*, *8*(4), 324-333.

Ye, M., Zyren, J., & Shore, J. (2005). A monthly crude oil spot price forecasting model using relative inventories. *International Journal of Forecasting, 21*(3), 491-501. doi:10.1016/j.ijforecast.2005.01.001

Zeng, T., & Swanson, N. R. (1998). Predictive evaluation of econometric forecasting models in commodity futures markets. *Studies in Nonlinear Dynamics & Econometrics*, *2*(4).

Exhibits

Exhibit 1: Signalling effect of oil inventory

Ye et al. (2002) argues that changes in global crude oil inventory levels provide a measure of whether production is in excess of demand and that changes in inventories thus can be used as an indicator of market pressure on price changes:

$$\Delta Inventories_t = Production_t - Demand_t$$

Assuming an initial equilibrium level of inventories, a change in inventory levels will thus contain information regarding the imbalance between market supply and demand. An increase in crude oil inventories signals that production in the last period exceeded the demand, and vice versa (Ye et al., 2002). Several studies has constructed forecasting models based the suggested signal effect proposed by Ye et al., (2002) and found that inventory does have predictive power on the spot price of crude oil in the short-run (Merino & Ortiz, 2005; Ye et al., 2002, 2005).

Exhibit 2: Derivation of the futures product spread model

A common view amongst oil analysts and market practitioners is that decreasing spreads between the price of refined oil products and the price of crude oil signals a price reduction on the future spot price of crude oil, and vice versa. This view on the dynamic relationship between the price of refined oil products and the future spot price of crude oil originates from the proposition by Verleger (1982), stating that the price of crude oil largely depends on the demand from oil refiners.

Verleger (2011; 1982) considers oil refiners to be price-takers in the product market. Oil refiners will produce refined oil products (such as gasoline, heating oil, etc.) as long as the price of crude oil allows for profitable production. If the price of crude oil does not allow for profitable production, the oil refiners will cut their production (i.e., if the crude oil price exceeds the price of refined oil products less transportation- and production costs). Consequently, a cut in production as a result of high crude oil price will in time reduce the demand for crude oil and lead to a drop in the price of crude oil. Subsequently, the short supply of refined products will lead to an increased price of refined products (P. Verleger, 2011). Based on this dynamic relationship the futures product spread is argued to have predictive power on the price of crude oil (Baumeister et al., 2013b).

Verleger (1982) argues that the value of one barrel of crude oil V_t is determined by the weighted average of the price of refined products $S_{j,t}^{refined\ prod.}$,

$$V_t = \sum_{j=1}^{n} w_{j,t} S_{j,t}^{refined\ prod.} \qquad \text{(E.1)}$$

where the weight $w_{j,t}$ defines the percentage weight of product j produced at time t using one barrel of crude oil. Note that the value of a barrel of crude oil, V_t, overstates the price a refiner is actually willing to pay for a barrel of crude oil. This is due to the fact that equation (E.1) does not consider transportation costs and cost of refining which an oil refiner incurs. By introducing transportation costs s_t and cost of refining c_t, one can express the price of crude oil refiners are willing to pay S_t^{oil} as follows (P. K. Verleger, 1982);

$$S_t^{oil} = V_t - s_t - c_t \qquad \text{(E.2)}$$

The costs will be treated as a constant, as this is the norm in empirical studies (Baumeister et al., 2013b). This means that by combining equation (E.1) and (E.2), one can express the price oil refiners are willing to pay for a barrel of crude oil as the weighted average price of refined oil products. In addition, the equation will be extended to include $t + h$ periods. This results in equation (E.3):

$$S_{t+h}^{oil} = \sum_{j=1}^{n} w_{j,t} S_{j,t+h}^{refined\ prod.} \tag{E.3}$$

which will be the starting point for deriving a product spread regression model on the change in future spot price of crude oil. Further, introducing the conditional expectation to equation (E.3) is the first step in order to obtain a forecasting model based the product spread:

$$\mathbb{E}_t[S_{t+h}^{oil}] = \sum_{j=1}^{n} w_{j,t} \mathbb{E}_t[S_{j,t+h}^{refined\ prod.}] \tag{E.4}$$

According to the expectations hypothesis the futures price of refined product *j* equals the expected spot price of refined product *j*. This hypothesis therefore assumes no risk premium. Introducing the expectations hypothesis to equation (E.4) results in the following equation:

$$\mathbb{E}_t[S_{t+h}^{oil}] = \sum_{j=1}^{n} w_{j,t} F_{j,t+h|t}^{refined\ prod.} \tag{E.5}$$

In order to derive the product spread both sides are divided by S_t^{oil}, resulting in the following equation:

$$\frac{\mathbb{E}_t[S_{t+h}^{oil}]}{S_t^{oil}} = \sum_{j=1}^{n} w_{j,t} \frac{F_{j,t+h|t}^{refined\ prod.}}{S_t^{oil}} \tag{E.6}$$

From the equation above, the regression model (E.7) naturally emerges:

$$\frac{S_{t+h}^{oil}}{S_t^{oil}} = \alpha + \beta \left[\sum_{j=1}^{n} w_{j,t} \frac{F_{j,t+h|t}^{refined\ prod.}}{S_t^{oil}} \right] + \varepsilon_{t+h} \tag{E.7}$$

$\hat{\alpha}$ and $\hat{\beta}$ are obtained from the regression model (E.7) (using either rolling window or recursive least squares method). By reformulating and setting the residual to zero, we get the following forecasting model:

$$\hat{S}^{oil}_{t+h|t} = S^{oil}_t \left\{ \hat{\alpha} + \hat{\beta} \left[\sum_{j=1}^{n} w_{j,t} \frac{F^{refined\ prod.}_{j,t+h|t}}{S^{oil}_t} \right] \right\} \tag{E.8}$$

From the forecasting model above, one can conduct forecasting based on a single future spread model, a crack spread futures model and a weighted product futures model. This model can also be extended to include time-varying parameters for the product spreads.

In case of limited historical futures data, one could also test the forecasting ability of product spreads determined by their spot prices:

$$\frac{S^{oil}_{t+h}}{S^{oil}_t} = \alpha + \beta \left[\sum_{j=1}^{n} w_{j,t} \frac{F^{refined\ prod.}_{j,t+h|t}}{S^{oil}_t} \right] + \varepsilon_{t+h}$$

Single future spreads:

$$\hat{S}^{oil}_{t+h|t} = S^{oil}_t \left\{ \hat{\alpha} + \hat{\beta} \left[\frac{F^{refined\ prod.}_{j,t}}{S^{oil}_t} \right] \right\}$$

Single spot spreads:

$$\hat{S}^{oil}_{t+h|t} = S^{oil}_t \left\{ \hat{\alpha} + \hat{\beta} \left[\frac{S^{refined\ prod.}_{j,t}}{S^{oil}_t} \right] \right\}$$

Exhibit 3: Underlying assumptions in linear regression and OLS

The following assumptions describes the underlying properties that need to be fulfilled in order for Ordinary Least Square to provide unbiased and consistent coefficient estimates (Wooldridge, 2016, p. 40)

1. The conditional error distribution has a mean of zero;
$$\mathrm{E}[u_i|X_{1i}, X_{2i}, \dots, X_{ni}] = 0$$
2. $(X_{1i}, X_{2i}, \dots, X_{ni}, Y_i)$, $i = 1, \dots, n$ are independently and identically distributed (i.i.d).
3. Large outliers are unlikely;
$$0 < \mathrm{E}\left[{X_i}^4\right] < \infty \;\; and \;\; 0 < \mathrm{E}\left[{Y_i}^4\right] < \infty$$
4. No perfect multicollinearity (no linear dependence)

Furthermore, in order for OLS to generate the Best Linear Unbiased Estimator (BLUE) the following assumptions regarding the residual must be satisfied;

5. The variance of the error terms is constant (i.e., homoskedasticity);
$$Var(u_i|X_{1i}, X_{2i}, \dots, X_{ni}) = \sigma^2$$

6. Normality assumption: the error terms is normally distributed with zero mean and constant variance;
$$u_i \sim \mathcal{N}(0, \sigma^2)$$

Exhibit 4: Wald test

The Wald test (Wooldridge, 2010, p. 46) can be used to jointly test multiple parameter restrictions. With Q restrictions of P parameters, the test statistic is:

$$W = (R\hat{\theta}_n - r)' \left[R \left(\frac{\hat{V}_n}{n} \right) R' \right]^{-1} (R\hat{\theta}_n - r)$$

r is a $Q \times 1$ nonrandom vector, and $\hat{\theta}_n$ is a vector of our sample estimator of the parameters. R is the Jacobian of r, and $\hat{V}_n$ is an estimator of the covariance matrix which can be chosen to be fully robust. This tests the following null hypothesis against the alternative:

$$H_0: R\theta = r$$

$$H_1: R\theta \neq r$$

Under the null hypothesis, $W \sim \chi^2_Q$.

