MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master‘s Thesis

„Oil Prices: Analysis of High-Frequency Data“

verfasst von / submitted by

Anna Kaushik BSc

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Master of Science (MSc)

Wien, 2016 / Vienna, 2016

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:	A 066 914
Studienrichtung lt. Studienblatt / degree programme as it appears on the student record sheet:	Internationale Betriebswirtschaft
Betreut von / Supervisor:	Univ.-Prof. Dr. Franz Wirl

Oil Prices: Analysis of High-Frequency Data

Abstract

Traditional macroeconomic demand and supply arguments fail to explain the incessant fluctuations in today's oil prices. This gives rise to the question: what drives oil prices and can forecasting help identify the future direction of oil prices in the presence of volatility? Mixed-data sampling (MIDAS) regression is a topic of growing interest since the variable in question is often desired at a lower frequency while the relevant information is also available at a higher frequency. In this context, the role of high-frequency financial and energy market data in predicting the price of crude oil is of increasing importance since these markets tend to catch the effects of volatility almost immediately in contrast to other macroeconomic indicators which are subject to longer periods of revision. I investigate the benefits of using univariate MIDAS models in predicting the future price of crude oil. Using a range of high-frequency predictors, I find that the oil futures market and global metal prices show promising results in predicting crude oil prices changes over multiple forecast horizons. The results for the one-step ahead forecast, however, show large prediction errors which cannot be overlooked. I conclude that high-frequency financial and energy predictors do not appear to have a very significant bearing on improving forecast performance in the short run. Although the high-frequency data may contain information rich signals, it is not powerful enough to compensate for the additional noise, which is said to be the cost of using high-frequency data in forecasting. Last but not least, MIDAS models are especially unsuitable and should be refrained from, when a market crash is in sight. This is because bad market news tends to have a severe negative impact on predictive accuracy and interferes with the performance of the forecast models.

Keywords

Brent crude oil price, Financial and energy markets, High-frequency data, Mixed-data sampling (MIDAS), Non-linear least squares (NLS), Forecasts.

Ölpreise: Analyse von Hochfrequenzdaten

Zusammenfassung

Traditionelle makroökonomische Nachfrage- und Angebotsargumente können die unaufhörlichen Schwankungen des heutigen Ölpreises nicht erklären. Daraus ergibt sich die Frage: Was treibt die Ölpreise an und können Prognosen helfen, die zukünftige Ausrichtung der Ölpreise bei Volatilität zu identifizieren? Mixed-Data-Sampling (MIDAS) Regression ist ein Thema von wachsendem Interesse, da die erfragte Variable oft mit einer niedrigeren Frequenz erwünscht ist, während die relevanten Daten in einer höheren Frequenz verfügbar sind. In diesem Zusammenhang ist die Rolle der hochfrequenz Finanz- und Energiemarktdaten bei der Vorhersage des Rohölpreises von zunehmender Bedeutung. Dies fußt darauf dass diese Märkte dazu neigen, die Auswirkungen der Volatilität nahezu sofort, im Gegensatz zu anderen längerfristig anhaltenden makroökonomischen Indikatoren, zu erfassen. Ich untersuche die Vorteile der Anwendung von univariat MIDAS-Modellen bei der Vorhersage des zukünftigen Rohölpreises. Unter Verwendung einer Reihe von Hochfrequenz-Prädiktoren, komme ich zu dem Schluss dass der Öl-Futures-Markt und die globalen Metallpreise bei der Vorhersage der Rohölpreisänderungen über mehrere Prognosehorizonte, viel versprechende Ergebnisse aufweisen. Die Ergebnisse für die Prognose am ersten Prognosehorizont weisen jedoch große Fehlberechnungen auf, die nicht übersehen werden sollten. Ich schließe daraus, dass die hochfrequenz Finanz- und Energieprognosen keine signifikanten Auswirkungen auf die kurzfristige Verbesserung der Ölpreisprognosen zu haben scheinen. Obwohl die Hochfrequentzdaten informationsreiche Signale enthalten können, sind diese nicht stark genug, um das zusätzliche Rauschen zu kompensieren, was oft als Nachteil bei der Verwendung von Hochfrequenzdaten gesehen wird. Darüber hinaus, sind MIDAS-Modelle bei einem vorraussichtlichen Markteinbruch besonders ungeeignet und sollten nicht angewandt werden. Dies liegt daran, dass schlechte Marktnachrichten tendenziell eine schwere negative Auswirkung auf die prädiktive Genauigkeit haben und die Leistung der Prognosemodelle beeinträchtigen.

Schlüsselwörter

Brent-Rohölpreis, Finanz- und Energiemärkte, Hochfrequenzdaten, Mixed-Data-Sampling (MIDAS), Non-linear least squares (NLS), Prognosen.

Acknowledgements

I would like to express deepest gratitude to my supervisor, Prof. Dr. Franz Wirl for his invaluable help, expert advice, understanding and support along every step of writing my master's thesis. Without his guidance, patience and encouragement, this task would have been an overwhelming pursuit.

I would also like to thank all members of the Chair for Energy, Industry and Environment from the Department of Business Administration at the University of Vienna for providing feedback during the master seminar. Their thoughtful questions and comments were greatly valued.

Last but not least, I am very grateful to my sister, Nina Kaushik for painstakingly reading my thesis and providing feedback so that I could improve it.

Author: Anna Kaushik BSc

Matriculation ID: 1252460

Table of Contents

Abbreviations

AR-MIDAS	Autoregressive Mixed-Data Sampling
AIC	Akaike Information Criterion
BIC	Bayesian Information Criterion
CPI	Consumer Price Index
EC	European Commission
EIA	Energy Information Administration
EMMI	European Money Markets Institute
EURIBOR	Euro Interbank Offered Rate
EW	Equal weights
FRED	Federal Reserve Economic Data
GARCH	Generalized Autoregressive Conditional Heteroscedasticity
GDP	Gross Domestic Product
ICE	Intercontinental Exchange
MF-VAR	Mixed-Frequency Vector Autoregression
MIDAS	Mixed-Data Sampling
MSPE	Mean Square Prediction Error
Nealmon	Normalized Exponential Almon
NLS	Non-linear Least Squares
NYMEX	New York Mercantile Exchange
NYSE	New York Stock Exchange
OECD	Organization for Economic Co-operation and Development
O(A)PEC	Organization of (Arab) Petroleum Exporting Countries
U-MIDAS	Unrestricted Mixed-Data Sampling
PDL	Polynomial Distributed Lag
RMSE	Root Mean Square Error
SPR	Strategic Petroleum Reserve
WTI	West Texas Intermediate

1. Introduction

"Oil is so much more than a fuel. It's a force even bigger than its trillion-dollar market. It's a weapon, a strategic asset, a curse. It's a maker and spoiler of fortunes, a leading indicator and an echo chamber. Each has a part in determining oil prices."

~ Brian Wingfield, Bloomberg

The price of oil changes throughout the day. Financial investors and speculators across the globe wake up in the middle of the night to see how the Tokyo Stock Exchange is opening and how the New York Mercantile Exchange is closing. If anything, the last 30 years have proven that oil prices are subject to such high levels of volatility, which sometimes defy all laws of economic gravity. Traditional macroeconomic demand and supply reasoning comes nowhere close to explaining the incessant fluctuations in today's oil prices. This gives rise to two inevitable questions. Firstly, what drives oil price volatility? And secondly, what can be done to minimize the volatility? While the first question can be answered, there is no straightforward answer to the second question. While it is true that the future of oil prices will depend on resource availability, geopolitical and economic stability of nations and on the evolution of global financial markets. However, it is also true that a large part of the volatility has less to do with changes in market forces and more to do with decisions made by policy makers. Meanwhile, forecasters, econometricians and financial analysts can only combine their expert attempts and hope to predict the expected volatility of oil markets correctly or at least identify the direction of future oil price development by putting the information and resources available to them, to the best use.

There is a lot of debate on the best approach to forecast oil prices both in the short and long run. Time series regression models are typically sampled at the same frequency. Mixed-data sampling (MIDAS) regression first proposed by Ghysels et al. (2004) is a topic of growing interest since the variable in question is often desired at a lower frequency while the relevant information is available at a higher frequency, revised on a tick by tick basis. MIDAS models are also very useful in cases of unavailability of same frequency data.

In fact, this approach has also been applied to cases where time series data was available at the same frequency, primarily for the purpose of comparing the performance of direct and iterated forecast models in case of multi-period forecasting[1].

While MIDAS regression analysis is trending, there exist several other approaches for handling data with different frequencies. Until the 21st century forecasters mainly relied on state space models which used the Kalman filter. The Kalman filter allows the use of both high- and low-frequency data to predict low-frequency macroeconomic variables[2]. The use of Mixed-Frequency Vector Auto Regressive (MF-VAR) models in handling mixed-frequency models is also quite common. In fact, Kuzin, Marcellino, and Schumacher (2011) find that it is difficult to rank MF-VAR and MIDAS models based on their efficiency. They consider them to be complements rather than substitutes, since the former performs better at longer horizons and the latter performs better at shorter forecast horizons. In contrast to other mixed-frequency models, an obvious advantage of using the MIDAS framework is that it is a parsimonious regression model and spares the imposition of endless assumptions and parameter estimation for the measurement of equations (Andreou et al., 2010). Bai et al. (2009) and Kuzin et al. (2011) show that MIDAS regressions can be represented with fewer equations, as a reduced form of a state space model. To no surprise, over the last decade the MIDAS approach gained substantial popularity amongst researchers. Nowadays the approach is being used to predict financial data volatility by using intraday data[3].

---

[1] See Marcellino, Stock, and Watson (2006) for more recent literature on direct and iterated forecasting techniques. Comparisons between single step and iterated models for multi-period forecasts have also been made by Findley (1983), Findley (1985), Lin and Granger (1994), Clements and Hendry (1996), Chevillon and Hendry (1996) and Bhansali (1999).

[2] Using mixed-frequency data for macroeconomic forecasting is not new, traditional literature tried to tackle the problem of information loss due to temporal aggregation approaches. The issue was commonly addressed using state space models where the low frequency is data is treated as “missing”. The Kalman filter is then applied to retrieve the missing data. For literature on the Kalman filter in the context of forecasting see for example, Harvey and Pierse (1984), Harvey (1989), Zadrozny (1990) Bernanke, Gertler, and Watson (1997) and Mariano and Murasawa (2003) among others. Bai, Ghysels, and Wright (2013) also provide a direct comparison of MIDAS regressions and state space models using the Kalman filter.

[3] See for example Andersen, Bollerslev, Christoffersen, and Diebold (2006).

It is also commonly deployed when forecasting macroeconomic variables such as GDP growth with intra annual data[4].

The role of financial market data in predicting the price of oil is of growing interest since financial markets tend to catch the effects of volatility almost immediately in contrast to other macroeconomic indicators or physical markets. This research explores the benefits of using high-frequency financial market data in predicting changes in the monthly and quarterly price of oil. According to Baumeister et al. (2014), financial data is highly useful due to its forward looking nature and availability in real time on a daily basis. It has the advantage that this information is not subject to long periods of revision. Yet many industrial data are only published on a monthly or quarterly basis and are subject to longer periods of revision, but may prove to be equally important for forecasting. This thesis finds its primary motivation in a similar research "Do High-Frequency Financial Data Help Forecast Oil Prices? The MIDAS Touch at Work", published by economists Baumeister, Guérin and Kilian in 2014. My research takes advantage of information-rich high-frequency data from global financial and energy markets and based on the predictive power of the dataset, it assesses the usefulness of the variables in determining changes in the real price of oil. The usefulness is measured via improvements in predictive accuracy of the oil price forecasts.

The underlying difference between the publication by the authors stated above and my thesis lies in terms of the variables used and the preferred MIDAS model estimation technique. While the authors compare the performance of the MF-VAR model, a non-linear least squares (NLS) MIDAS model and the unrestricted MIDAS (U-MIDAS) model. I focus exclusively on the original MIDAS model framework with NLS estimation by Andreou et al. (2010). Although the U-MIDAS ordinary least square estimation proposed by Foroni et al. (2014) has been proven to produce rather accurate forecasts under Monte Carlo simulations, I chose not to focus on it for two reasons. Firstly, I wanted to make new contributions to the existing literature on mixed-data sampling in the context of oil price forecasts. Secondly, not only are MIDAS regressions more parsimonious but they have also shown to be less sensitive to specification errors when using NLS lag polynomials.

[4] Daily financial data is also often used to predict monthly macroeconomic variables. Armesto, Engemann and Owyang (2010) offer a user friendly familiarization to MIDAS regressions in this context.

Foroni et al. (2012) compared the performance of MIDAS with U-MIDAS and recommended using U-MIDAS models since they outperform original MIDAS models, but only when forecasting quarterly variables using monthly data. Imposing non-linear polynomial lag restrictions delivered better results in the presence of larger differences in sampling frequencies. The finding that U-MIDAS models tend to be less accurate than polynomial lag MIDAS model specifications when forecasting the monthly real price of oil has been reinforced previously (see Baumeister et al., 2014). In this research, the variables of primary interest are used to produce monthly oil price forecasts. It is only in the case of data unavailability for certain variables, that I create quarterly forecasts.

It should be noted that depending on how a MIDAS model is implemented, the length of the sample period and the variables under consideration matter to a large extent. The literature will point to different and often contradictory findings for most empirical studies deploying MIDAS regressions. Therefore, it is not my intent to draw conclusions based on the best forecasting models. My intention is to answer the following question: *can financial and energy market data predict changes in the price of crude oil in the short run? And do the information rich signals contained in high-frequency data compensate for the additional noise, which is often said to be the cost of using high-frequency data in forecasting?* The following section delves deeper into the key drivers of oil prices and the relevance of the oil futures market in determining oil prices today. Section 3 discusses the development of oil prices in the context of financial and energy markets by introducing the set of high-frequency predictors used in the research. The high-frequency predictors include the spread between oil futures prices and crude oil spot prices; the spread between spot prices of diesel and crude oil; oil company stock returns; crude oil inventories; the global metal and agricultural raw materials price index and finally, interest rates. I investigate the link and movements between the individual predictors and the price of oil, based on existing literature. Section 4 reviews the data sources and explains how the weekly variables were constructed. Section 5 provides a description of the implemented MIDAS polynomial lag models. Section 6 describes the methodology adopted to proceed with the quantitative research. Section 7 discusses the performance of the high-frequency predictors, whereas section 8 analyzes the main findings of the paper. The conclusions drawn for the research can be found in section 9.

2. Historical Background

This section aims to bring the reader a step closer towards understanding the price formation of crude oil and the accompanying volatility in both, the short and long run. In the 1920s several American, British and French oil companies struck a deal under the Red Line Agreement which ultimately led to the foundation of the Seven Sisters by the mid-20th century (United States Department of States, 2016). The biggest players included multinationals, today known as Exxon Mobil, Esso, Chevron, Royal Dutch/ Shell, British Petroleum and Texaco which continue to influence the oil industry even today. These oil giants controlled about 85 percent of the world’s petroleum reserves. It was not until the formation of the OPEC cartel in 1960 that the oil market witnessed a shift in power that led to the emergence of state owned oil companies across developing nations outside the OECD region. It included Russia, China, Malaysia, Iran, Venezuela, Saudi Arabia and Brazil, forming the new seven sisters. By the 21st century, the new seven sisters controlled about one third of the world’s oil and gas production and reserves, while the former global players were reduced to producing 10 per cent and holding only 3 per cent of world’s oil and gas reserves (Hoyos, 2007). Today the OPEC, a consortium of 13 countries controls about 40 per cent of the world’s oil. Thus, the supply market for oil can be viewed as an oligopoly, steered by a few sellers. The oil market is in fact, a hermetic market characterized by imperfect competition and heavily controlled by the OPEC. The market has been affected by major shifts in political and economic power.