Exhibit 5: Diebold & Mariano test

The DM-test is used to test whether two forecasts are equally good. Considering two forecasts, $\hat{y}_{1t}$ and $\hat{y}_{2t}$ for $t = 1, \dots, T$.

Define the forecast errors as:

$$e_{it} = \hat{y}_{it} - y_t, \quad i = 1,2$$

And a loss function, e.g. as squared errors:

$$g(e_{it}) = e_{it}^2$$

Define the loss differential between the two forecasts as:

$$d_t = g(e_{1t}) - g(e_{2t})$$

DM tests for the null hypothesis against the alternative (a two-sided test):

$$H_0: E(d_t) = 0 \quad \forall t$$
$$H_1: E(d_t) \neq 0$$

Under the null hypothesis, the test statistic becomes:

$$DM = \frac{\bar{d}}{\hat{\sigma}_d} \rightarrow N(0,1)$$

where $\bar{d}$ is the sample mean loss differential and $\hat{\sigma}_d$ is a consistent estimate of the standard deviation of $\bar{d}$. We use the heteroscedasticity and autocorrelation robust (HAC) standard errors defined by Newey & West (1987).

Exhibit 6a: Descriptive statistics for the raw data

Monthly data	Mean	Median	Max	Min	St. dev.	Skewness	Kurtosis	J-B	p-value	Period	n
S_t^{oil}	43,917	30,332	133,88	11,361	29,373	0,886	2,609	54,856	0	1986/1 – 2019/4	400
$F_{t+h\|t}^{oil}$											
$h = 1$	42,779	29,693	134,016	11,311	28,471	1,004	2,876	73,257	0	1983/3 – 2019/4	434
$h = 2$	42,862	29,518	134,523	11,305	28,645	0,985	2,819	70,833	0	1983/3 – 2019/4	434
$h = 3$	42,901	29,196	134,780	11,355	28,778	0,971	2,771	69,079	0	1983/3 – 2019/4	434
$h = 4$	42,907	28,977	134,888	11,489	28,875	0,959	2,732	67,792	0	1983/3 – 2019/4	434
$h = 6$	42,862	28,242	135,367	11,634	28,993	0,941	2,673	65,974	0	1983/3 – 2019/4	434
$h = 12$	46,325	29,579	135,123	13,017	29,997	0,669	2,170	37,847	0	1983/3 – 2019/4	404
$S_t^{gasoline}$	53,401	37,336	138,27	12,902	34,727	0,789	2,335	48,304	0	1986/6 – 2019/4	395
$F_{t+h\|t}^{gasoline}$											
$h = 1$	53,062	36,364	144,356	14,218	34,704	0,867	2,502	55,876	0,000	1985/1 – 2019/4	412
$h = 2$	62,967	59,678	143,550	15,076	34,859	0,435	2,012	21,959	0,000	1994/1 – 2019/4	304
$h = 3$	52,556	34,783	142,906	13,906	34,355	0,836	2,420	53,940	0,000	1984/12 – 2019/4	413
$h = 4$	62,569	60,980	137,022	16,829	34,442	0,390	1,932	22,166	0,000	1994/1 – 2019/4	304
$h = 6$	89,409	83,566	136,904	52,302	22,098	0,217	1,865	9,846	0,007	2006/1 – 2019/4	160
$h = 12$	88,730	87,550	145,432	45,722	21,685	0,256	2,242	5,579	0,061	2006/1 – 2019/4	160

Exhibit 6b: Descriptive statistics for the raw data

Monthly data	Mean	Median	Max	Min	St. dev.	Skewness	Kurtosis	J-B	p-value	Period	n
S_t^{heat}	53,911	36,679	159,63	12,79	37,075	0,870	2,538	53,307	0	1986/6 – 2019/4	395
$F_{t+h\|t}^{heat}$											
$h = 1$	51,270	34,072	160,106	13,086	35,346	1,087	3,016	92,911	0	1980/1 – 2019/4	472
$h = 2$	51,307	33,824	161,173	13,308	35,494	1,076	2,984	91,148	0	1980/1 – 2019/4	472
$h = 3$	51,386	33,889	162,551	13,564	35,639	1,069	2,963	89,893	0	1980/1 – 2019/4	472
$h = 4$	51,446	34,125	163,723	13,850	35,760	1,063	2,947	88,955	0	1980/1 – 2019/4	472
$h = 6$	51,520	33,811	165,743	14,568	35,932	1,053	2,923	87,327	0	1980/1 – 2019/4	472
$h = 12$	53,896	33,109	160,813	13,915	37,337	0,830	2,380	55,206	0	1980/1 – 2019/4	422
$invUS$	329,43	318,14	531,84	250,33	53,02	1,818	6,188	429,68	0	1982/8 – 2019/4	441
$prodUS$	7054,6	6626,5	12200	3960,75	1574,15	0,669	2,931	32,632	0	1983/1 – 2019/4	436
$rigcountGlob$	2699,6	2336,0	6227	1156,00	966,42	1,036	3,957	115,23	0	1975/1 – 2019/3	531
$S\&P500\ index$	1065,4	1088,1	2903,8	151,08	685,33	0,680	2,8756	33,015	0	1984/1 – 2019/5	425
$US\ 3m\ Tbill\ yield$	0,0359	0,0366	0,1095	0	0,0282	0,296	2,1000	20,512	0	1984/1 – 2019/4	424
$US\ 10y\ Tbond\ yield$	0,0539	0,0503	0,1362	0,0145	0,0261	0,6526	2,9987	30,097	0	1984/1 – 2019/4	424
$baa\ corp\ bond\ yield$	0,0770	0,0755	0,1515	0,0422	0,0235	0,7365	3,2689	39,610	0	1984/1 – 2019/4	424
Mkt_{FF}	0,0068	0,0113	0,1247	-0,2324	0,0439	-0,8603	5,6968	180,780	0	1984/1 – 2019/4	424
SMB_{FF}	0,0003	-0,00025	0,2171	-0,1687	0,0303	0,7525	11,3990	1286,400	0	1984/1 – 2019/4	424
HML_{FF}	0,0020	-0,00045	0,1290	-0,1110	0,0288	0,2093	5,3615	101,620	0	1984/1 – 2019/4	424

$invUS =$ Crude Oil US Stocks excluding SPR (million barrels), $prodUS =$ US Crude Oil Production (thousand barrels/day), $rigcountGlob =$ Global rig count from Baker Hughes

Exhibit 7a: Descriptive statistics for oil price change, and crude oil futures and gasoline futures normalised by the crude oil spot

Monthly variables	Mean	Median	Max	Min	St. dev.	Skewness	Kurtosis	J-B	p-value	Period	n
S_{t+h}^{oil}/S_t^{oil}											
$h=1$	1,006	1,012	1,480	0,674	0,086	0,141	5,966	147,613	0	1986/1 – 2019/3	399
$h=2$	1,015	1,017	1,816	0,536	0,136	0,390	7,386	329,147	0	1986/1 – 2019/2	398
$h=3$	1,024	1,028	2,007	0,395	0,173	0,621	7,806	407,658	0	1986/1 – 2019/1	397
$h=4$	1,033	1,033	2,159	0,352	0,201	0,587	6,692	247,626	0	1986/1 – 2018/12	396
$h=6$	1,049	1,056	1,956	0,307	0,237	0,226	4,154	25,234	0	1986/1 – 2018/10	394
$h=12$	1,095	1,072	2,444	0,410	0,331	0,650	3,948	41,817	0	1986/1 – 2018/4	388
$F_{t+h\|t}^{oil}/S_t^{oil}$											
$h=1$	1,000	1,000	1,023	0,979	0,004	-0,146	10,071	834,833	0	1986/1 – 2019/4	400
$h=2$	1,000	1,003	1,112	0,903	0,023	0,294	6,748	239,877	0	1986/1 – 2019/4	400
$h=3$	1,000	1,003	1,177	0,851	0,038	0,307	5,911	147,481	0	1986/1 – 2019/4	400
$h=4$	0,999	1,002	1,220	0,820	0,049	0,304	5,190	86,090	0	1986/1 – 2019/4	400
$h=6$	0,997	0,998	1,279	0,794	0,068	0,308	4,411	39,510	0	1986/1 – 2019/4	400
$h=12$	0,989	0,988	1,393	0,723	0,106	0,371	3,789	18,832	0	1986/1 – 2019/4	385
$F_{t+h\|t}^{gasoline}/S_t^{oil}$											
$h=1$	1,235	1,224	1,603	0,907	0,111	0,544	3,452	23,149	0	1986/1 – 2019/4	400
$h=2$	1,235	1,228	1,712	0,932	0,112	0,622	4,171	36,951	0	1994/1 – 2019/4	304
$h=3$	1,225	1,216	1,769	0,958	0,108	0,958	5,633	176,823	0	1986/1 – 2019/4	400
$h=4$	1,229	1,216	1,789	0,986	0,119	1,140	5,847	168,539	0	1994/1 – 2019/4	304
$h=6$	1,236	1,208	1,768	1,003	0,133	1,245	5,474	82,128	0	2006/1 – 2019/4	160
$h=12$	1,227	1,212	1,563	0,989	0,120	0,547	2,790	8,280	0,016	2006/1 – 2019/4	160

Exhibit 7b: Descriptive statistics for heating oil futures normalized by the crude oil spot, change in US oil inventory and global rig count

Monthly variables	Mean	Median	Max	Min	St. dev.	Skewness	Kurtosis	J-B	p-value	Period	n
$F_{t+h\|t}^{heating}/S_t^{oil}$											
$h = 1$	1,211	1,186	1,628	1,000	0,112	0,930	3,675	65,200	0	1986/1 – 2019/4	400
$h = 2$	1,210	1,189	1,549	0,985	0,109	0,759	3,323	40,132	0	1986/1 – 2019/4	400
$h = 3$	1,210	1,198	1,552	0,950	0,113	0,638	3,226	28,023	0	1986/1 – 2019/4	400
$h = 4$	1,210	1,197	1,563	0,939	0,118	0,518	3,104	18,084	0	1986/1 – 2019/4	400
$h = 6$	1,209	1,199	1,595	0,924	0,131	0,374	2,952	9,375	0,009	1986/1 – 2019/4	400
$h = 12$	1,208	1,201	1,741	0,868	0,163	0,471	3,151	14,830	0,001	1986/1 – 2019/4	391
inv_t/inv_{t-h}											
$h = 1$	1,001	1,002	1,089	0,934	0,026	0,120	2,962	0,992	0,609	1986/1 – 2019/4	400
$h = 2$	1,003	1,002	1,176	0,904	0,044	0,327	3,214	7,895	0,019	1986/1 – 2019/4	400
$h = 3$	1,005	1,002	1,247	0,881	0,056	0,474	3,545	19,918	0,000	1986/1 – 2019/4	400
$h = 4$	1,006	1,000	1,293	0,866	0,064	0,570	3,829	33,162	0	1986/1 – 2019/4	400
$h = 6$	1,009	1,001	1,324	0,813	0,073	0,663	4,605	72,271	0	1986/1 – 2019/4	400
$h = 12$	1,014	1,009	1,309	0,805	0,086	0,498	3,822	27,824	0	1986/1 – 2019/4	400
$rigs_t/rigs_{t-h}$											
$h = 1$	1,000	1,007	1,150	0,800	0,049	-0,782	3,951	55,823	0	1986/1 – 2019/4	400
$h = 2$	1,001	1,008	1,189	0,687	0,084	-0,648	3,527	32,616	0	1986/1 – 2019/4	400
$h = 3$	1,003	1,014	1,251	0,613	0,107	-0,590	3,686	31,078	0	1986/1 – 2019/4	400
$h = 4$	1,004	1,016	1,313	0,555	0,124	-0,548	3,910	33,845	0	1986/1 – 2019/4	400
$h = 6$	1,006	1,021	1,446	0,507	0,152	-0,430	3,924	26,534	0	1986/1 – 2019/4	400
$h = 12$	1,008	1,039	1,486	0,504	0,200	-0,418	2,966	11,663	0,003	1986/1 – 2019/4	400

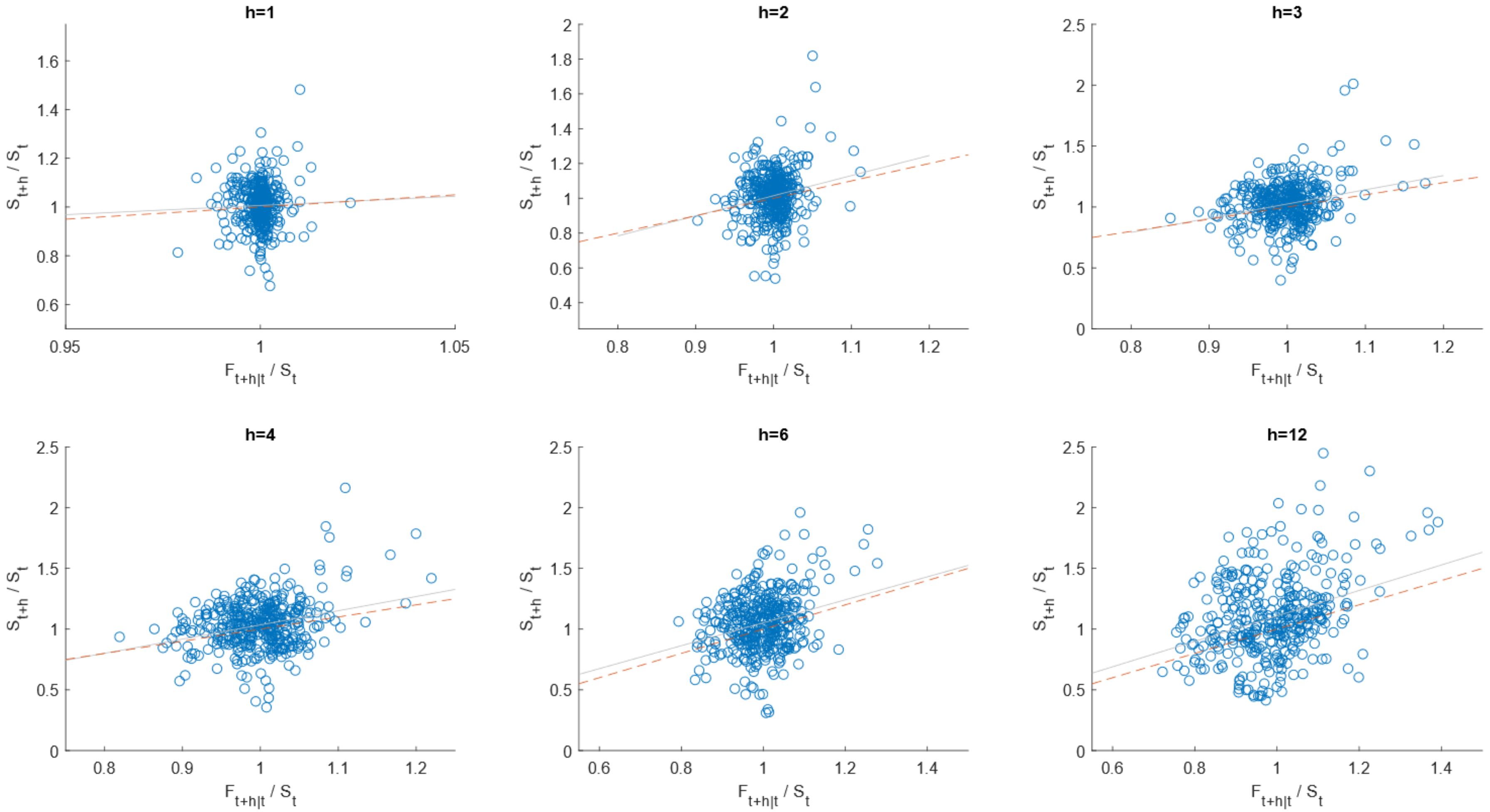
Exhibit 7c: Descriptive statistics for change in US crude oil production, S&P 500 excess return and change in yield curve