Crude Oil Price Reactions to Geopolitical and Economic Events

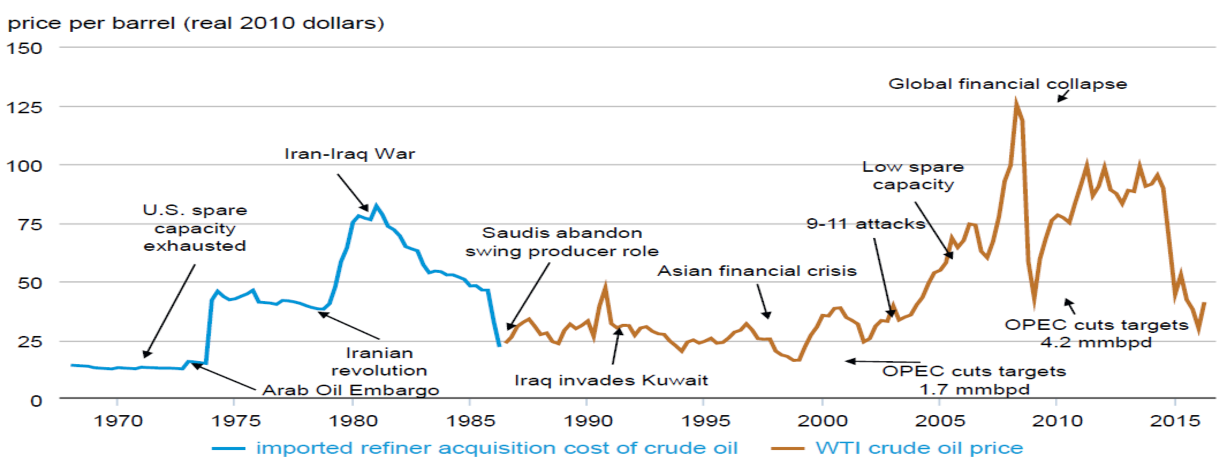

Figure 1. Source: U.S. Energy Information Administration, Thomson Reuters

Oil markets have also been hit hard by severe geopolitical and financial crises in the last four decades. The figure above captures some of the major events resulting in oil price shocks[5]. These include two major oil crises in the 1970s resulting from the OAPEC embargo followed by the Iranian Revolution and Iran-Iraq war. Later, the 1990s witnessed the Iraqi Invasion of Kuwait leading to another oil price spike with the Persian Gulf War. And finally the 2000s financial crises and global recession pushed up oil prices to reach the absolute peak of 136.55$ per barrel (measured in December 2015 $) on June 2008, which has been the greatest spike in the history of oil price shocks (McMohan, 2016).

A brief walk through history helps uncover some of the underlying reasons for the oil price volatility in the long run. But it does not suffice to fully explain short term volatility or what drives the price of oil, when looking at the sudden oil price crash which left oil markets in a slump towards the end of 2014, for instance. To keep this paper in reasonable focus, I emphasize on the big picture and avoid digging too deep into the economic fundamentals and detailed oil market analysis. The purpose is to present a helicopter view of the key drivers of oil prices before proceeding with the analysis of the high-frequency data derived from financial and energy markets. Scanning through the literature on key oil price drivers almost immediately reveals major disagreements amongst politicians and analysts on what will shape the future of oil prices. The discovery of new oil fields, changes in consumer behavior and industrial activity are important factors contributing to the oil price dynamics. The central arguments affecting demand and supply include resource scarcity, cartel behavior, political stability and economic growth of emerging economies resulting from industrialization and globalization to name a few (EIA, 2016).

Narrowed down, the price of oil is said to depend on three factors. Firstly, the economic forces of demand and supply are influential price drivers of crude oil. The demand and supply dynamics for oil can change almost immediately and the oil price is therefore very sensitive to an entire range of events including wars and terrorism, alternative energy sources or substitutes, environmental catastrophes (deep water oil spill), changes in consumption patterns, the rate of global economic growth and investment activity, supply disruptions or news on new supply.

---

[5] U.S.Energy Information Admistration (2016). What drives Crude Oil Prices? See EIA Statistics: http://www.eia.gov/finance/markets/reports_presentations/eia_what_drives_crude_oil_prices.pdf

Depending on the nature of the event the implications may be long or short lived. Furthermore, if the changes in supply and demand fail to offset one another by a comparable magnitude, the consequences for the price volatility can be severe. In reality however, oil is not a product but a commodity and its prices cannot be determined by competitive market forces alone.

Secondly, OPEC- viewed as a major player and accused of abusing its power as a cartel- significantly influences the price of oil. This is because a large portion of the oil supply is concentrated in the Middle East and OPEC volume adjustments are common (Chevillon and Rifflart, 2009). It is advocated that in the 1970s oil prices were solely determined by politics and OPEC which controlled the supply of oil. It was not until the late 1990s, where for the first time oil markets were affected by a major demand side event, that is, the Asian Financial Crises of 1997 (Simmons, 2005). In the time period 2003-2008, OPEC began increasing oil production to sober down the rising prices but the counter attempt was unsuccessful. It is often argued that political unrest within the OPEC member countries resulting in a loss in production capacity coupled with increased worldwide demand, might be an indication that the link between OPEC's control over supply and the oil price trends is weakening. However, in the course of writing this thesis, I read that OPEC had decided to cut production for the first time in 8 years which immediately caused Brent crude oil prices to rise by almost 6% to nearly $49 a barrel on the news[6]. Similarly, in 2014 OPEC maintained its production levels despite lower demand in China and Europe which caused an excess of oil supply resulting in the fall of oil prices to less than 50$ a barrel. This confirms the tremendous power the "*clumsy cartel*"[7] continues to have over oil markets even today and that the cartel's influence should not be underestimated.

This leaves the third and the most debated factor open for discussion, namely, the role of the oil futures in determining the price of oil. Futures contracts are a binding agreement for commodities, giving traders the right to buy oil by the barrel at a predetermined price on a predefined date in the future, whereby the buyer and seller are both obliged to fulfill the agreement.

---

[6] BBC Business News, September 29th, 2016. Oil rallies after OPEC Ministers announce Output Cut: http://www.bbc.com/news/business-37502538

[7] One author referred to OPEC as a "clumsy" cartel (Adelman 1980).

Since the buyer and seller are both in agreement upon the price, futures are expected to reflect direct information about expectations regarding the price of oil in the future. But the matter is complicated since traders in the futures market not only include hedgers (commercial traders in the petroleum business) but also speculators (non-commercial traders). In the last decade high oil prices have been blamed on speculative behavior. Market reactions and investor sentiments may change quickly and contribute heavily to oil price volatility based on the principle of bounded rationality. To make it worse, the power of OPEC extends to the futures market since the cartel often announces production quotas leading to instant movements in crude oil futures' prices. This is due to the changing market sentiment stemming from anticipated changes in production (Hamilton 2009a).

Speculation as a culprit causing inefficiencies in the futures market, is strongly discarded by many-including the U.S. Energy Information Administration and the ICE futures market who argue that physical and financial arbitrage constraints limit speculators' possibilities to drive oil prices above the market equilibrium. Speculation is ruled out under the belief that any mispricing will be reversed due to arbitrage (Happonen, 1999).

However, arbitrage is said to be costly and risky, causing rational investors to refrain from corrective movements and the mispricing to typically continue (Shleifer and Vishny, 1997). Speculations in the oil futures market are also categorically denied by those who blame low inventory levels for high oil prices. Many claim there exists no evidence of supply being withheld from the markets. But there is copious contrary evidence of speculative bubbles in the oil futures market available. Since 2003, the relationship between oil prices and inventory levels has weakened and the total long positions held by non-commercial traders (a measure of speculative activity) on the NYMEX has increased considerably and resulted in excessive trading (Merino and Ortiz, 2005). It is not only demand and supply shocks, but also excessive trading that causes markets to overreact which in turn increases price volatility (De Bondt and Thaler, 1990). Last but not least, all futures commodity trading commissions and institutions claim that oil prices are purely driven by demand and supply dynamics which sometimes results in overconfidence in the oil commodity futures market.

I hope to have conveyed the intrinsic message, namely, no single factor is expected to exclusively explain the movements observed in oil markets.

3. Literature Review: Oil Price Development in the Context of Energy and Financial Markets

Crude oil is a physical commodity and in the long run its price can be determined by demand and supply fundamentals and inventory (Barsky and Kilian, 2004; Hamilton, 2009b). Although in the short run, crude oil prices behave similar to financial assets, yet not identically. The oil futures and derivatives markets have become increasingly liquid with more hedgers and speculators. Arbitrage and risk management theories in financial markets influence market expectations and play an important role in determining prices of oil futures (Huntington, 1994). The previous section highlighted that the futures market dominates oil market trading activity. Yet there are grounds to suspect that the price of oil is also linked to the price of other commodities since the correlation among commodity prices has risen. Financial activity has not only increased in oil markets but also in industrial or non-oil commodity markets. Using information directly measuring economic activity such as prices of non-oil commodities, oil inventories or interest rates has shown that the predictive accuracy of forecasts for crude oil prices can be improved (Alquist et al., 2013). Literature also offers evidence of a close relationship between financial and energy markets and the physical market for oil. Therefore, high-frequency data capturing the important details of current price movements may prove to be highly beneficial in predicting short term oil price movements.

3.1. Crude Oil and Futures Spread

The perception that oil futures prices contain information about the future spot price of oil is not new. Baumeister, Kilian and Zhou (2013) investigate the power of several models using product spreads to forecast the real price of oil, including an oil futures and oil spot price spread model. They reason that in the absence of a risk premium, arbitrage ensures that the expected spot price of a product equals the current price of the same product in the futures market. In reality, futures are risky assets and thereby include risk premiums for the possibility that spot prices deviate from the futures price. Wu and McCallem (2005) argue that although these risk premiums are practically nil in the case of oil futures, over time the prices can reflect large volatility. There also exist valid concerns that futures trading can trigger artificial movements in oil spot prices (Huntington, 1994). This suggests that using oil futures may not be optimal to forecast future oil prices. Alquist and Kilian (2010) consider two types of models to forecast oil prices based on the price of oil futures.

The first model assumes that the oil futures of today reflect tomorrow's oil price. This model is believed to be unbiased in the sense that simply using oil futures prices has an equal probability to over- or under-predict the future spot price of oil. The second model uses the spread between the futures and spot prices of oil to predict the direction of future oil price development. The figure below displays the weekly development of crude oil spot and oil futures prices from the period January 1995 to June 2016. One can see that the spread reflects large variations, especially during the Global Financial Crises of 2008.

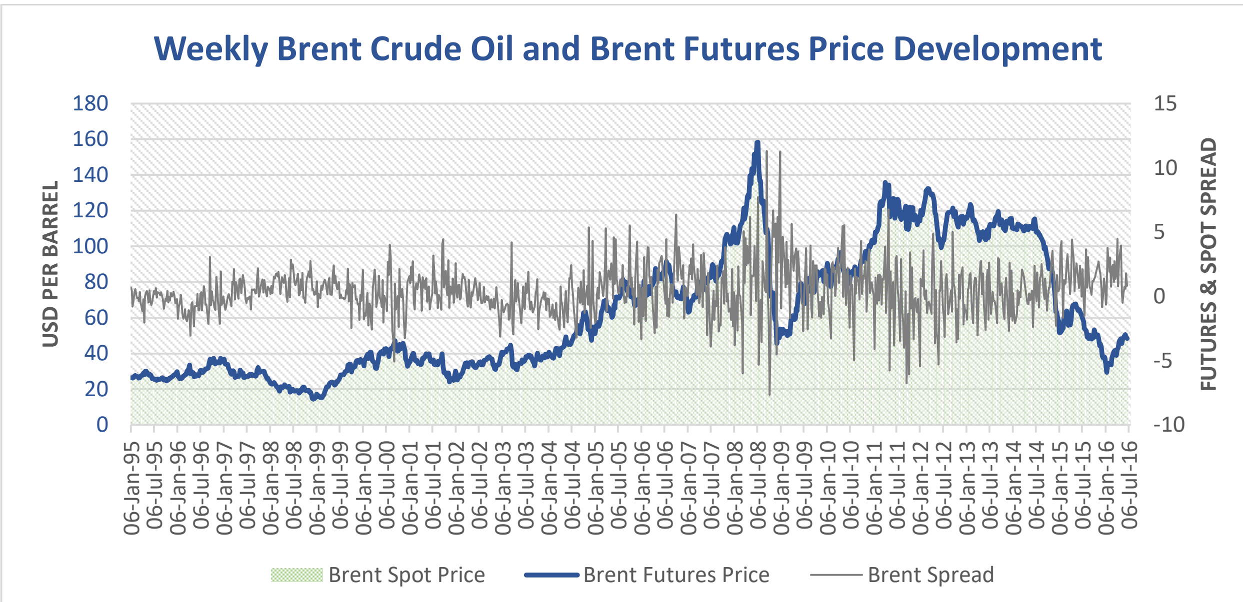

Figure 2. Source: FRED and Quandl database, 2016

While none of the papers used mixed-frequency sampling, Alquist and Kilian (2010) show that whether only using futures or the futures and spot spread, both models were more biased and less accurate than the no-change or random walk forecast (using the no-change benchmark model is standard in literature when evaluating forecast performance). Manescu and Robays (2014) came to the same conclusion.

Limited availability of information on futures prices at longer maturities seems to be a major drawback. Therefore, forecast accuracy evaluation using futures spread is only possible for shorter horizons. A short horizon is said to range anywhere between 1 and 24 months. The futures market at longer horizons lacks liquidity implying little traded volumes of oil futures for contracts exceeding 1 year maturities (Baumeister et al., 2013). Although the total trading volume has substantially increased over the last 25 years, the oil futures market seems to be most liquid only at 1, 3 and 6 month horizons. Yet, there is continued interest in using product spreads from the futures market for predicting commodity prices, at least for very short time horizons. Wu and McCallem (2005) show that using the futures spread model produces more accurate forecasts compared to only using oil futures, but only at short horizons up to 4 months and that the prediction errors remain substantial. New findings suggest that futures prices at varying maturity dates are starting to move more closely with each other and with the spot prices (Behmiri and Manso, 2013). Finally, most relevant to this research, Baumeister et al. (2014) show that using MIDAS models indeed compares more favorably than the no-change forecast for short horizons. So far the empirical evidence does not suffice to reject the usefulness of oil futures spreads in predicting oil prices, specifically when applying a mixed-data sampling approach.

3.2. Crude Oil and Diesel Spread

Refined crude oil yields petroleum products such as gasoline, heating oil and diesel. There is good reason to suspect that these prices may influence crude oil prices since diesel and crude oil prices show a very high degree of correlation. Typically research focuses on whether crude oil prices are beneficial in forecasting diesel prices and not the other way around. Investigating the power of product spreads in predicting the real price of oil has been a topic of interest for a while now. The spread reflects the extent to which the petroleum product prices deviate from crude oil prices and is widely viewed as a predictor of crude oil spot price changes. The views presented in this paragraph and the next are based on findings by Baumeister et al. (2013), else otherwise stated. Using refined crude oil product spreads to predict crude oil prices can be problematic since crude oil prices are more likely to be determined by those refined products which have the highest demand. Because the refined crude oil products are produced in fixed proportions, changes in demand for the product in one market may fail to correctly predict the price of crude oil in another. Also, depending on the grade of crude oil inputs- resulting in different levels of refined petroleum product outputs, it is hard to tell which geographical areas may suffer a shortage or abundance of the product in question and its implications for the future price of crude oil. For instance, the United States was the dominant producer and consumer of gasoline, and the demand for it is likely to affect WTI oil prices. Whereas in European markets, the demand for diesel and heating oil is higher than gasoline, which in turn might not have an impact on WTI crude oil prices at all but have very relevant implications for Brent crude oil prices. The graph below plots the development of diesel and crude oil prices (per barrel). Due to data limitations the diesel spot prices in European markets are only available as of January 2005. Since then one can see significant differences in the diesel price development when comparing American and European spot prices[8]. It is mostly due to shifts in the demand for refined products that we notice a discrepancy between diesel prices in the U.S. and in Europe. One must exercise caution in predicting oil prices using product spreads, since these may vary across different parts of the world.