Monthly variables	Mean	Median	Max	Min	St. dev.	Skewness	Kurtosis	J-B	p-value	Period	n
$prod_t/prod_{t-h}$											
$h=1$	1,001	1,000	1,165	0,780	0,024	-2,127	39,889	22981,293	0	1986/1 – 2019/4	400
$h=2$	1,002	1,000	1,259	0,751	0,035	-0,282	25,845	8703,716	0	1986/1 – 2019/4	400
$h=3$	1,003	1,000	1,267	0,771	0,041	-0,008	16,669	3114,012	0	1986/1 – 2019/4	400
$h=4$	1,004	0,999	1,272	0,739	0,046	0,041	12,363	1461,299	0	1986/1 – 2019/4	400
$h=6$	1,006	0,999	1,371	0,745	0,056	0,692	10,482	964,911	0	1986/1 – 2019/4	400
$h=12$	1,010	0,991	1,341	0,782	0,078	1,109	4,441	116,587	0	1986/1 – 2019/4	400
$R_{t-h,t}^{S\&P} - R_{t-h,t}^{Tbill}$											
$h=1$	0,005	0,009	0,120	-0,205	0,035	-1,059	7,410	398,884	0	1986/1 – 2019/4	400
$h=2$	0,010	0,014	0,192	-0,275	0,055	-1,095	6,815	322,409	0	1986/1 – 2019/4	400
$h=3$	0,015	0,020	0,223	-0,313	0,070	-0,995	6,025	218,604	0	1986/1 – 2019/4	400
$h=4$	0,020	0,024	0,236	-0,317	0,082	-0,936	5,306	147,018	0	1986/1 – 2019/4	400
$h=6$	0,030	0,038	0,379	-0,379	0,104	-0,808	5,063	114,400	0	1986/1 – 2019/4	400
$h=12$	0,063	0,080	0,521	-0,434	0,155	-0,764	3,764	48,665	0	1986/1 – 2019/4	400
$\Delta\left(R_t^{Tbond} - R_t^{Tbill}\right)$											
$h=1$	-0,005	0,000	0,874	-0,869	0,244	0,341	3,734	16,728	0	1986/1 – 2019/4	400
$h=2$	-0,010	-0,113	1,415	-1,105	0,384	0,522	3,678	25,867	0	1986/1 – 2019/4	400
$h=3$	-0,017	-0,020	1,624	-1,105	0,484	0,437	3,180	13,286	0,001	1986/1 – 2019/4	400
$h=4$	-0,024	-0,066	1,880	-1,437	0,583	0,332	3,107	7,532	0,023	1986/1 – 2019/4	400
$h=6$	-0,036	-0,123	2,509	-2,089	0,759	0,242	2,995	3,909	0,142	1986/1 – 2019/4	400
$h=12$	-0,068	-0,237	3,970	-2,569	1,177	0,343	3,000	7,857	0,020	1986/1 – 2019/4	400

Exhibit 7d: Descriptive statistics for change in the credit spread and realized volatility of crude oil futures

Monthly variables	Mean	Median	Max	Min	St. dev.	Skewness	Kurtosis	J-B	p-value	Period	n
$\Delta\left(R_t^{corp} - R_t^{Tbond}\right)$											
$h = 1$	-0,005	0,000	0,874	-0,869	0,244	0,341	3,734	16,728	0,000	1986/1 – 2019/4	400
$h = 2$	-0,010	-0,113	1,415	-1,105	0,384	0,522	3,678	25,867	0,000	1986/1 – 2019/4	400
$h = 3$	-0,017	-0,020	1,624	-1,105	0,484	0,437	3,180	13,286	0,001	1986/1 – 2019/4	400
$h = 4$	-0,024	-0,066	1,880	-1,437	0,583	0,332	3,107	7,532	0,023	1986/1 – 2019/4	400
$h = 6$	-0,036	-0,123	2,509	-2,089	0,759	0,242	2,995	3,909	0,142	1986/1 – 2019/4	400
$h = 12$	-0,068	-0,237	3,970	-2,569	1,177	0,343	3,000	7,857	0,020	1986/1 – 2019/4	400
$\frac{\sum_{i=1}^{n}\left(F_i^{oil} - \bar{F}^{oil}\right)^2}{n-1}$											
$h = 1$	3,944	1,389	125,018	0,029	8,823	8,062	97,104	151 926	0	1986/1 – 2019/4	400
$h = 2$	8,547	2,759	269,557	0,070	20,294	7,657	83,245	111 229	0	1986/1 – 2019/4	400
$h = 3$	13,666	3,874	417,552	0,108	35,700	7,358	69,748	77 865	0	1986/1 – 2019/4	400
$h = 4$	19,317	5,311	591,149	0,139	55,123	7,666	71,837	82 894	0	1986/1 – 2019/4	400
$h = 6$	31,257	7,439	1111,691	0,316	96,470	7,682	71,457	82 041	0	1986/1 – 2019/4	400
$h = 12$	62,814	14,514	1429,529	0,778	172,457	5,560	37,744	22 180	0	1986/1 – 2019/4	400

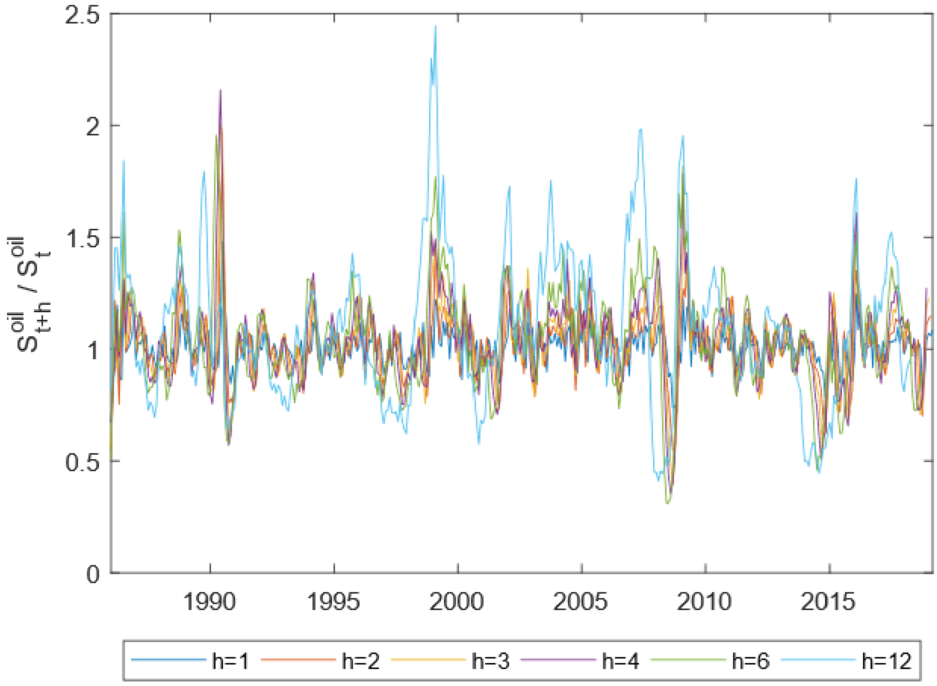
Exhibit 8: Scatterplot of futures price and future spot price, normalized by contemporaneous spot price.



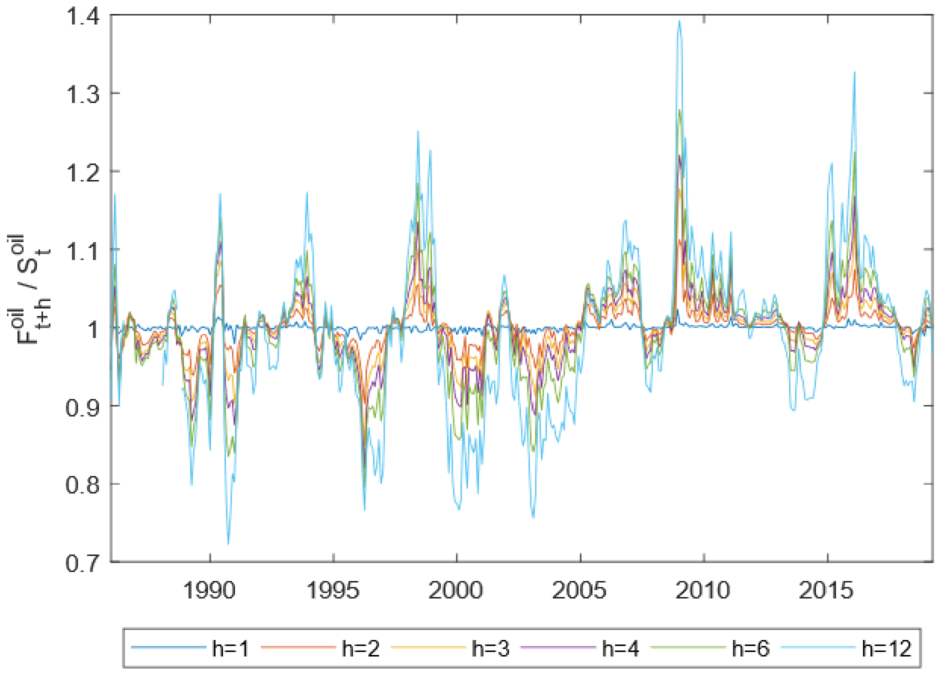
These variables can also be interpreted as spot price changes (Y) and futures spread (X) in gross values. Grey line is superimposed least-squares line, and dashed line is the diagonal from origo.

Exhibit 9: Unit root test and plots

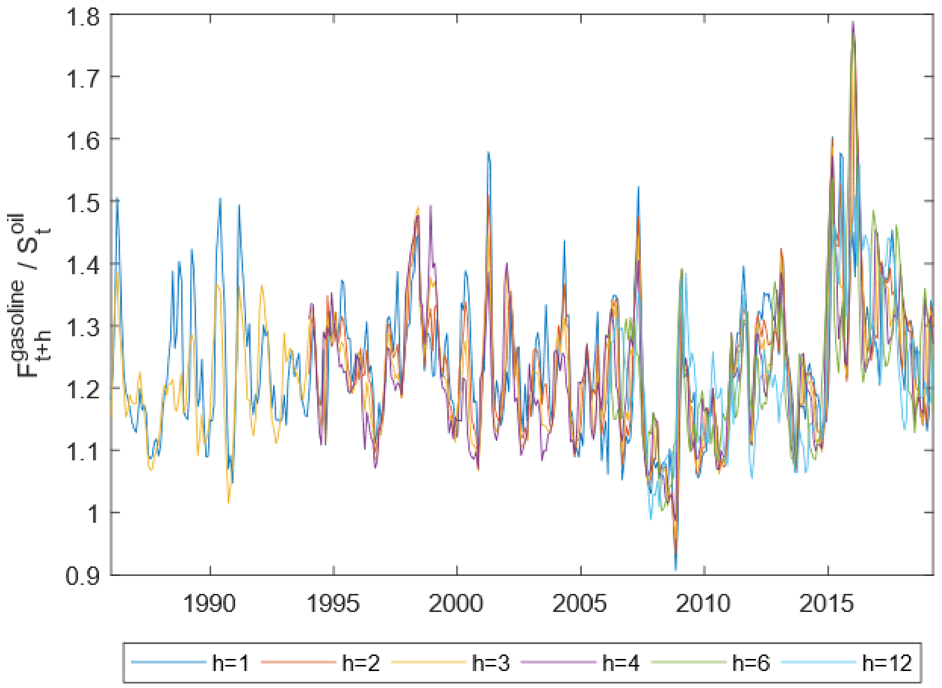
Spot change	Horizons					
	$h = 1$	$h = 2$	$h = 3$	$h = 4$	$h = 6$	$h = 12$
ADF p-value	0	0	0	0	0	0
ADF-stat	-15,579	-9,275	-6,690	-7,865	-5,535	-4,281
No. lags	0	5	6	5	7	12



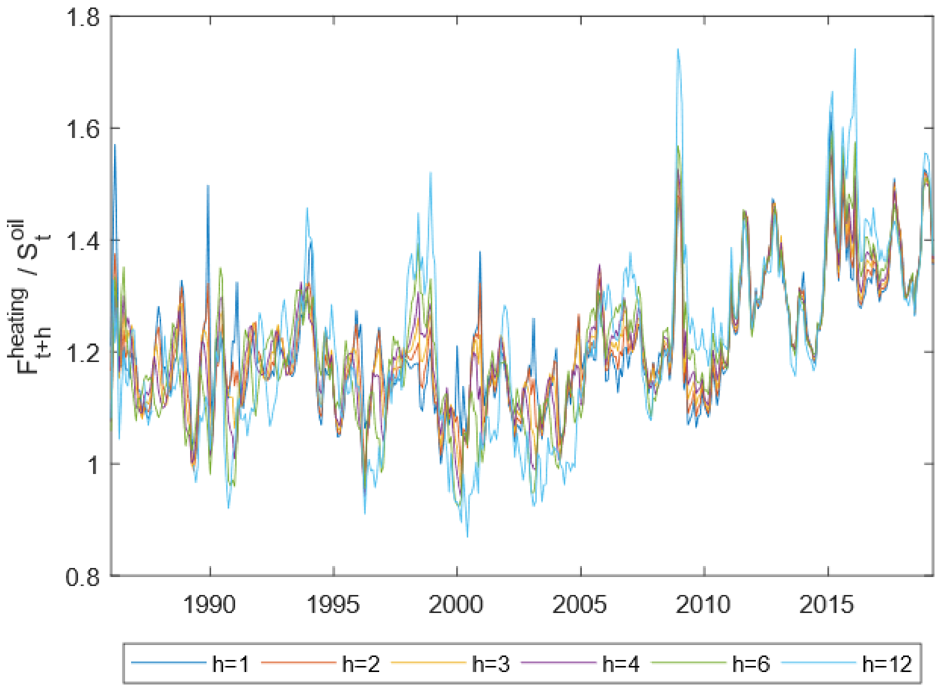
Futures price	Horizons					
	$h = 1$	$h = 2$	$h = 3$	$h = 4$	$h = 6$	$h = 12$
ADF p-value	0	0	0	0	0	0
ADF-stat	-9,592	-6,118	-6,204	-5,948	-5,642	-5,215
No. lags	1	0	1	1	1	1