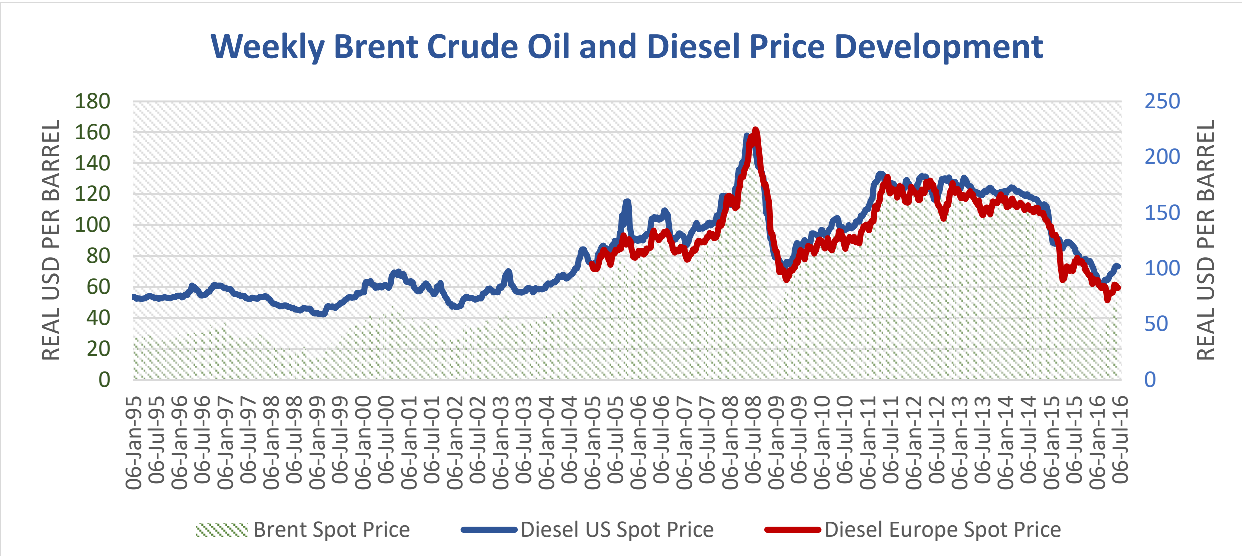

Figure 3. Source: FRED, EIA and European Commission Energy database, 2016

The before mentioned authors show that gasoline and crude oil spread models improve the forecast accuracy up to 2 years, although the prediction accuracy relative to the no-change forecasts is more impressive at shorter forecast horizons up to 6 months. Baumeister et al. (2014) showed similar results using MIDAS models, however the gains in accuracy over the no-change forecast are not particularly impressive. According to Verleger (2011), Europe is becoming the new marginal market for diesel products. For this purpose, Baumeister et al. (2013) extended their research to European markets and used Rotterdam heating oil and Brent crude oil spot price spreads to predict

[8] Prior to 2005, historical diesel spot prices across Europe were only available for a few selected countries. After this period, many countries joined the EU member countries and diesel spot prices across the EU were published regularly.

the price of crude oil and showed no improvements over the no-change forecast model. The performance of diesel spreads using MIDAS models to predict Brent crude oil prices is yet to be assessed.

3.3. Crude Oil and Stock Market Returns

There is a lot of clamour surrounding views on the movements of returns on stocks and oil prices. While some expect falling oil prices to be good news for stock prices due to expected increases in net oil imports for oil importing countries such as China, others are convinced that stock markets and oil prices move together and that low oil prices hurt profits and signal a decline in global aggregate demand for oil. There is also evidence of Brent futures markets and S&P 500 plunging together (Watts, 2016). The positive correlation has been linked to the recent elevations in market volatility which signals, that in times of high uncertainty, investors tend to refrain from both commodities and the stock market (Bernanke, 2016). The correlation between oil and stock prices has witnessed a lot of volatility over the last decade and it may swing towards positive or negative extremes. By large however, the correlation remains positive. In order to understand the correlation, it is a good idea to identify whether the oil price shock is caused by the aggregate demand, the precautionary demand or the supply. While business cycle fluctuations can cause aggregate demand side shocks, uncertainty regarding the future supply of oil due to changing expectations of future demand may trigger a precautionary demand shock- for instance due to wars and terrorism. On the other hand, supply side shocks are generally caused due to falling oil production and these are characterized as exogenous shocks (Kilian, 2009). A positive correlation between oil prices and stock returns in case of aggregate demand shocks and a negative correlation in the case of precautionary demand side shocks has also been confirmed (Degiannakis et al., 2011). Precautionary demand shocks are especially relevant in the short run since these affect investors' reactions to future shortfalls in oil supply. Shifts in aggregate demand are more relevant in the long run as a result of common expectations regarding future economic activity, allowing stock and oil prices to move together (Rafailidis and Katrakilidis, 2014). While most empirical research suggests that supply side shocks fail to explain the link between stock returns and oil prices, Ready (2016) found a strong negative correlation between the stock markets and oil in the presence of supply shocks. The finding however was limited to the consumer goods industry and oil importing countries. Hence, the relationship between stock markets and oil markets is very complex and sensitive. While the majority of literature explores the impact of oil markets on stock markets, the reverse relationship has seldom been explored. The graph below plots the development of the oil stock returns and Brent crude oil prices from the period of January 1995 to June 2016.

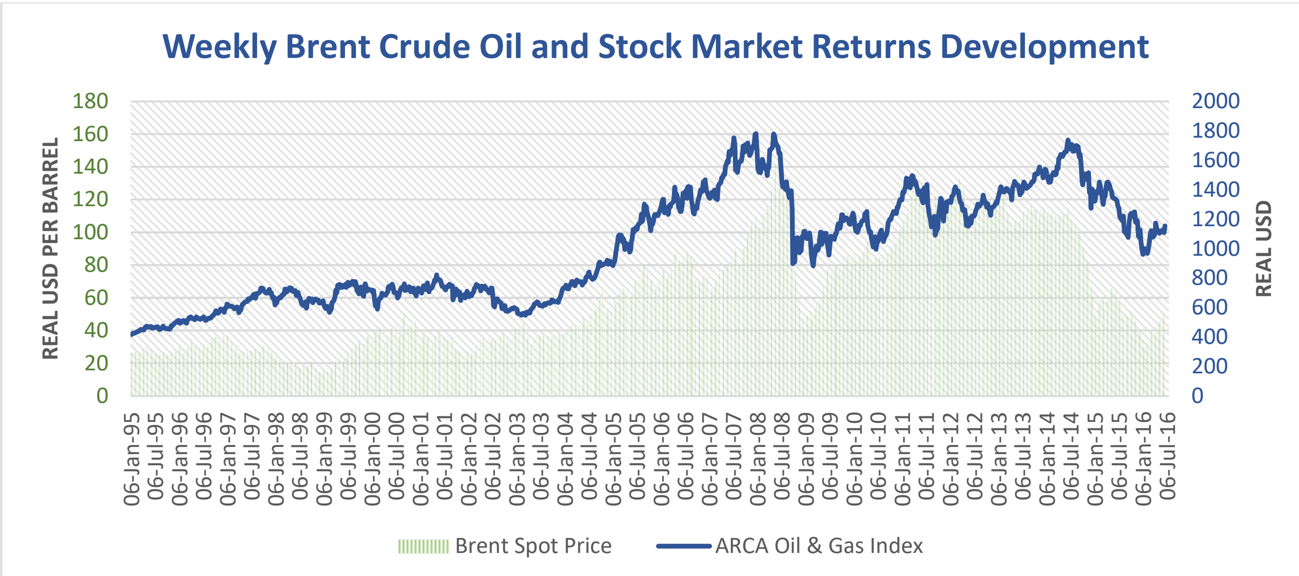

Figure 4. Source: FRED and Yahoo Finance database, 2016

According to Chen (2014) oil-sensitive stock price indices are powerful tools for predicting nominal and real crude oil prices at 1 month horizons relative to the no-change forecast. They showed MSPE reductions of up to 28% for Dubai crude oil and 22% for WTI crude oil and attested directional accuracy using the Arca Oil & Gas Index. Baumeister et al. (2014) did the same using MIDAS models and were able to show improvements in predictive power for horizons up to 15 months. However, the MSPE reductions of 6% relative to the no-change forecast and improvements in directional accuracy remain rather small for all MIDAS model specifications. Stock markets and oil markets are one of the most difficult to forecast. The contradictions in empirical studies are, in fact, quite overwhelming. Therefore, using oil and energy stocks might prove to be more reliable and successful in predicting the future course of crude oil prices. The section on the MIDAS results will shed more light upon the relationship between the two markets.

3.4. Crude Oil and Oil Inventories

The inventory model approach is a long recognized traditional approach used to explain oil price developments. The model was first developed by Pindyck (1994) who explained the effect of demand and supply shocks on commodity prices. It was later extended to petroleum and crude oil commodities (Pindyck, 2001). It has been found that, the price of oil is less responsive to changes in production levels and that oil inventories serve as a better proxy for capturing the effect of market volatility on oil prices in the short run (Ye et al., 2002). The rationale behind maintaining oil inventories can be attributed to demand fluctuations in the commodities markets. Holding inventory allows producers to reduce losses in situations of stockout during periods of peak demand. It also helps save on costs when a need to change production levels arises, for instance due to reduced demand. According to Pindyck (1994) the market clearing price is determined by changes in inventory, production and consumption. And the level of inventory, depends on the inventory holding costs. Pindyck (2001) showed that in the case of temporary oil demand shocks, the inventory used as a buffer caused crude oil pot prices to rise in the short run. After the demand returned back to normal, the spot price of crude oil fell again. But these prices remained higher than the original prices in the market- before the demand shock. This is because oil production remained unchanged and it needed to catch up with consumption levels, so that oil inventories could be replenished again. In the case of more permanent demand shocks, oil prices reacted differently to inventories. This was because both, production and inventories, were increased to meet the growing demand. This results in higher crude oil spot prices but also higher levels of oil inventories which in turn creates a new price equilibrium in the market compared to the pre shock levels. While changes in crude oil inventories are useful in explaining oil price movements in the short run, based on rational demand and supply dynamics, the impact of inventories on the price of oil is said to be limited. The reason is that movements in the oil futures market, stock markets and other commodities markets also steer the course of future oil prices. In simple words, changes in expectations about the real price of crude oil are reflected in changes in crude oil inventories, but only under an all else equal assumption (Alquist and Kilian, 2010). Figure 5 depicts the development of weekly U.S. commercial crude oil inventory together with the safety inventory and the real price of Brent crude oil from the period January 1995 to June 2016.

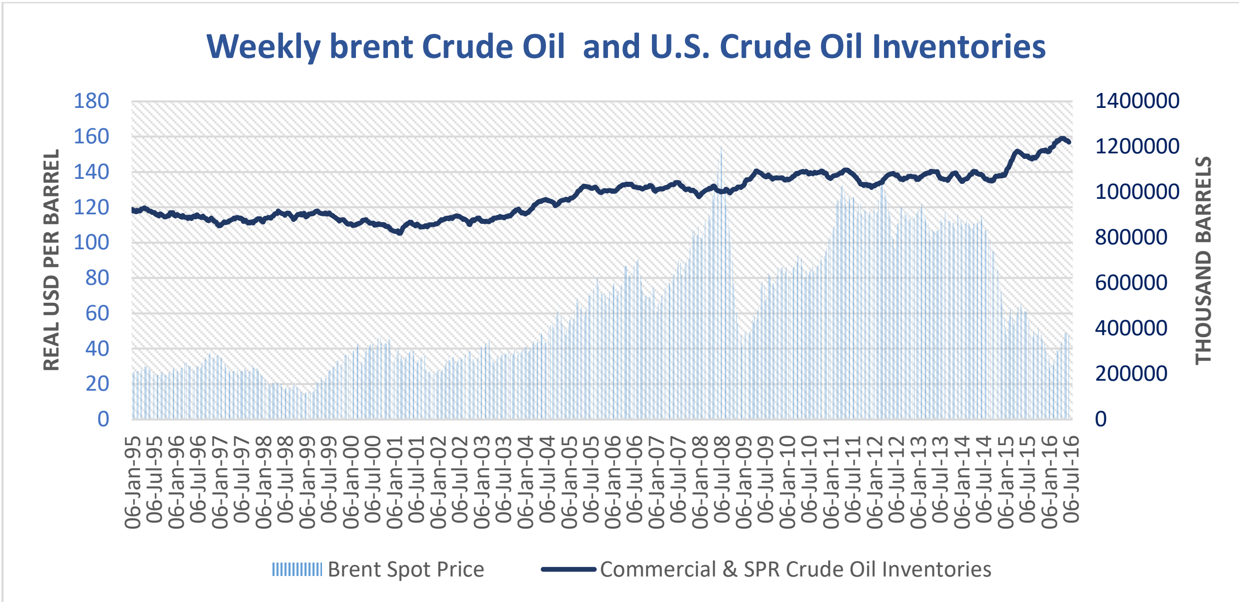

Figure 5. Source: FRED and EIA database, 2016

Baumeister et al. (2014) tested the predictive power of changes in U.S. crude oil inventories on the real price of WTI crude oil using different MIDAS regression models. They found that oil inventories are very successful in predicting changes in the price of crude oil at forecast horizons exceeding 6 months. They achieved MSPE reductions of up to 28% with high directional accuracy. The directional accuracy was found to be as high as 80%, which meant that in 4 out of 5 cases a rise or fall for the price of crude oil was correctly predicted. At shorter horizons they established that MIDAS models, using inventories, performed similar to the no-change forecast model. Their empirical findings are quite remarkable.

3.5. Crude Oil and Industrial Commodity Indices

The price movements of many energy commodities are beginning to converge over the years, mainly due to increased financial activity in both, oil and non-oil or industrial commodity futures markets. A part of the volatility in oil prices is believed to stem from the spot prices of other industrial commodities. It is believed that non-oil commodities could prove to be very useful in predicting future oil price movements. This is because it is easier to predict the price of industrial commodities since their prices are determined by more reliable and predictable variations in global economic activity (Baumeister et al., 2014). While there exist several ways to measure global economic activity, the relationship between oil and gold prices for instance, is considered to be weak. However, a good proxy to measure economic activity are industrial inputs such as the global agricultural raw material index and the global metal index. In fact, there are several reasons to suspect that non-oil commodities and oil markets are correlated. In times of globalization, food production is becoming increasingly fuel dependent, worldwide. According to Chavdarov (2015), agricultural raw material prices affect food prices and there exists a strong positive correlation between oil and food prices. Oil is required for agricultural chemicals, farming equipment and for the transportation of the raw materials and finished goods. Furthermore, there is an increased demand for substitutes such as biofuels which are produced from agricultural raw materials like wheat, sugarcane or corn for instance (Chavdarov, 2015). The demand for these oil substitutes is believed to affect the demand for crude oil commodities. Switching to more renewable energy sources like biofuels or the usage of more efficient transportation modes and environment friendly farming technology is expected to shift the demand for crude oil in the long run, which will have implications for the spot price of crude oil.