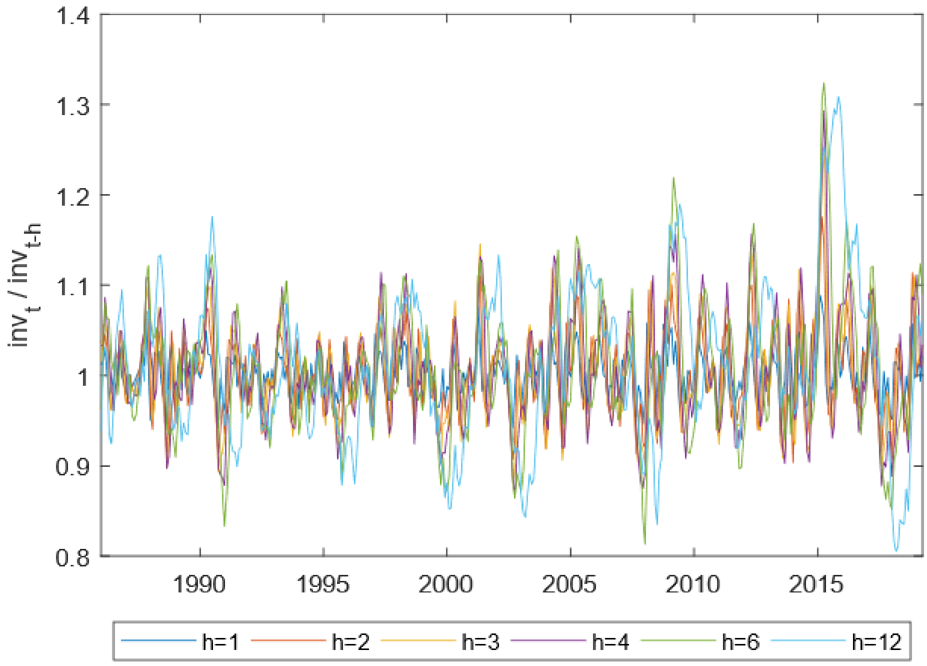
Gasoline spread	**Horizons**					
	$h = 1$	$h = 2$	$h = 3$	$h = 4$	$h = 6$	$h = 12$
ADF p-value	0	0	0,008	0,013	0,262	0,037
ADF-stat	-8,603	-6,686	-3,594	-3,419	-2,087	-3,074
No. lags	1	1	12	12	6	0



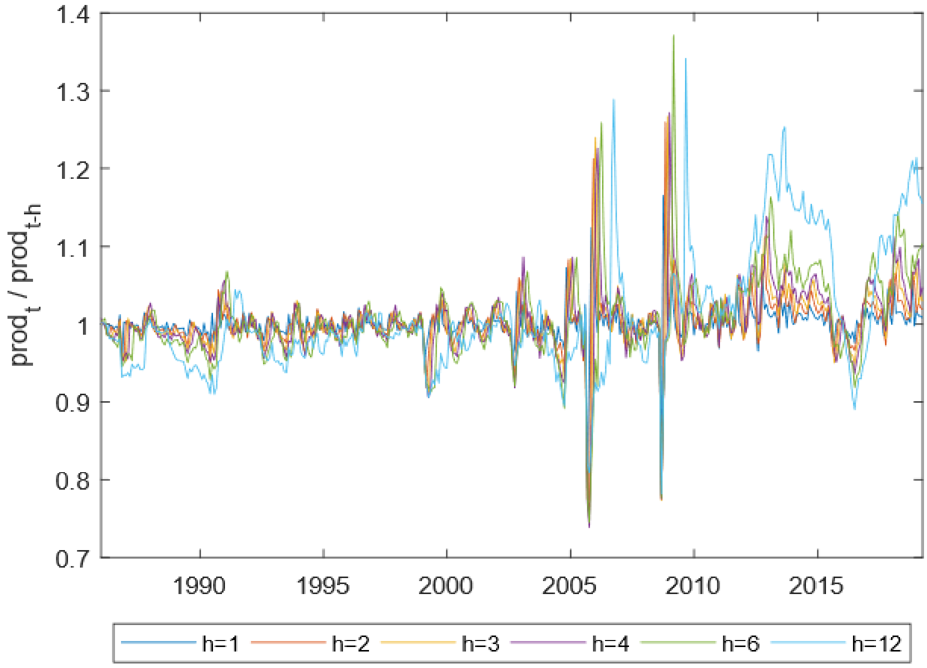
Heating oil spread	**Horizons**					
	$h = 1$	$h = 2$	$h = 3$	$h = 4$	$h = 6$	$h = 12$
ADF p-value	0	0	0	0	0	0,001
ADF-stat	-6,003	-5,458	-5,199	-5,046	-4,777	-4,136
No. lags	1	1	1	1	1	1



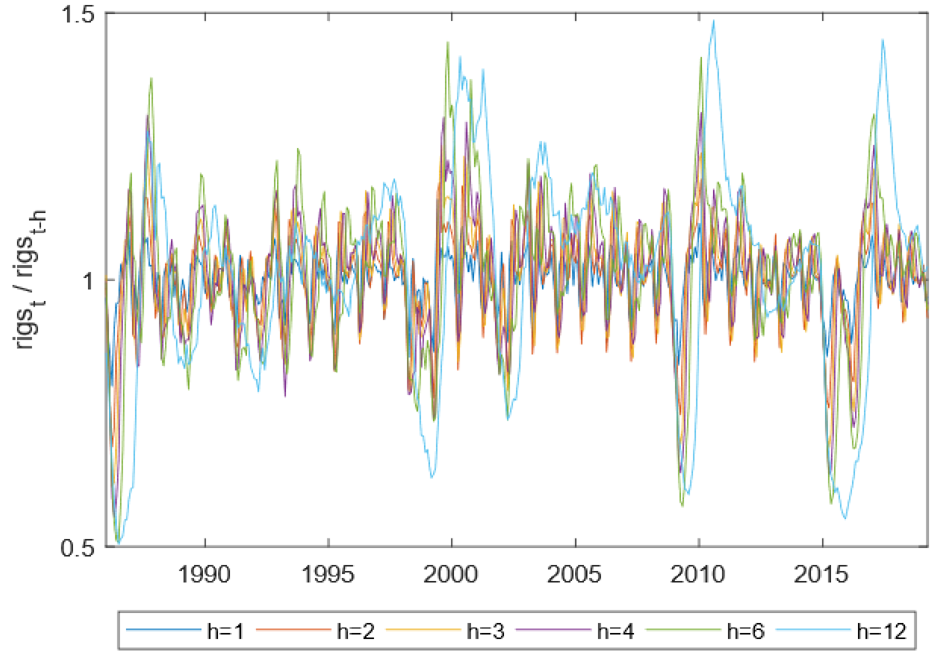
US inventory changes	**Horizons**					
	$h = 1$	$h = 2$	$h = 3$	$h = 4$	$h = 6$	$h = 12$
ADF p-value	0	0	0	0	0	0,012
ADF-stat	-4,361	-4,772	-4,663	-4,833	-4,762	-3,443
No. lags	11	11	10	9	7	12