In the recent years several other industrial commodities were also affected by the oil price drops witnessed in 2014. The global metal index lists prices of very important metals including copper. The demand for copper is considered to be a signal of economic growth since the metal is heavily used in the manufacturing, IT, construction and health sectors. Low metal prices signal a weakening of the global economy. Furthermore, metals and crude oil prices seem to be affected by very similar economic factors. The difference is that the demand for metals is not as volatile as the demand for oil and there are less speculations regarding the future supply of metals (Burstein, 2015). Thus there is a certain element of stability attached to predicting the prices of industrial commodities. Empirical research shows that using these commodities to predict oil prices yields significant directional accuracy and MSPE reductions. The graph below depicts the development of the global metal index, the global agricultural raw material index and Brent crude oil prices from January 1995 to June 2016.

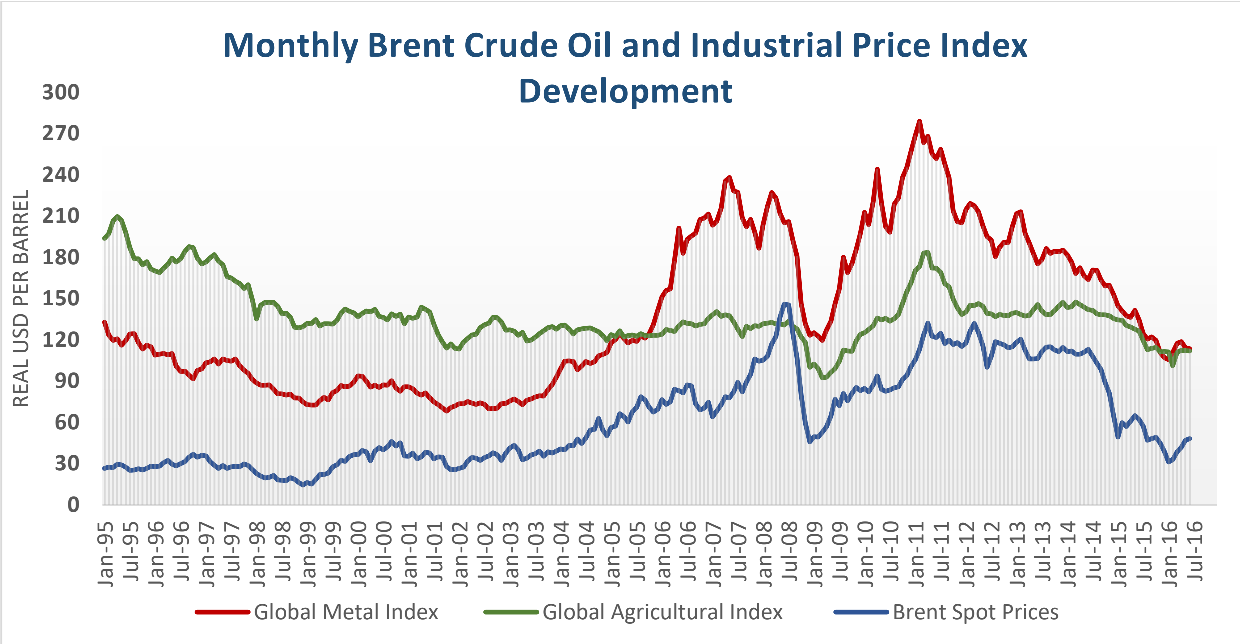

Figure 6. Source: FRED database, 2016

Baumeister et al. (2014) predicted the monthly real price of WTI crude oil using the spot prices of industrial raw materials, provided by the Commodity Research Bureau (CRB) Index. They showed that the MIDAS modeling approach did not produce significant gains in forecast accuracy and that the benefit of using high-frequency weekly data was limited. Never the less I would like to test the relationship, using global industrial indices to forecast Brent crude oil prices changes.

3.6. Crude Oil and Interest rates

The relationship between interest rates and product prices is not new. Hotelling theorized as early as 1931, that under a certain set of assumptions, the opportunity cost of storing oil is nothing but the foregone interest rate. This suggests that interest rates and oil prices move in the same direction since the expected rate of return for holding oil should be identical to the interest rate. In practice however, the relationship between interest rates and crude oil prices is said to be inverse and Hotelling's theory has been widely rejected. Low interest rates reflect looser economic policies and are associated with a higher demand for crude oil since the prices of commodities rises (Baumeister et al., 2014). Similarly, high interest rates make holding oil supply in storage expensive. When the opportunity cost from the foregone interest rate is higher, the incentive to hold crude oil inventories diminishes. The reduced crude oil inventories result in a tightening of oil markets. Moreover, investors tend to shift their investments from the commodities markets to capital assets, when interest rates are high (Piotrowski, 2015). Ultimately this can dry up the liquidity in the oil futures markets and cause the reduced inventory levels to push the price of crude oil upward. But in the short run the implication is a fall in the price of oil. The reverse holds true when interest rates fall. It is only when the crude oil futures are in contango[9] with spot prices that an incentive to hold inventory exists. In simple words, one can expect rising interest rates to increase production and consumption costs, which will result in a lower demand for crude oil due to the reduced economic activity. As long as the supply of oil suffices and interest rates rise, the price of oil is expected to fall, indicating an inverse correlation. Similarly, a fall in interest rates will encourage consumers and producers to borrow and spend more, and thereby push the demand and price of oil upward.

However, the dynamics are not the same for all crude oil importing countries. The exchange rate is also said to impact the price of oil since it affects the purchasing power of oil importing countries and export revenues in fast growing economies with high oil consumption. With oil priced in U.S. dollars, higher interest rates help strengthen the dollar against other currencies. This, for example, allows oil companies in the U.S. to purchase more oil as a result of the dollar appreciation, which in turn makes oil cheaper due to the increased supply. But the effect will be the opposite for China since they have to pay more Chinese Yuan to purchase oil. Similarly, when interest rates are low, the dollar will depreciate and buying oil in the U.S. will become more expensive and push the price of crude oil upward (Piotrowski, 2015). Hence, depending upon how changing interest rates influence fluctuations in the value of the dollar relative to other currencies, it may become cheaper or more expensive for oil importing countries to buy oil (Baumeister et al., 2014). The complexity may further increase since a stronger dollar could also clear out speculative traders in the oil futures and commodities market and thereby affect the price of oil (Piotrowski, 2015). Figure 7 plots the movements in EURIBOR interest rates and Brent crude oil prices between January 1999 and June 2016. The period before 1999 plots the Euro Area Interbank rates.

[9] Contango is a situation where forward or futures prices of commodities are priced higher than spot prices.

The last four years reflect an era of zero interest rates, in fact- negative interest rates. Raising interest rates may cause the price of oil to plunge further, as long as the supply exceeds demand and there are no shortages in crude oil inventories.

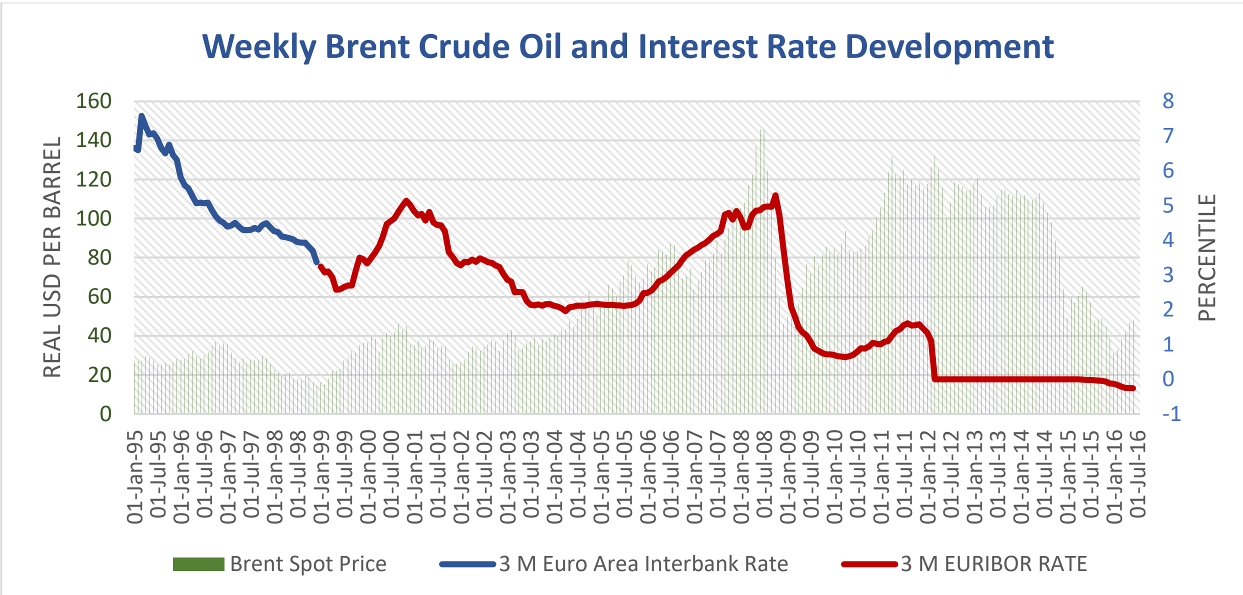

Figure 7. Source: FRED and European Money Markets Institute database, 2016

Baumeister et al. (2014) investigated the predictive power of U.S. interest rates and the U.S. exchange rate on changes in the price of crude oil using MIDAS specifications. In the case of interest rates, MIDAS models outperformed the no-change forecast between horizons of 6 to 18 months. The improvements were very modest and the forecasts showed no directional accuracy. The U.S. exchange rate as a predictor failed to produce any MSPE reductions. While a theoretical link between interest rates, exchange rates and oil prices may exist, lack of predictive power among the variables suggests that the quantitative importance is yet to be established. I aim to test this finding using EURIBOR interest rates.

This section discussed the theoretical relationship between a range of high-frequency predictors from energy and financial markets and the price of oil. Now we are sufficiently equipped to establish whether the high-frequency data contains signals and information, beneficial for forecasting, or if it is simply additional noise, which should be ignored. Using MIDAS models, I explore the possible influence of the predictors in forecasting changes in price of crude oil and compare my findings to the empirical results in the reference paper on the use of high-frequency financial data in forecasting oil prices by Baumeister et al. (2014).

4. The Data

The global crude oil pricing system is anchored on many types of crude oil including West Texas Intermediate (WTI), Dubai Crude and the North Sea Brent as depicted in figure 8. Lately, Brent has become the world's most widely referenced crude oil price benchmark, with 60 percent of the world's traded oil being priced using the Brent reference. Brent is a mix of crude oil from 15 different oil fields in the North Sea and is often mixed with OPEC reference baskets[10]. Baumeister et al. (2014) explored the relevance of high-frequency financial data in forecasting the real price of WTI crude oil. However, in 2010 the crude oil dynamics in North America changed as a result of an oversupply of oil in Cushing, Oklahoma. Since then the WTI has been traded at a significant discount to Brent and many U.S. traders started using a weighted average of Brent and WTI prices as a benchmark.

Worldwide Crude Oil Pricing

Figure 8. Source: ICE Crude and Refined Oil Products Report, 2016

---

[10] For more information on OPEC reference baskets, please refer to OPEC Monthly Oil Market Report, pp.8: http://www.opec.org/opec_web/static_files_project/media/downloads/publications/MOMR%20September%202016.pdf

It is only recently that this gap seems to be closing slowly[11]. It is also worth mentioning that the MIDAS results presented in the paper by Baumeister et al. (2014) may contain some mismeasurements since the results are based on data ranging from the years 1992 to 2012. This also covers the time period where WTI crude oil was no longer considered to be the most reliable benchmark. And this might have influenced the predictive ability of the futures and gasoline spread models. Due to its new status as the international benchmark for crude oil, I used Brent crude oil spot prices in this thesis.

Data Sources

The majority of the data was retrieved from the Federal Reserve Economic Data (FRED) database made available by the Federal Reserve Bank of St. Louis. This included data for the Brent crude oil spot prices, the global price of metals[12], the global price of agricultural raw materials[13], the 3 month Interbank Rate for the Euro area and the U.S. consumer price index (CPI) for all urban consumers. The U.S. ending stocks of commercial crude oil, the U.S. ending stocks of SPR crude oil[14] and the U.S. number 2 Diesel prices were obtained from the Energy Information Administration (EIA) database. The Brent futures front month continuous contracts B1[15] were extracted from the Quandl database. The diesel prices in Europe were obtained from the Weekly Oil Bulletin European Commission database.

---

[11] Intercontinental Exchange (2016). ICE Crude & Refined Oil Products Report: https://www.theice.com/publicdocs/ICE_Crude_Refined_Oil_Products.pdf

[12] The global metal index composes of copper, aluminium, iron ore, tin, nickel, lead and uranium prices.

[13] The global agricultural raw materials index composes of timber, cotton, wool, rubber and hides prices.

[14] U.S. Ending stocks of crude oil consist of commerical and SPR crude oil inventories. The Strategic Petroleum Reserve (SPR) is an emergency fuel storage of petroleum maintained underground in Louisiana and Texas by the United States Department of Energy. It acts as safety inventory and with a capacity to hold up to 727 million barrels it serves as world’s largest emergency supply.

[15] The Brent futures contracts are front month contracts with a maturity of 1 month.

The EURIBOR rates were retrieved from the European Money Markets Institute (EMMI) data base. The NYSE ARCA Oil & Gas index[16] prices were retrieved from Yahoo Finance.

Data Construction

The in-sample period for all data begins in January 1995 and ends in December 2014. The data from January 2015 to June 2016 serves as the out-of-sample period for all the forecasts. All spot prices in the data set were in nominal U.S. dollars and were adjusted for inflation to reflect today's prices using the CPI index of April 2016. I used the U.S. CPI index since all the spot prices use the dollar as the benchmark currency. Another option would have been to use nominal prices but this approach is less common in the literature on forecasting. In order to generate a balanced weekly data set such that each month consists of only 4 weeks, I reduced all months consisting of 5 weeks to 4 weeks. The reduction was undertaken by computing the average of weeks 4 and 5 for each month consisting of 5 weeks and thereby creating a new observation for the 4th week. The unavailability of data across variables over longer periods was highly problematic. Data availability ranged anywhere between the period 1988 to 2016. The U.S. diesel prices could be retrieved starting from the year 1995 while diesel prices in Europe were only available starting from the year 2005. Due to the lack of consistent data sets for diesel spot prices in Europe prior to January 2005, I used U.S. diesel spot for the entire in-sample period. The alternative would have been to work with a shorter in-sample period or to combine U.S. and European diesel spot prices[17]. The diesel spot prices were reported in U.S. dollars per gallon and were converted to the barrel by multiplying the spot price by 42 gallons per barrel.

---

[16] The ARCA Oil & Gas index comprises of several oil companies such as Anadarko Petroleum, British Petroleum, Chevron, Exxon mobil, ConocoPhillips, Hess, Marathon Oil and Total SA among others.

[17] Due to limitations in data on diesel spot prices prior to 2005, U.S. diesel no. 2 prices are used to construct the product spread and maintain consistency. While diesel is an important fuel in Europe, gasoline is more relevant for U.S. markets. Thus is should be taken into consideration that the spot prices may only be sub-optimal for predicting Brent crude oil prices, firstly due to their reduced relevance in European markets and second, due to the use of the Brent crude oil benchmark instead of WTI.

Unfortunately, the data for global crude oil inventories is not published at a weekly or monthly frequency, it is only for the U.S. oil inventories that regular data is available. Very relevant for this research is the question whether U.S. crude oil inventories will have any predictive power in determining changes in the real price of Brent crude oil. Due to the change in the international benchmark for crude oil, it would not be reliable to use WTI crude oil prices at the moment. Neither is it optimal to predict Brent crude oil prices using U.S. oil inventories, currently there seems to be an excess supply of oil in America. An alternative would have been to use data capturing global crude oil production at a monthly frequency. This might not succeed in capturing the gains in forecast accuracy with the inventory model approach. This is because the inventory model approach rests on basic economic fundamentals, namely the interaction between demand and supply, where crude oil inventories are actually a reflection of the imbalance between the demand for oil and oil production levels. The 3 month EURIBOR interest rates were only available starting from 1999. Gaps in the sample period before 1999 are covered using the Euro area 3 month interbank rates[18]. Data for the global metal index and the global agricultural raw material index was only available at a monthly frequency.