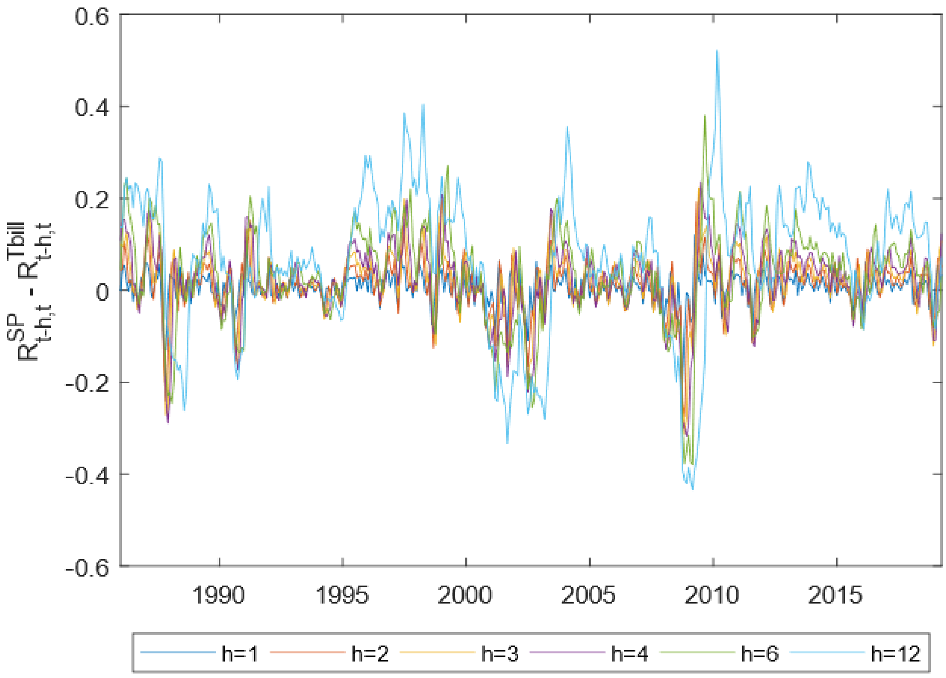
US production changes	**Horizons**					
	$h = 1$	$h = 2$	$h = 3$	$h = 4$	$h = 6$	$h = 12$
ADF p-value	0	0	0,003	0,003	0,004	0,186
ADF-stat	-20,029	-6,611	-3,905	-3,912	-3,803	-2,273
No. lags	0	6	9	8	6	12



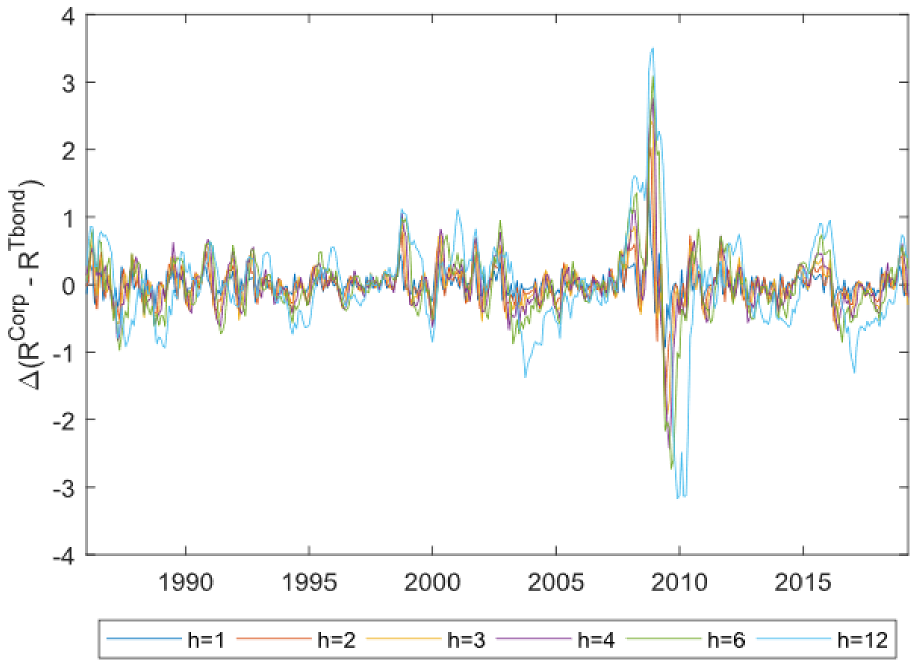
Global rig changes	**Horizons**					
	$h = 1$	$h = 2$	$h = 3$	$h = 4$	$h = 6$	$h = 12$
ADF p-value	0	0	0	0	0	0
ADF-stat	-4,572	-5,770	-6,974	-6,473	-6,452	-6,901
No. lags	12	12	12	11	9	3



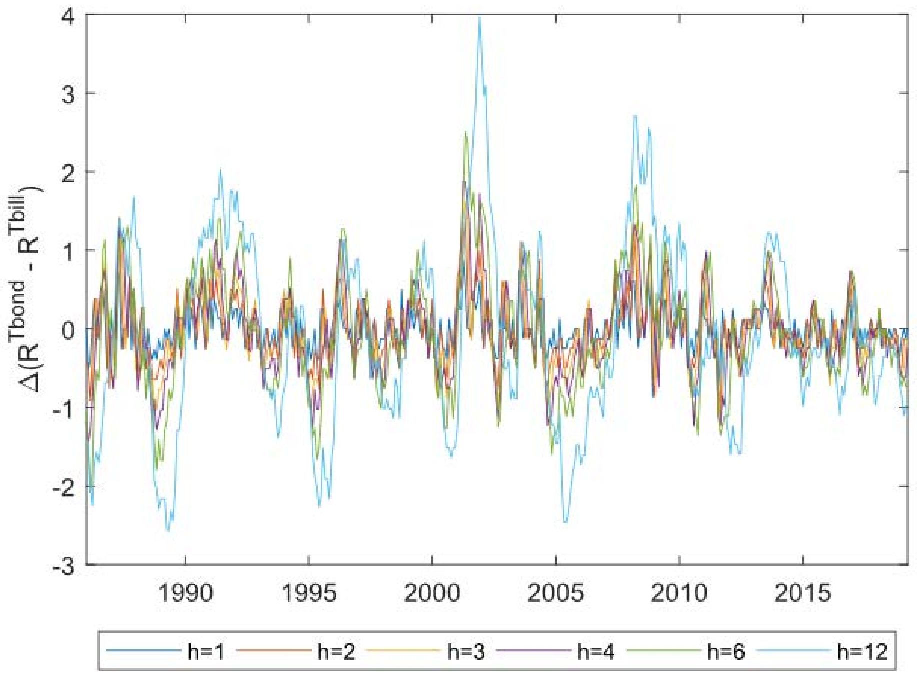
SP500 excess return changes	**Horizons**					
	$h = 1$	$h = 2$	$h = 3$	$h = 4$	$h = 6$	$h = 12$
ADF p-value	0	0	0	0	0,008	0,027
ADF-stat	-15,272	-6,453	-4,696	-5,144	-3,558	-3,136
No. lags	0	7	9	9	12	12



Credit spread changes	**Horizons**					
	$h = 1$	$h = 2$	$h = 3$	$h = 4$	$h = 6$	$h = 12$
ADF p-value	0	0	0	0	0	0
ADF-stat	-14,153	-7,682	-7,552	-6,249	-4,279	-3,578
No. lags	0	7	7	9	12	12



Yield curve changes	**Horizons**					
	$h = 1$	$h = 2$	$h = 3$	$h = 4$	$h = 6$	$h = 12$
ADF p-value	0	0	0	0	0	0
ADF-stat	-9,431	-5,540	-5,117	-5,861	-4,293	-3,359
No. lags	2	7	10	11	12	12



Realized volatility	**Horizons**					
	$h = 1$	$h = 2$	$h = 3$	$h = 4$	$h = 6$	$h = 12$
ADF p-value	0	0	0	0	0	0,001
ADF-stat	-4,888	-7,113	-6,334	-4,893	-5,908	-4,160
No. lags	2	1	2	3	4	4

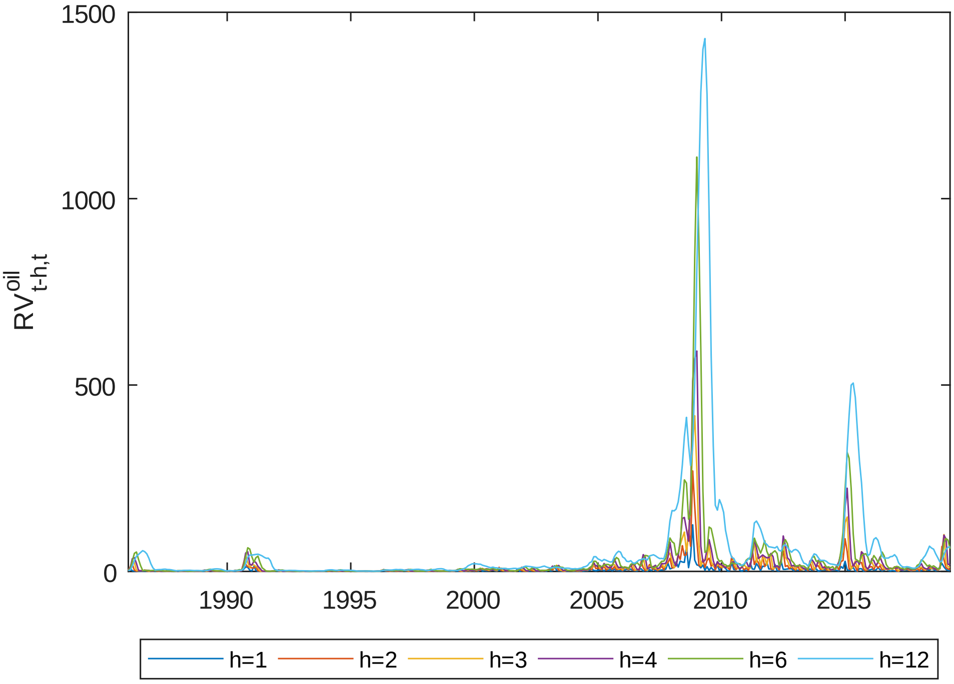


Exhibit 10: Correlation matrix

$h = 1$	dSpot	Fut	Gas	Heat	Crack	dInv	dProd	dRig	SP	dSlope	dCred	RV
dSpot	1,00											
Fut	0,03	1,00										
Gas	0,14	0,08	1,00									
Heat	-0,01	0,17	0,18	1,00								
Crack	0,12	0,14	0,91	0,57	1,00							
dInv	0,00	0,26	0,15	0,15	0,19	1,00						
dProd	0,00	0,05	-0,14	0,17	-0,04	0,19	1,00					
dRig	-0,06	-0,10	-0,28	-0,16	-0,30	-0,34	0,03	1,00				
SP	0,07	0,07	0,09	-0,01	0,07	0,00	-0,07	-0,07	1,00			
dSlope	0,01	-0,10	0,01	-0,09	-0,03	0,05	0,05	-0,13	-0,06	1,00		
dCred	-0,22	-0,03	-0,12	0,05	-0,07	0,03	0,05	-0,02	-0,42	-0,26	1,00	
RV	-0,17	0,13	-0,26	0,21	-0,13	0,09	0,20	-0,04	-0,18	0,03	0,24	1,00

$h = 2$	dSpot	Fut	Gas	Heat	Crack	dInv	dProd	dRig	SP	dSlope	dCred	RV
dSpot	1,00											
Fut	0,20	1,00										
Gas	0,21	0,29	1,00									
Heat	0,02	0,46	0,28	1,00								
Crack	0,15	0,42	0,91	0,64	1,00							
dInv	0,08	0,33	0,34	0,14	0,33	1,00						
dProd	0,00	0,11	-0,07	0,24	0,04	0,16	1,00					
dRig	-0,06	-0,18	-0,35	-0,08	-0,34	-0,44	0,01	1,00				
SP	0,16	-0,02	0,10	-0,06	0,06	-0,02	-0,05	-0,07	1,00			
dSlope	-0,04	-0,08	-0,01	-0,16	-0,07	0,07	0,04	-0,12	-0,06	1,00		
dCred	-0,21	-0,02	-0,15	0,12	-0,07	0,04	0,03	-0,01	-0,49	-0,19	1,00	
RV	-0,15	0,24	-0,27	0,29	-0,10	0,11	0,25	-0,07	-0,30	0,01	0,38	1,00

$h = 3$	dSpot	Fut	Gas	Heat	Crack	dInv	dProd	dRig	SP	dSlope	dCred	RV
dSpot	1,00											
Fut	0,25	1,00										
Gas	0,30	0,43	1,00									
Heat	0,04	0,58	0,34	1,00								
Crack	0,25	0,57	0,92	0,67	1,00							
dInv	0,14	0,37	0,40	0,16	0,38	1,00						
dProd	0,05	0,15	0,07	0,24	0,15	0,15	1,00					
dRig	-0,14	-0,22	-0,46	-0,08	-0,39	-0,50	-0,01	1,00				
SP	0,11	-0,06	0,17	-0,10	0,09	-0,05	-0,02	-0,07	1,00			
dSlope	0,02	-0,06	-0,03	-0,16	-0,09	0,08	-0,03	-0,09	-0,05	1,00		
dCred	-0,16	0,04	-0,15	0,18	-0,04	0,04	0,01	-0,01	-0,53	-0,16	1,00	
RV	-0,08	0,33	-0,11	0,34	0,05	0,14	0,22	-0,10	-0,35	0,03	0,44	1,00

$h = 4$	dSpot	Fut	Gas	Heat	Crack	dInv	dProd	dRig	SP	dSlope	dCred	RV
dSpot	1,00											
Fut	0,29	1,00										
Gas	0,32	0,53	1,00									
Heat	0,06	0,67	0,46	1,00								
Crack	0,24	0,66	0,94	0,75	1,00							
dInv	0,14	0,40	0,39	0,21	0,38	1,00						
dProd	0,04	0,14	0,18	0,21	0,22	0,16	1,00					
dRig	-0,18	-0,24	-0,41	-0,13	-0,39	-0,49	-0,01	1,00				
SP	0,06	-0,08	0,10	-0,10	0,03	-0,06	0,00	-0,08	1,00			
dSlope	0,03	-0,04	-0,07	-0,14	-0,11	0,06	-0,03	-0,07	-0,05	1,00		
dCred	-0,07	0,08	-0,03	0,20	0,07	0,03	-0,02	-0,01	-0,56	-0,14	1,00	
RV	0,02	0,38	-0,02	0,35	0,13	0,19	0,21	-0,13	-0,37	0,04	0,45	1,00

$\boldsymbol{h=6}$	dSpot	Fut	Gas	Heat	Crack	dInv	dProd	dRig	SP	dSlope	dCred	RV
dSpot	1,00											
Fut	0,27	1,00										
Gas	0,36	0,54	1,00									
Heat	0,08	0,77	0,73	1,00								
Crack	0,31	0,56	0,98	0,86	1,00							
dInv	-0,02	0,44	0,13	0,39	0,21	1,00						
dProd	-0,10	0,10	0,00	0,25	0,02	0,14	1,00					
dRig	-0,14	-0,27	-0,11	-0,32	-0,18	-0,46	-0,01	1,00				
SP	-0,01	-0,07	-0,05	-0,02	-0,12	-0,08	0,01	-0,08	1,00			
dSlope	-0,05	-0,06	-0,35	-0,11	-0,38	0,02	0,02	-0,04	-0,09	1,00		
dCred	0,05	0,15	0,07	0,17	0,14	0,05	-0,03	-0,04	-0,59	-0,11	1,00	
RV	0,15	0,43	0,09	0,34	0,16	0,25	0,16	-0,16	-0,39	0,06	0,46	1,00

$\boldsymbol{h=12}$	dSpot	Fut	Gas	Heat	Crack	dInv	dProd	dRig	SP	dSlope	dCred	RV
dSpot	1,00											
Fut	0,34	1,00										
Gas	0,45	0,57	1,00									
Heat	0,19	0,85	0,68	1,00								
Crack	0,45	0,67	0,96	0,86	1,00							
dInv	0,20	0,62	0,60	0,52	0,63	1,00						
dProd	-0,10	0,12	-0,12	0,37	-0,06	0,03	1,00					
dRig	-0,24	-0,31	-0,39	-0,28	-0,41	-0,37	0,01	1,00				
SP	-0,01	-0,08	-0,09	-0,01	-0,16	-0,14	0,13	0,04	1,00			
dSlope	-0,13	-0,08	-0,39	-0,15	-0,39	-0,03	0,04	-0,02	-0,28	1,00		
dCred	0,04	0,22	0,05	0,19	0,15	0,14	-0,14	-0,24	-0,53	0,08	1,00	
RV	0,07	0,39	0,16	0,30	0,17	0,32	0,15	-0,29	-0,40	0,18	0,38	1,00