While the EURIBOR and Euro area interbank rates were available at a higher frequency, these were converted to a monthly frequency and incorporated into the quarterly forecast model because interest rate data for 90 days shows little variance on a weekly basis. All other data was available at weekly frequency. Hence, depending on the high-frequency variable, I implemented MIDAS models which produced monthly or quarterly crude oil forecasts.

[18] Globalization has brought world financial markets a lot closer over the last two decades. Hence it is assumed that using EURIBOR or LIBOR rates should not affect results with a bias. It is also standard to use interest rates with a maturity of 3 months (90 days) in forecasting. Furthermore, the difference between the Interbank Interest Rate for the Euro Area and the EURIBOR is negligible. Therefore, the former was used to fill the data gaps between the period 1995 to 1999. Thereafter, the EURIBOR returns were used.

5. The Mixed-Data Sampling Approach for Forecasting

Handling variables of mixed frequencies typically involves reducing the higher frequency data to the same frequency as the dependent variable after which standard regression models can be estimated. Aggregation or summing coefficients are traditional approaches for handling mixed-frequency data. The first approach is a simple aggregation approach which aggregates the higher frequency information and transforms it into a lower frequency. It uses equal weighted sums of higher frequency data converted to the lower frequency variable in order to perform a same frequency regression:

$$R_{t+h} = \beta_0 + \lambda R_t + \beta_1 \sum_{k=1}^{K} \frac{1}{\lambda} L^{k/w} x_t^{(w)} + \varepsilon_{t+h} \qquad (1)$$

The approach estimates a single 1/λ value which acts as the common coefficient through which all higher frequency lags enter the lower frequency regression.

The second approach is the individual coefficient approach which takes the distinct coefficient of each high-frequency component and adds it as a regressor into the lower frequency regression. This means that one includes all high-frequency data corresponding to the current low frequency observation:

$$R_{t+h} = \beta_0 + \lambda R_t + \beta_1 \sum_{k=1}^{K} \alpha_k L^{k/w} x_t^{(w)} + \varepsilon_{t+h} \qquad (2)$$

This approach estimates a distinct α coefficient for each of the high-frequency lag regressors.

In the previous sections I explained that financial markets are subject to high levels of volatility. The problem with the first approach is that aggregation may lead to a loss of important information by weakening the impact of the data. Whereas the second approach may leave us with many lags in the data and therefore a very large number of coefficients to estimate in the regression. MIDAS regressions offer a middle ground solution between these two approaches by allowing flexible and parsimonious parameterization to deal with the effect of high-frequency variables on the lower frequency variable. It reduces the parameters to estimate while trying to minimize information loss by incorporating non equal weights. Ghysels et al. (2004) show that MIDAS regressions share some commonalities with finitely distributed polynomial lag models and are built on (non-linear) least square estimations. Aggregation biases due to the unavailability of continuous time data collected at equally distant discrete points in time, is a common issue faced even when using MIDAS models. A classic example is reducing a month with 5 weeks to 4 weeks.

Univariate MIDAS Regression Models

There are various efficiency gains of deploying MIDAS regression due to their superiority relative to simple aggregation techniques used in distributed lag models- in spite of the discretization and aggregation biases, where only independent variables can be sampled more frequently. The MIDAS weighting functions reduce the number of parameters in the model by placing restrictions on the effects of high-frequency variables with varying lag lengths (Ghysels et al., 2004). The MIDAS augmentations shown in this paper use unrestricted autoregressive terms which enter linearly with unconstrained coefficients (Ghysels et al., 2016). The model below represents a univariate MIDAS regression model forecasting h periods ahead:

$$R_{t+h} = \beta_0 + \lambda R_t + \beta_1 B\left(L^{1/w} ; \theta\right) x_t^{(w)} + \varepsilon_{t+h} \tag{3}$$

where R_{t+h} , the dependent variable, is the change in the real price of Brent Crude Oil tomorrow. R_t is the first lag of the dependent variable. In other words, it is the autoregressive term which exhibits the change in the real price of crude oil today. Both are sampled at a low frequency (monthly or quarterly in this paper). While $x_t^{(w)}$ is the independent variable sampled at the higher frequency *w*. The high-frequency data in this paper is weekly or monthly, depending on the variable. In short, x_t denotes the predictor observed in week $w \in \{1,2,3,4\}$ of month t or the predictor observed in month $w \in \{1,2,3\}$ of quarter t. The predictor may also depend on horizon h of the forecast and it is defined as the cumulative change in $x_t^{(w)}$ between a current week and the same week h months ago. It can be represented as $x_t^{(w,h)}$ (this is not shown in equation 3). When the dependent and the independent variables are sampled at the same frequency, *w*= 1. If the change in the real price of Brent crude oil is sampled monthly while the explanatory variable is sampled weekly, the frequency *w*= 4. Similarly, in case where the change in the real price of Brent crude oil is sampled quarterly and the change in the explanatory variable is sampled at a monthly frequency, *w*= 3. β_1 is the common slope coefficient of the high-frequency variables converted to a lower frequency by applying MIDAS weights. In other words, it is the impact parameter.

$B(L^{1/w}\,;\,\theta)$ is the polynomial lag operator for $x_t^{(w)}$, such that the high-frequency lags can enter the low-frequency regression. It is the weighting function for each lagged observation and is represented as:

$$B(L^{1/w}\,;\,\theta) = \sum_{k=1}^{K} B(k;\theta)L^{k/w} \tag{4}$$

where k is the number of lags selected and $\boldsymbol{\theta}$ is the vector of hyper-parameters to be estimated for the normalized weighting function $B(k;\theta)$, which contains the MIDAS weight coefficients that determine the shape of the above weighting function. $L^{1/w}$ is the lag operator or common coefficient such that $L^{1/w}x_t^{(w)} = x_{t-1/w}^{(w)}$. Thus the impact parameter β_1 (in equation 3) is the common slope coefficient obtained by applying the MIDAS weights (θ) to the overall slope ($L^{1/w}$) for k lags. This is how MIDAS regressions allow for high-frequency lags to enter the low frequency regression with a common coefficient. This coefficient depends on the type of polynomial weighting scheme used. The MIDAS application in this paper uses the original distributed lag polynomials as weighting functions provided by Ghysels et al. (2006a, 2006b).

5.1. Exponential Almon MIDAS Regression

The normalized exponential Almon (Nealmon) weighting function uses exponential weights where Q denotes the degree order of the polynomial. The weights of this polynomial sum up to unity and produce positive coefficients due the expressions in the denominator:

$$B(k,\boldsymbol{\theta}) = \frac{e^{(\theta_1 k^1 + \cdots + \theta_Q k^Q)}}{\sum_{k=1}^{m} e^{(\theta_1 k^1 + \cdots + \theta_Q k^Q)}} \tag{5}$$

A lag polynomial of degree order 2 yields the functional form of two parameters with $\boldsymbol{\theta} = [\,\theta_1, \theta_2]$:

$$B(k,\theta_1,\theta_2) = \frac{e^{(\theta_1 k + \theta_2 k^2)}}{\sum_{k=1}^{m} e^{(\theta_1 k + \theta_2 k^2)}} \tag{6}$$

The expanded form of the Exponential Almon Lag MIDAS regression model can be represented by the following equation:

$$R_{t+h} = \beta_0 + \lambda R_t + \beta_1 \sum_{k=1}^{K} \frac{e^{(\theta_1 k + \theta_2 k^2)}}{\sum_{k=1}^{m} e^{(\theta_1 k + \theta_2 k^2)}} L^{k/w} x_t^{(w)} + \varepsilon_{t+h} \tag{7}$$

where k is the selected number of lags and $L^{k/w}$ becomes the slope of the coefficient which is common across the lags. The Exponential Almon function can change shapes with $\theta_1 = \theta_2 = 0,$ implying equal weights for instance. Thus differential response comes via the exponential weighting function and the lag polynomial which depends on the two MIDAS coefficients θ_1 and θ_2. The lag selection in such models is largely data driven, where the rate of weight decline determines the number of lags included in the regression and therefore one must ensure to not choose too few lags. This regression model is highly non-linear in the parameters of the model.

5.2. PDL Almon MIDAS Regression

The Almon lag weighting function by Shirley Almon (1965) also known as the polynomial distributed lag (PDL) function works well while placing restrictions on the lag coefficients of autoregressive models and is thereby also suitable for MIDAS regressions. Here the weight on each lag is computed as:

$$B(k;\, \boldsymbol{\theta}) = \sum_{q=1}^{Q} \theta_q\, k^q \tag{8}$$

where Q represents the order of the polynomial $\boldsymbol{\theta} = \left(\theta_1\,, \theta_2\,, \theta_3 \dots \theta_Q\right)$ and k is the number of lags. Since the weights of this polynomial do not sum up to unity parameter β_1 is not estimated. The number of coefficients or hyper-parameters to be estimated depends on the polynomial order. A second order Almon polynomial, meaning Q = 2 will estimate four parameters $\theta_1, \theta_2, \theta_3$ and β_0. The expanded form of the PDL Almon Lag MIDAS regression model may be represented by the equation:

$$R_{t+h} = \beta_0 + \lambda R_t + \sum_{q=1}^{3} \theta_q\, k^q L^{k/w}\, x_t^{(w)} + \varepsilon_{t+h} \tag{9}$$

where k is the number of lags and $\boldsymbol{\theta}$ is a distinct coefficient associated with each of the Q sets of constructed variables. Since the chosen number of lags is not data driven, the model is sensitive to the choice of lags.

5.3. Beta MIDAS Regression

The normalized Beta weighting function discussed by Ghysels at al. (2006) is also deployed in the MIDAS regression. This beta lag polynomial has two parameters given by:

$$f(k,\theta_1,\theta_2) = \frac{f(\frac{k}{K},\theta_1,\theta_2)}{\sum_{k=1}^{K} f(\frac{k}{K},\theta_1,\theta_2)} \tag{10}$$

This function is often used in Bayesian econometrics to impose parsimonious yet flexible distributions. The beta function can take many shapes with $\theta_1 = \theta_2 = 1,$ implying equal weights for instance. Slowly declining weights imply that $\theta_1 = 1$ and $\theta_2 > 1$. Faster declining rates can be obtained as θ_2 increases. A further restriction is to set θ_3= 0 or θ_3= 1. Thus various parameterizations can obtain humped shaped or strictly decreasing weighting functions. As in the case of the exponential Almon Lag polynomial function, the rate of weight decline determines the number of lags included in the MIDAS regression.

These yield positive coefficients which sum up to unity, where:

$$f\left(\frac{k}{K},\theta_1,\theta_2\right) = \frac{\left(\frac{k}{K}\right)^{\theta_1-1}\left(1-\frac{k}{K}\right)^{\theta_2-1}\tau(\theta_1+\theta_2)}{\tau(\theta_1)\tau(\theta_2)} \tag{11}$$

and

$$\tau(\theta_p) = \int_0^\infty e^{\frac{-k}{K}}\left(\frac{k}{K}\right)^{\theta_p-1} d\,\frac{k}{K} \tag{12}$$

For simplicity let us denote $\left(\frac{k}{K}\right)$ by ω. Hence equation:

$$f\left(\frac{k}{K},\theta_1,\theta_2\right) = f(\omega,\theta_1,\theta_2) = \frac{(\omega)^{\theta_1-1}(1-\omega)^{\theta_2-1}\tau(\theta_1+\theta_2)}{\tau(\theta_1)\tau(\theta_2)} \tag{13}$$

can be reformulated to:

$$f(k,\theta_1,\theta_2) = \frac{(\omega)^{\theta_1-1}(1-\omega)^{\theta_2-1}\tau(\theta_1+\theta_2)}{\sum_{k=1}^{K}(\omega)^{\theta_1-1}(1-\omega)^{\theta_2-1}\tau(\theta_1+\theta_2)} \tag{14}$$

Hence the expanded form of normalized Beta Polynomial MIDAS regression can be represented by the following equation:

$$R_{t+h} = \beta_0 + \lambda R_t + \beta_1 \sum_{k=1}^{K} \frac{(\omega)^{\theta_1-1}(1-\omega)^{\theta_2-1}\tau(\theta_1+\theta_2)}{\sum_{k=1}^{K}(\omega)^{\theta_1-1}(1-\omega)^{\theta_2-1}\tau(\theta_1+\theta_2)} L^{k/w} x_t^{(w)} + \varepsilon_{t+h} \tag{15}$$

where k is the number of lags. The number of parameters of the beta weighting model is at most three so that it does not increase with the number of lags. This paper restricts itself to two parameters. The resultant regression model is also highly non-linear in terms of parameters.

6. Methodology and Model Specification

Forecast Method

I used three types of MIDAS weighting functions to predict the change in the real price of Brent Crude Oil over the period January 2015- June 2016. The model's forecast performance is assessed by splitting the data set into an in-sample period which begins in January 1995 and ends in December 2014. This period is used for initial model selection and parameter estimation. January 2015 to June 2016 serves as the out-of-sample period, to which the performance of the forecasts covering a horizon of 1 to 18 months is later compared. In simple words, the in-sample period over which the model is estimated is used to forecast the out-of-sample period using the static forecast method. When producing forecasts, a model which is best in terms of fit for an in-sample data, does not necessarily provide the more accurate forecasts. But a robust model with adequate functional constraints and lag length is a good starting point. A true of out-of-sample forecast is only possible for one-period ahead (one-step ahead) meaning at horizon 1. Therefore, the results discussed in the next section for multiple forecast horizons exceeding one period should be interpreted as pseudo forecasts. It is common to produce out-of-sample predictions over a desired forecast horizon, which does not require new data for forecasting at each horizon (Ghysels et al., 2016).

Model Estimation

The models are estimated in programming language R version 3.2.4 using the midasr package. The most important factor for reliable MIDAS results is perhaps the model specification itself. Due to the fact that the weighting schemes adopted in this research for high-frequency regressors result in NLS estimation, only one variable was tested at a time to avoid multicollinearity complexities. Andreou et al. (2011) offer an alternative solution to tackle multicollinearity issues, that is by estimating several univariate models and then using a forecast combination to produce final forecast. This however is currently not the subject of interest since I attempt to determine the predictive power of a range of individual financial and energy market variables. MIDAS are typically not autoregressive models. It is assumed that the real price of Brent Crude oil will exhibit a seasonal response, whether or not the high-frequency explanatory variables carry seasonal variation. Ghysels et al. (2006b) warn that the model carries the risk of efficiency losses when lagged dependent variables are introduced into the MIDAS regressions and that this specification should only be used if there are seasonal patterns in the independent variables. Results can be interpreted as whether the change in the real price of oil also linearly depends on the change in its own previous values and on a stochastic term. Since the results are model estimates for the given realizations of explanatory variables, a further risk accompanying this methodology is that the in-sample period performance can be quite sensitive to outliers or sudden shocks in the data. This is why, out-of-sample forecasts are better at reflecting information available in real time. There exist several methods to improve and re-estimate a non-linear least square model with a poor fit. I used the *optim* function in the R package which applies the Nelder Mead algorithm to optimize the NLS function. The function may also be used to find more suitable start values. The obtained start values are then used to re-estimate the NLS model. This however requires running the Nelder Mead algorithm for many iterations and is only successful if the algorithm reaches convergence, which is often difficult to achieve (Ghysels et al., 2016). Several optimization functions can be chosen from to perform the minimization for the NLS estimation. There also exist hybrid optimization methods for example in the E-views software package which might be more successful in reaching convergence (for examples please refer to the Eviews guide on MIDAS). Table 1 provides a summary of these NLS estimates for the hyper-parameters from the in-sample data for MIDAS models. The summary results reflect only initial model selection and parameter estimation.

Model Restrictions

The imposed restrictions are the weighting scheme applied across the model, the maximum lag order of high-frequency variables used in the low frequency regression and the parameters of weighting functions in each lag. The regression model estimates the impact parameter and hyper-parameters mentioned in the above section. The impact parameter is the slope of the coefficient which is common across the high-frequency variable transformed to the lower frequency and the hyper-parameters are the two actual lag or MIDAS coefficients denoted as θ on which the lag pattern primarily depends. In case of the PDL Almon weighting scheme, three lag coefficients are obtained as hyper-parameters, while by default, the slope coefficient is not estimated. The midasr package requires providing start values for the weights of the first low frequency lag in order to estimate these parameters. The start values implicitly define the number of parameters of the constraint functions used for each data series (Ghysels et al., 2016). Furthermore, depending on the start values and lags selected, the NLS problem may or may not converge. This makes a MIDAS model specification highly sensitive. Different start values for the weighting functions were tried and the forecast results are based on the start values where the NLS problem successfully converged.

In order to validate the forecast results, it was necessary to test the adequacy of the restrictions for the MIDAS regressions by performing some standard tests, with the null hypothesis being that the functional restrictions on the empirical model are adequate. Else the forecasts would have little relevance. Table 2 in the appendix summarizes the results of some of the constraint adequacy tests. Three tests were performed. The Hah-test by Kvedaras and Zemlys (2012) was used to test whether the restrictions on the MIDAS regression coefficients hold. This test is a heteroscedasticity and autocorrelation weight specification test. The Hahr-test is simply the robust version of the Hah test. I also performed the Deriv test to check whether the NLS restricted MIDAS regression problem converged. This tests gives values of the gradient and hessian of the optimization function and tells us whether the conditions of local optimum are met. However, if the functional adequacy cannot be rejected at an appropriate significance level, there exist several other options to make a selection of the best candidate in the midasr package. I applied the *select and forecast*[19] feature built for restricted regression models which is designed such that it selects the best forecasting equations in terms of model specification and the in- and out-of-sample precision measures for each forecasted horizon.

Lag Determination

The results discussed are based on a high-frequency maximum lag length fixed at sixteen for monthly forecasts and twelve for the quarterly forecasts. The minimum lag length is four, in the case of the monthly forecasts and three, in the case of quarterly forecasts. A lag length of sixteen, implies that prices for June 2016 are explained by the movements in the weeks of March, April,

[19] The *select and forecast* feature of the midasr package automatically selects the best models at each forecast horizon from a set of potential model specifications containing the lag orders and functional restrictions. It should be noted that the forecast combinations obtained from this function are based on the MIDAS coefficient restrictions and not on the selected lag length. The best lag order can be decided using an information criterion.

May and June. Determining the lag length of the explanatory variables is very important since choosing too few lags results in under-parameterizing of the MIDAS models with the consequence of very high prediction errors (Götz et al., 2014). Provided that the sample is large enough, for cases where more than necessary lags are chosen, the MIDAS lag polynomial assigns a zero weight to the unimportant lags without affecting predictive accuracy (Asimakopoulos et al., 2013). In mixed-data sampling literature, there is a lot of discussion on the optimal lag lengths, which are selected based on some information criterion. The lag lengths can be selected based on several selection criteria. Foroni et al. (2014) for instance, test whether the Akaike information criterion(AIC) provides better results than the Bayesian Information Criterion (BIC) since it puts a lower loss on the number of parameters than BIC when the lag lengths are increased[20]. They show that there are no gains from switching the selection criterion. As the lag length increases, the problem is not the omitted regressors, but rather the estimation of too heavily parameterized models. The in-sample fit is optimized based on the AIC lag selection criterion in this research. The AIC criteria gives a warning when the lags selected are too high. I tested the model by varying the lag lengths, however the results do not vary much as long as the minimum lag length is covered. This is why the forecast results use different lag lengths and start values, since achieving convergence for the NLS estimates at a fixed lag length with the same start values is not always guaranteed.

Forecast Evaluation

In order to present and compare the predictive power of the variables using different MIDAS specifications, I rely on the root mean square error (RMSE) observed over the forecast horizon. The RMSE is computed by evaluating Brent crude oil forecasts against the actual crude oil prices from the out-of-sample window. The forecast accuracy of the predictors is presented in the form

[20] The Akaikes information criteria usually tries to find unknown models. The Bayesian information criteria on the other hand only comes across true models. Thus AIC cannot determine the quality of a model in an absolute sense. In the case where all the candidate models fit poorly relative to another, AIC can only give a warning of that. For simple definitions or more information on AIC and BIC, see for instance, Acquah de Graft (2010):

http://www.academicjournals.org/article/article1379662949_Acquah.pdf of RMSE ratios. The ratio is computed by taking the RMSE of the monthly (quarterly) forecasts relative to the RMSE value of the no-change (random walk without drift) forecast. This is standard in literature. According to oil experts, Alquist and Kilian (2010), no-change forecasts are more accurate than forecasts based on econometric models or survey forecasts. The no-change forecast can be interpreted as a naive forecast. It has been shown that, whether or not crude oil spot prices follow a random walk, naive forecasts tend to be attractive in terms of their RMSE. This is because excluding other predictors in small samples results in reduced variance which more than offsets the omitted variable bias (Alquist and Kilian, 2010). For the benchmark forecast, the future forecast values are simply set to be the values of the last observation. Since I use a random walk model without drift and compare predictive accuracy over a horizon, it is possible to work under the assumption that the price of crude oil remains the same throughout the 18 months or 6 quarter forecast horizon for the no-change forecast. An RMSE ratio below 1 indicates that a MIDAS model is more accurate than the no-change forecast model. This suggests that high-frequency data from financial markets do contain information which is useful in predicting oil prices. Section 7 discusses these results for the horizon of one, three, six, nine, twelve, fifteen and eighteen months in the case where the frequency of the dependent variable is monthly. When the frequency of the dependent variable is quarterly, the RMSE ratios are provided at quarterly intervals over the forecast horizon of one to six quarters. Whether the change in the real price of Brent Crude Oil is being measured at a monthly or quarterly frequency, the out-of-sample period January 2015- July 2016 is fully covered. I discuss the individual forecast results of the specific MIDAS models relative to the no-change forecast. I also present the overall accuracy of MIDAS models over the no-change forecast using an equal weighted average of the Exponential Almon, PDL and Beta MIDAS forecast performance. I also evaluate the directional accuracy of the forecasts with the help of the Pesaran Timmerman test. Under this tests, the null hypothesis states that there is no directional accuracy. A success ratio is used to report the directional accuracy of the forecasts, measuring the number of times the forecast model is able to correctly predict whether the real price of crude oil rises or falls. Success ratios higher than 0.5 imply rejecting the null hypothesis of no directional accuracy.

7. Empirical Results

This section discusses the predictive power for a range of individual high-frequency predictors in determining the real price of crude oil. The theoretical relationship between the predictors and the price of oil as well as the most important findings from the current literature were already discussed in section 3. It should be taken into consideration that the efficiency of the MIDAS models not only depends on the model specification itself but also on the predictive power of the variables since some variables have a stronger correlation to oil prices than others. While it is possible to evaluate the performance of the parameters using MIDAS regression models for the historical data, currently there exists no valid test to confirm the statistical significance of the RMSE reductions for the out-of-sample forecasts generated (Baumeister et al., 2014). Thus there is no reliable method to confirm whether the forecasted model suffers from parameter estimation uncertainty. Before discussing the results of the univariate MIDAS models, it should also be mentioned that the results should be interpreted as changes measured through logarithmic differences, which is similar to using percentage changes. The real price of crude oil is measured by taking the difference between the natural logarithm of crude oil prices from one period to the next. The no-change forecast for the real price of Brent crude oil also uses the natural logarithmic difference to measure the change in the price of oil from one period to the next. The forecast accuracy measured through the RMSE shows how far the forecasted values deviate from the actual realized values in the out-of-sample period. Finally, the MIDAS forecast performance relative to the no-change forecast presented in the form of RMSE ratios is assessed.

7.1. Forecasting Brent Crude Oil Prices with Futures

There is a widespread consensus that the oil futures market is a good predictor of the future price of oil. It was explained in section 3 that in the absence of risk premium, arbitrage should ensure the price of crude oil futures to be the conditional expectation of the spot price of crude oil. The historical Brent futures prices and Brent spot prices were converted into their natural logarithm and the spread was computed by taking the difference between the two. The spread can be written as:

$$\Delta Spread_t = F_t - S_t \tag{16}$$

The change in the spread from one week to the next forms the high-frequency variable $x_t^{(w)}$introduced in section 5 and is measured for weeks w= 1,2,3,4 for month t. The underlying difference between the Brent futures used in this research and the WTI Crude oil futures modeling approach by Baumeister et. al (2014) lies in terms of maturity of the futures contracts. Due to limitations in accessing data, the futures contracts used in this paper have a maturity of 1 month, while the futures contracts used in the reference paper have a maturity varying from 1 to 18 months. I worked under the assumption that the difference between futures contracts with longer maturities is modest. Considering the empirical evidence that futures prices with varying maturity dates are beginning to move closer together along with the spot prices, this is reasonable assumption to make (Behmiri and Manso, 2013).

Table 1 summarizes the regression estimates of the in-sample data for the MIDAS models with the normalized exponential Almon weights, the PDL Almon weights and the Beta weights. With the exception of the intercept and the first MIDAS coefficient of the PDL Almon model, all variables in the in-sample data are statistically significant. This means that changes in tomorrow's crude oil prices can be explained by the change in the current oil price and by the common slope obtained by converting the change in weekly futures spread to the monthly frequency with the help of the MIDAS weight coefficients. The current oil price has a positive impact on the tomorrow's oil price at a 5% significance level for all models, which is a reasonable expectation. The exponential Almon and Beta models show that the change in the futures spread had an overall negative impact on the change in price of crude oil at a 1% and 0% significance level, respectively. This implies that as the oil futures and spot spread increases, the price of oil drops. For the PDL Almon specification, the impact on crude oil prices is solely explained by the slope of the PDL weights. For this model, the second and third MIDAS weight coefficients are statistically significant at 1%. The change in the spread can be explained twofold. Firstly, herd behavior among speculators and panic among hedgers in the futures market may cause the oil future prices to rise by a higher magnitude than the crude oil spot price. Furthermore, a temporary shock triggering a spike in oil prices may cause the increase in the futures prices to be of a smaller magnitude than oil spot prices. The effects of such a shock however are not expected to last long and should dissipate in the market.

Table 2 tests the adequacy of the imposed restrictions. The null hypothesis cannot be rejected for either model, implying that none of the MIDAS weight restrictions suffer from heteroscedasticity and autocorrelation. Hence the restrictions are appropriate. The robust version of the Hah test concludes similar results. The Deriv test confirms that the NLS MIDAS problem has converged for all MIDAS specifications. The results of the MIDAS models differ in lag lengths (determined using AIC). The purpose of discussing the model fitting is to ensure that the forecasts results are reliable, since predictive accuracy of the variables would have little benefit if the NLS problem did not converge.

The figure below depicts the individual and average predictive accuracy (EW MIDAS) of the MIDAS models relative to the no-change forecast using the oil futures spread. A true forecast only applies to horizon 1. As shown in the graph, none of the MIDAS forecasts beat the no-change forecast at horizon 1. The RMSE ratios relative to the no-change forecast are very high at values close to 5 for all three MIDAS specifications. For a model to perform better than no-change forecast one would expect an RMSE ratio below 1. This however changes for forecast horizons 3, 6, 9, 12, 15 and 18. The colored lines lying lowest in the 3D graph reveal the superior performing model with the largest RMSE reductions.

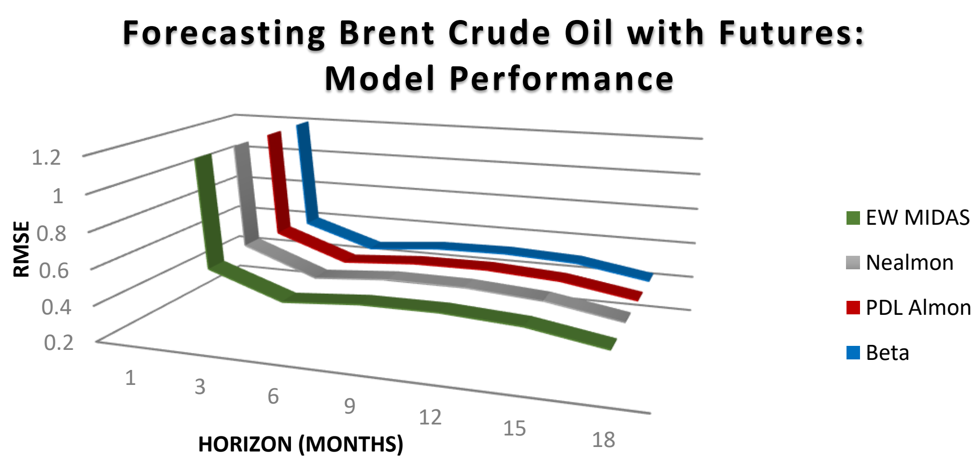

Figure 9

Table 3 in the appendix shows that Beta MIDAS forecasts have the highest RMSE reductions relative to the no-change forecast at every horizon between 3 and 18 months. The exponential Almon and PDL Almon models give similar results. The gains in accuracy up to 40% for the Beta MIDAS specification are substantial at a horizon of 3 months and are maintained at longer horizons by all model specifications. The largest RMSE reduction is 57% and is achieved at a forecast horizon of 18 months by the PDL Almon and Beta MIDAS models. Both the no-change forecast and the MIDAS forecasts predict the direction of change in oil prices correctly at horizon 1. The directional accuracy for the MIDAS models is poor at all other horizons except at the horizon of 18 months where it performs slightly better compared to the no-change forecast. The directional accuracy however is statistically insignificant. The failure of either model to compare favorably with the no-change forecast at horizon one is striking and at a first glance indicates that there are no gains from employing MIDAS models with futures spread for one period ahead forecasts.

7.2. Forecasting Brent Crude Oil Prices with Diesel

The concept of using product spreads to predict changes in the price of crude oil was introduced in the literature review on oil price development in section 3. Product price spreads reveal information about changes in the demand for petroleum products since they reflect how far the price of diesel deviates from the spot price of crude oil. It hereby serves as a topic of interest to examine how this may affect expectations regarding the future development of crude oil spot prices. Similar to the futures spread model, the weekly U.S. Diesel spot prices and Brent crude oil spot prices were converted to their natural logarithm and the spread reflects the price difference between the two products.

Table 1 shows that changes in today's spot price of crude oil have a positive impact on the change in tomorrow's oil price at a 0% significance level for all three MIDAS model specifications. The regression estimates of the in-sample data for the exponential Almon and Beta models show an overall negative relationship between the change in the monthly real price of crude oil and the change in the real price of the diesel spread at a 0% significance level. The MIDAS coefficients of these two models also show the highest statistical significance. For the PDL Almon model, only the second MIDAS weight coefficient is statistically significant at a 5 % level. The regression estimates indicate that an increase in the diesel and crude oil spot spread today will cause the spot price of crude oil to drop tomorrow. There are many possibilities to explain this observation. One argument would be that as the demand for diesel rises, it pushes up the price while the demand for crude oil remains unchanged or increases by a smaller magnitude. The resultant increase in the product price spread may result in lower demand for crude oil, triggering crude oil prices to fall.

Testing the adequacy of the restrictions in table 2 confirms convergence of the NLS optimization function for all three MIDAS models. However, it is only in the case of PDL Almon specification that the MIDAS coefficient restrictions are free of heteroscedasticity and autocorrelation. The null hypothesis for the Exponential Almon and Beta MIDAS specification is rejected quite strongly at a 0% significance level, implying that the imposed restrictions are not optimal. Hence I only discuss the performance of the PDL Almon specification relative to the no-change forecast in detail.

The 3D graph below presents the individual and average MIDAS model performance for all three weighting schemes relative to the no-change forecast. Similar to the futures spread, none of the MIDAS models outperforms the no-change forecast model at horizon 1. However, gains in predictive accuracy are remarkably high after the first month.

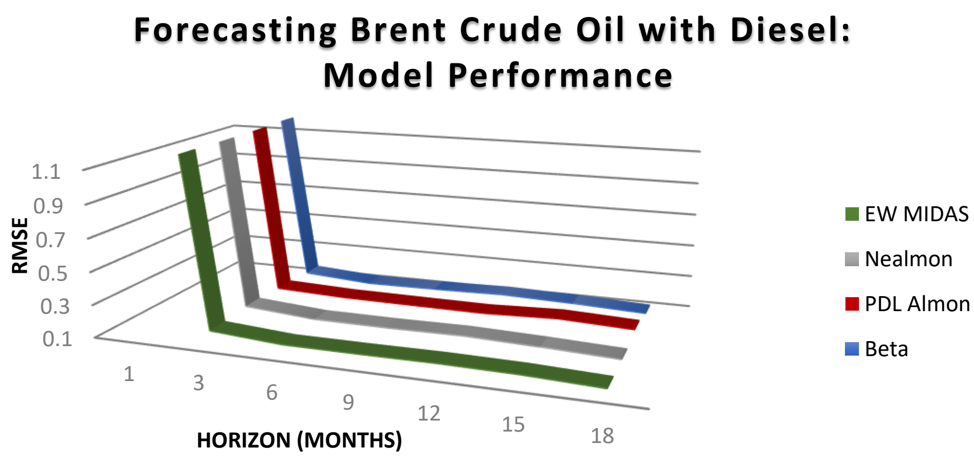

Figure 10

Table 4 in the appendix shows that PDL Almon MIDAS forecasts yield the highest RMSE reduction of 74% relative to the no-change forecast at horizon 18. The RMSE reduction at a horizon of 3 months is as high as 72% and the RMSE reductions are maintained over the horizons of 3 to 18 months. Although the exponential Almon and Beta models provide even better results and the model parameters are highly significant, the imposed restrictions suffer from heteroscedasticity and autocorrelation and therefore the results are not discussed. The majority of the PDL Almon MIDAS regression parameters on the other hand are statistically insignificant. In terms of directional accuracy, the exponential Almon shows the best results, followed by the Beta model. The PDL Almon model also shows high directional accuracy and improvements over the no-change forecast and the results are statistically significant at forecast horizons of 6, 9, 12, 15 and 18 months. Directional accuracy for all the models is successful at month 1. It is plausible that U.S. diesel prices may not serve as the best proxy for applying the product spread to the Brent market.

It can be concluded that the high predictive accuracy of PDL Almon model relative to the no-change forecast cannot compensate for the lack of statistical significance of the in-sample data and that little is lost by ignoring diesel product spreads when forecasting the real price of Brent crude oil.

7.3. Forecasting Brent Crude Oil Prices with Oil Stock Market Returns

In my research I also discussed the claim of oil stocks as a powerful predictor in determining future oil prices. One would expect oil stocks to reflect important signals regarding the price of oil, which might be weakened when looking at general stock returns from the S&P 500 index, for instance. For this purpose, I used the NYSE Arca Oil and Gas Index, which consists of many important international oil companies, to see whether it can help predict the real price of oil. The return on the oil stocks, recorded at a weekly frequency, was converted to its natural logarithm and the difference from one week to the next was used to predict the monthly change in the real price of Brent crude oil.

For all three MIDAS regressions, the model summary in table 1 shows that changes in today's spot price of crude oil have a positive impact on the change in tomorrow's oil price, however the results are statistically insignificant. The regression estimates of the in-sample data for the exponential Almon and Beta models show an overall positive relationship between change in the monthly real price of crude oil and change in the oil stock returns at a 0% significance level. The MIDAS weight coefficients of the Beta model show higher statistical significance at 0% compared to the exponential Almon model. The hyper-parameters of the PDL Almon model are also statistically significant with the exception of the first weight coefficient which lacks explanatory power in determining the price of oil. The positive common slope coefficient, obtained by converting changes in oil stock returns from the weekly to the monthly frequency through the MIDAS weight coefficients, signal that oil stock prices and oil prices move together. This is in line with most empirical findings which show that, in times of high uncertainty, investors tend to shy away from both commodities and the stock market and vice versa. This causes the price of oil to move in the same direction as the stock market.

Although the NLS problem converges for the exponential Almon, PDL Almon and Beta models, testing the adequacy of the remaining restriction in table 2 shows that all model restrictions suffer from heteroscedasticity and autocorrelation. Since the model suffers from parameter estimation uncertainty, it is difficult to assess the reliability of the MIDAS models in terms of their predictive accuracy.

The graph below presents the individual and average MIDAS model performance and it is clear that it was not possible to beat the no-change forecast at forecast horizon of 1 month. All models predict the direction of change of the oil prices correctly at the first horizon. The RMSE ratio for the MIDAS models is almost 9 compared to the no-change forecast for the one period ahead forecast, indicating that the no-change forecast is clearly the better predictor. However predictive accuracy improved impressively for forecast horizons of 3, 6, 9, 12, 15 and 18 months. The 3D graph in fat shows that the RMSE ratios over the horizons are almost monotonically decreasing for all MIDAS specifications.

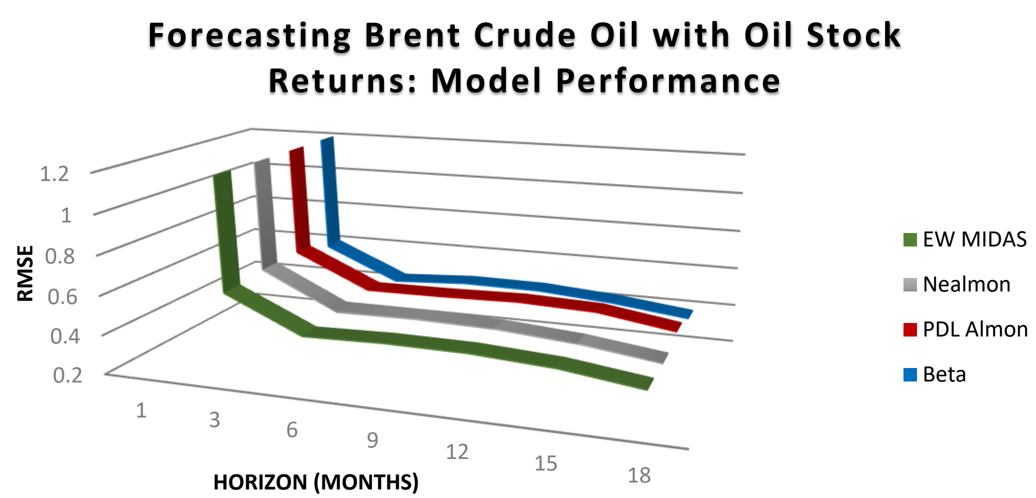

Figure 11

Table 5 in the appendix shows that all three MIDAS models provide similar results. Substantial improvements in RMSE ratios can be seen between forecast horizons of 3 and 6 months. The largest RMSE reductions are achieved for the Beta MIDAS specification. At horizon 3 this models shows 40% predictive accuracy, which increases to 56% at a forecast horizon of 6 months. The reductions in the RMSE ratio are maintained throughout the forecast horizon and biggest improvement of 62% can be seen at 18 months. The PDL Almon model is slightly less accurate than the exponential Almon and Beta models. Statistically significant improvements in directional accuracy over the no-change forecast can be seen in case of the exponential Almon and Beta model, where the latter clearly outperforms all other models through forecast horizons of 3 to 18 months. The PDL Almon model has higher directional accuracy only at horizon 18, it however lacks statistical significance. However, as mentioned, due to parameter estimation uncertainty and the poor performance at forecast horizon of one month there is no reason to prefer either MIDAS model over the no-change forecast.

7.4. Forecasting Brent Crude Oil Prices with U.S. Crude Oil Inventories

The role of crude oil inventories in capturing the future price of oil is of high importance in economic literature since in the short run, oil prices are believed to be more responsive to changes in oil inventories as compared to changes in oil production levels. Hence, changes in crude oil inventories are expected to reflect changes in expectations about the real price of oil, all else equal (Alquist and Kilian, 2010). Baumeister et. al (2014) showed that using weekly U.S. oil inventories was very successful in forecasting the monthly real price of WTI crude oil, particularly at longer forecast horizons using unrestricted MIDAS models. Data for global crude oil inventories was not available at a weekly frequency and data for global oil production was only available at a quarterly or annual frequency. Therefore, the logarithmic difference of U.S. weekly crude oil inventories had to be used to predict the change in the monthly real price of Brent crude oil.

Table 1 summarizes the regression results for the exponential Almon, PDL Almon and Beta MIDAS models. All models suffer from parameter estimation uncertainty. The first lag of Brent crude oil, the common slope coefficient and most of the MIDAS weight coefficients are statistically insignificant. The failure of this model does not come as a surprise since the international crude oil benchmark has shifted to the Brent over the recent years and currently there is an oil oversupply in the U.S. Thus, it is not realistic to expect U.S. oil inventories to contain explanatory power which may influence the future price of Brent crude oil.

Furthermore, in table 2 one can see that while the MIDAS coefficient restrictions are adequate for all three MIDAS models, the non-linear least square problem fails to converge for the Beta MIDAS specification. Hence, only the exponential Almon and PDL MIDAS model forecasts can be discussed although these results lack explanatory power. Unfortunately, these results can be regarded as irrelevant.

The graph below shows that the exponential Almon and PDL Almon models both perform very poorly relative to the no-change forecast at a horizon of 1 month. Big improvements in predictive accuracy can be seen between months 3 and 6. After these horizons the forecast improvements remain constant.

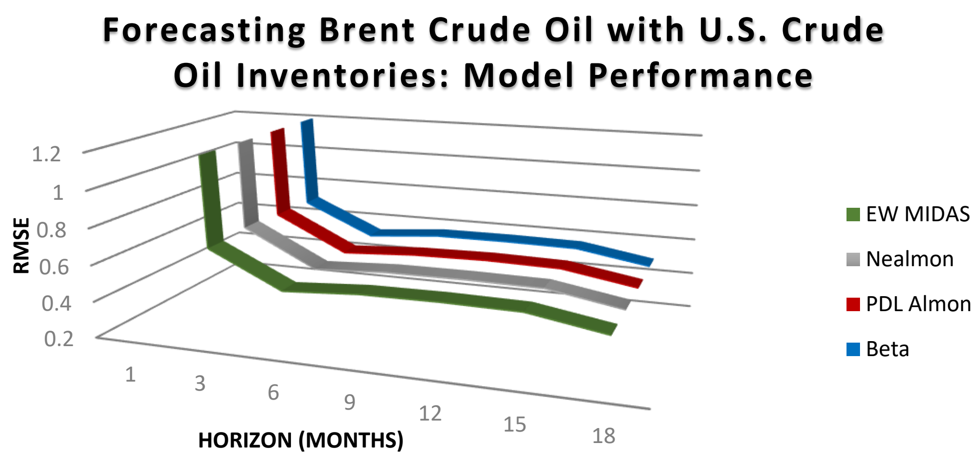

Figure 12

Table 6 shows that the largest RMSE reduction of 52% relative to the no-change forecast is realized at a forecast horizon of 18 months by the exponential Almon and PDL Almon MIDAS models. At a forecast horizon of 3 months, RSME reductions are moderate at 17% for both models. Between horizons of 6 and 15 months, predictive accuracy of the MIDAS models improves and the RMSE ratio reductions become constant at about 45%. While the no-change forecast shows superior performance at a horizon of 1 month, there is no difference between MIDAS and no-change models in terms of directional accuracy performance for the one-step ahead forecast. Directional accuracy of the MIDAS models compared to the no-change forecast at other horizons is poor. Modest improvements can be seen at forecast horizon 18, however only the success ratio for the exponential Almon specification is statistically significant at 67%. Hence there is no reason to prefer one MIDAS specification over the other. One may conclude that trying to use U.S. crude oil inventories to predict changes in the real price of Brent crude oil has no benefit, whatsoever.

7.5. Forecasting Brent Crude Oil Prices with Industrial Commodity Prices

It is also worthwhile to explore the predictive relationship between changes in non-oil commodities and changes in the real price of crude oil for two reasons. Firstly, a fraction of the observed volatility in oil prices is said to stem from the spot prices of industrial commodities, all else equal. Secondly, price movements of many energy commodities have started to converge over the years due to increased financial activity in both oil and non-oil or industrial commodity futures markets. The advantage of using non-oil commodities in forecasting oil prices is that their prices are determined through more transparent and foreseeable variations in global economic activity. This may prove to be useful in forecasting oil prices, especially in times of low volatility. I computed the change in the global metal prices and global agricultural raw materials prices, respectively, by taking the natural logarithmic difference between one month and the next. These explanatory variables were then implemented into two separate univariate MIDAS models.

The MIDAS regression summary results, using the global metal index, in table 1 show high statistical significance for the exponential Almon model followed by the Beta model. The parameters of the PDL Almon specification lack statistical significance. In the case of the exponential Almon specification, the change in the current price of oil has a positive impact on the change on tomorrow's crude oil price at a 1% significance level. Furthermore, the common slope coefficient is positive, implying that change in monthly metal prices has a positive impact on the change in the quarterly real price of crude oil at a 1% significance level. The MIDAS weight coefficients are also statistically significant at a 5% level. The regression results for the Beta MIDAS specification are similar, with the exception of changes in the current quarterly real price of crude oil, which lacks explanatory power. This gives good reason to suspect that oil prices variations can be explained through variations in other important non-oil commodities and that these prices move together.

The regression summary of the in-sample data is less promising in the case of the agricultural raw materials index. The exponential Almon and PDL Almon model are, by large, statistically insignificant. The Beta model finds the intercept to be statistically significant at a 10% level. The common slope coefficient and the first MIDAS weight coefficient also have high explanatory power at a 0% significance level. In line with the literature, a positive relationship between movements in oil prices and agricultural raw materials prices can be confirmed through the Beta MIDAS specification.

The Deriv test in table 2 shows that the NLS problem converges for all MIDAS specifications and that none of the MIDAS restrictions suffer from heteroscedasticity or autocorrelation. This holds true for both global metal prices and the global agricultural price MIDAS forecasts. Due to a lack of statistical significance for most models, only the more promising MIDAS results will be discussed.

The 3D graph below compares the performance of the individual MIDAS forecasts using monthly global metal prices relative to the no-change forecast. It can be seen that the exponential Almon model outperforms the Beta and PDL Almon model. In terms of directional accuracy for the one-step ahead forecast, all models predict the direction of change in oil prices correctly.

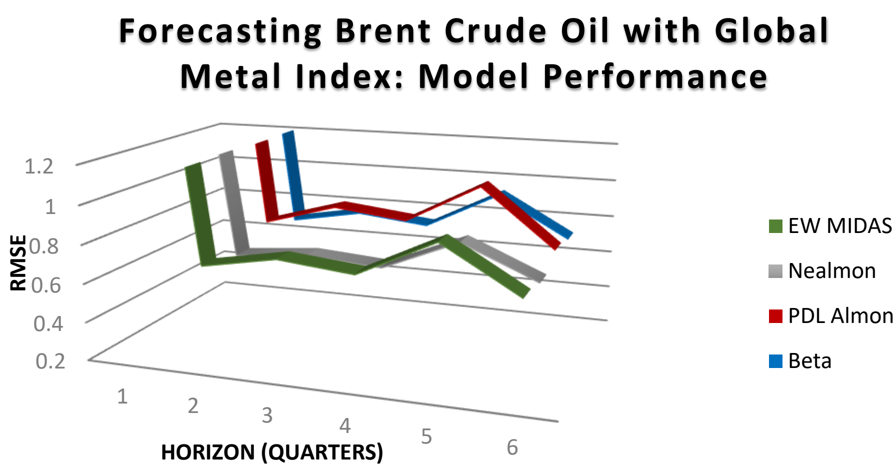

Figure 13

Table 7 shows that while none of the MIDAS specifications outperform the no-change forecast at horizon 1, the RMSE reduction at a forecast horizon of 2 quarters is 32% in the case of the exponential Almon MIDAS model. The gains in predictive accuracy are maintained at forecast horizons of 2, 3, 4 and 6 quarters. In quarter 5 however, one can see a decline in the predictive accuracy with an RMSE reduction of only 16%. A similar pattern is identified for the Beta and PDL MIDAS models. Directional accuracy of the MIDAS models compared to the no-change forecast is worse and statistically insignificant. Although the success ratio is greater than 0.5, the null hypothesis of no directional accuracy cannot be rejected.

When forecasting the changes in the quarterly real price of crude oil using changes in the monthly agricultural raw materials prices, the graph below shows that there is no reason to prefer one MIDAS model over the other. As in the case of all variables till now, the MIDAS models for using non-oil commodities were also not able to outperform the no-change forecast for the one-step ahead forecasts.

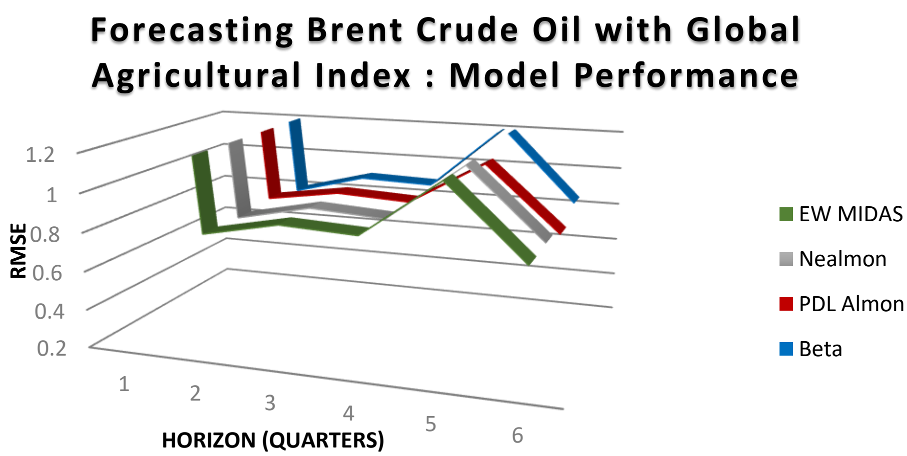

Figure 14

Table 9 compares the performance of the exponential Almon, the PDL Almon and Beta MIDAS models. However, due to parameter estimation uncertainty I only discuss the performance of the Beta model relative to the no-change forecast. The largest RMSE reduction is 20% at a horizon of 2 quarters. At horizons 3 and 4 quarters. a deterioration in the RMSE reductions at 10% is evident. At the forecast horizon over 5 quarters the no-change forecast clearly performs better than the Beta MIDAS model. Directional accuracy was achieved successfully for all models over the first forecast horizon. In terms of directional accuracy for the remaining horizons, the findings are the same as in the case of forecasting changes in crude oil prices using the changes in the global metal index. The Pesaran Timmerman test results are statistically insignificant and perform poorly relative to the no-change forecast. Since improvements in predictive accuracy are erratic over longer forecast horizons, one may conclude that using agricultural raw materials for predicting crude oil prices is only beneficial at very short forecast horizons, if at all.

7.6. Forecasting Brent Crude Oil Prices with EURIBOR Interest Rates

Last but not least, the literature review discussed the predictive relationship between interest rates and the real price of oil. An inverse correlation has been found between changes in the interest rate and changes in the spot price of crude oil. Furthermore, interest rates also effect the exchange rate and purchasing power of countries, which in turn has varying implications for oil importing and exporting countries. I investigate this predictive relationship using MIDAS models. Since the interest rates were already in percentages, changes in the monthly interest rate were measured by simply taking the difference in the EURIBOR rates from one month to the next.

The MIDAS regressions in table 1 show that changes in tomorrow's quarterly real price of crude oil are positively affected by changes in the current real price of crude oil at a 5% significance level. However the exponential Almon and the Beta MIDAS models both suffer from parameter estimation uncertainty with most of the MIDAS coefficients being statistically insignificant. Never the less, a negative relationship between the common slope coefficient (representing the higher frequency interest rate changes converted to the lower frequency) and the low frequency crude oil price changes is evident. This is in line with the findings in section 3. It is only in the case of the PDL Almon specification that the model is statistically significant.

Table 2 tested the adequacy of the MIDAS restrictions. While all three MIDAS models found convergence for the NLS optimization function, the null hypothesis stating that the MIDAS coefficient restrictions are adequate was rejected strongly at a 0% significance level in the case of the exponential Almon and Beta MIDAS augmentations. For the PDL Almon model the restrictions were adequate and did not suffer from autocorrelation and heteroscedasticity. Hence I focus on comparing the performance of the PDL Almon MIDAS model relative to the no-change forecast in predicting changes in the real price of crude oil.

The red line in the 3D graph below shows that the PDL Almon MIDAS model is not better than the no-change forecast at horizon 1. The second quarter sees improvements in predictive accuracy relative to the no-change forecast, these are however not sustainable through all the quarters of the forecast horizon.

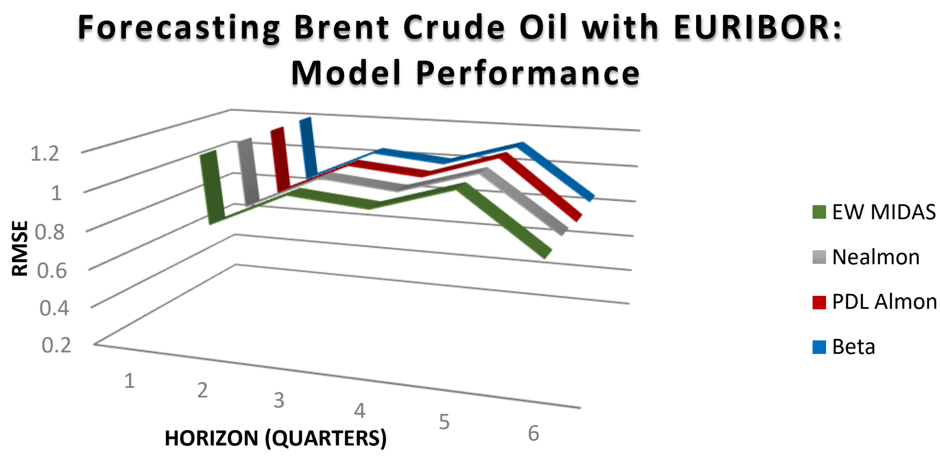

Figure 15

In table 9 of the appendix one can see that the PDL Almon MIDAS model beats the no-change forecast only at quarters 2 and 6 of the forecast horizons. The gains in predictive accuracy over longer forecast horizons are minimal. Over a horizon of 2 quarters the improvement in predictive accuracy is 14% and at horizon 6 the RMSE reduction rises to 16%. In the first quarter, all MIDAS models and the no-change model predict the direction of change in crude oil spot prices successfully. For the remaining horizons, MIDAS forecasts show low and statistically insignificant directional accuracy compared to the no-change forecast through quarters 2, 3, 4 ,5 and 6. Baumeister et al. (2014) showed similar results and suggested that while theoretically there is a strong link between the oil prices and interest rates, from a quantitative perspective, one should be skeptical when linking oil price fluctuations to interest rates.

8. Analysis in Context: Failure of MIDAS Models at the First Forecast Horizon

The preceding section demonstrated that the no-change forecast is the most accurate predictor of the change in the real price of crude oil at the first forecast horizon which is the one-step ahead or true forecast. In case of the one-step ahead forecasts, none of the financial and energy market high-frequency variables, including promising predictors such as an oil futures spread and global metal prices succeed in improving predictive accuracy compared with the random walk (without drift) model. This section investigates the reason behind the failure of the MIDAS models at the first forecast horizon. I open my argument with the claim that the type and impact of the news influences forecasting volatility over both, single and multiple forecast horizons. Chen and Ghysels (2011) show that sudden very good or bad news, makes it difficult to predict the volatility in the S&P 500 futures market with the latter having a more severe impact on the forecast performance. For this purpose, they split the in-sample data into two evaluation periods to produce volatility forecasts using different types of forecasting models including a semi parametric MIDAS regression model and an asymmetric GARCH model. The first evaluation period produced forecasts for a rather calm out-of-sample period and the second evaluation period ended right before the brink of the financial crises of 2008. They showed that the differences in predicted volatility for the out-of-sample period and actual realized volatility were rather small during calm times. However, during the crises the root mean square forecast errors were 20 times larger in comparison to the evaluation period used to forecast results in calm times. They also showed that whilst all forecast models performed poorly where the evaluation period ended right before the beginning of the financial crises, differences between the type of forecasting model used were a lot more pronounced. The semi parametric MIDAS models did not fare as well as the simple asymmetric GARCH models during periods of extremely good and bad news (Chen and Ghysels, 2011).

Having said that bad news tends to have a more severe impact on forecasts, the in-sample period used for the forecast evaluation in this thesis ends in December 2014. It was just after mid-June 2014 that the real price of Brent crude oil began to drop steadily. However, a real crash where crude oil spot prices almost fell by half, did not occur until December 2014. The no-change forecast predicts the change in the crude oil price for January 2015 to be equal to the drop in crude oil prices from November to December 2014. Thus the no-change forecast was more responsive in capturing the effects of the oil price crash compared to the MIDAS forecasts which use 20 years of evaluation period and changes in the high-frequency financial market data over several lags to predict the change in the real price of oil for January 2015. The rationale for using such a long evaluation period is to minimize the risk of spurious forecasts. Nevertheless, I wanted to confirm the finding whether the RMSE ratio at the first forecast horizon was unusually high as a result of the unforeseen oil price crash or if there was ground to suspect a general failure of the MIDAS models implemented in this research. For this purpose, I reproduced the MIDAS forecasts based on a different evaluation period which roughly coincides with the evaluation period used by Baumeister et al. (2014) in the reference paper. The out-of-sample period for which the new MIDAS forecasts were produced was calm. It begins in January 2013 and ends in June 2014. The results are summarized in tables 12 to 18 of the appendix and will be discussed shortly. But first, another point open for discussion is the impact of bad news in forecasting volatility over multiple forecast horizons. One might question why the MIDAS predictive accuracy begins to improve substantially and outperforms the no-change forecast model after forecast horizons exceeding 1 period. The explanation is straightforward but is only relevant where an unforeseeable upward or downward price trend after the end of a forecast evaluation period is identified. As explained in the section on the forecast methodology, the forecasts for the remaining horizons should be interpreted as pseudo forecasts. Generating pseudo forecasts for the out-of-sample period of January 2015- June 2016 implies that we assume all monthly or quarterly values realized during this period to be unknown. Implicitly, the random walk model without drift is not revised at every period and it will project the change in the price of oil to be constant through the entire length of the out-of-sample period. However, the price of oil continued to drop through 2015 which resulted in very large forecast errors, even for the no-change model. Furthermore, the prediction errors over longer forecast horizons comprise of the cumulative mean square errors over time, which further increases the RMSE values of the no-change forecasts. Hence, the effect of high predictive inaccuracy at horizon 1, diminished as the forecast horizons become longer and as the prediction errors for the random walk model also increased.

Table 10 in the appendix summarizes the MIDAS regression results and table 11 tests the adequacy of the restrictions using a slightly shorter evaluation period and a calmer out-of-sample period. The length of the out-of-sample period remains unchanged, whereas the in-sample period beginning in January 1995 ends in December 2012. In terms of predictive accuracy, the MIDAS estimates for the one-step ahead forecast (at horizon 1) perform substantially better compared to the empirical results discussed in section 7. In fact, some of the MIDAS models outperformed the no-change forecast model at the first forecast horizon. I would briefly like to discuss the performance of those MIDAS models, for which the regression estimates were statistically significant and where the adequacy of the restrictions could be confirmed.

With the exception of the one-step ahead forecasts, predicting changes in the real price of crude oil with the help of the futures spread proved to be promising. Table 12 in the appendix shows that, the exponential Almon and PDL Almon MIDAS forecasts yield high RMSE reductions relative to the no-change forecast between horizons of 3 and 18 months. However, the most significant gains in accuracy, as high as 32% for the PDL model and 30% for the exponential Almon model, were achieved between forecast horizons of 3 and 6 months. At these horizons, the gains in directional accuracy of the forecasts were as high as 83% and statistically significant.

Very impressive improvements in predictive accuracy were shown using diesel product spreads to forecast changes in the real price of crude oil (in table 13). The exponential Almon model showed the most impressive results with an RMSE reduction as high as 83% for the first forecast horizon. For the remaining forecast horizons, the RMSE reductions were less impressive, but sustainable between 20 and 30%. The gains in directional accuracy between horizons of 3 and 6 months were statistically significant and high at 83%. Although the directional accuracy of the exponential Almon MIDAS model declined over longer forecast horizons, it still showed statistically significant and superior performance at 67% compared to the no-change model.

Using oil stocks to forecast oil prices also paid off between forecast horizons of 1 and 9 months. Table 14 shows that the PDL Almon model specification beat the no-change model for the one-step ahead forecasts with 74% predictive accuracy. However, after the first forecast horizon, there was a rapid decline in the predictive accuracy as well as directional accuracy. The RMSE reductions fluctuated between 8% and 16% for forecast horizons of 3, 6 and 9 months. The directional accuracy dropped from 100% to 78% in this period, but remained statistically significant. For the remaining forecast horizons, directional accuracy dropped steadily and the MIDAS models failed to outperform the no-change forecast. This is in line with the finding by Chen (2014) who showed that oil-sensitive stock price indices are powerful tools for predicting crude oil prices at 1 month horizons relative to the no-change forecast. Baumeister et al. (2014) also showed that there are advantages in predicting crude oil prices using oil stocks, but these gains are not very high over multiple forecast horizons.

The MIDAS regression results for the crude oil inventory model were statistically insignificant, which was in line with my expectation. As far as the forecast models relying on non-oil commodities are concerned, although the regression estimates were partially significant, either the NLS problem did not converge or the MIDAS coefficient restrictions were not adequate. Hence, due to parameter estimation uncertainty, these results (available in tables 15-17) were not very helpful. Last but not least, the results for interest rates (in table 18) as a predictor, were in line with the findings of the reference paper by Baumeister et al. (2014). That is, the MIDAS forecasts performed poorly relative to the no-change forecast with RMSE ratios above one and the forecasts typically lacked directional accuracy.

Comparing the forecast results in tables 12-18 to the results in tables 3-9 of the appendix, it can be concluded that working with MIDAS models using financial data harbors advantages as well as disadvantages. On the one hand, we now know that it is largely due to the unexpected crash in world oil markets, that the MIDAS forecasts were highly inaccurate compared to the no-change forecasts at horizon 1. While this is good news, it is inevitably accompanied by some bad news- namely, that in times of extreme volatility, MIDAS models should not be favored since they do not prove to be reliable in predicting movements in the oil market. While forecasting with any model will lead to large prediction errors in times of sudden and extreme volatility, simple but powerful forecast models such as the random walk model without drift still perform better in such times.

9. Conclusions

In order to forecast oil prices, I implemented the original MIDAS models using a range of potential high-frequency financial and energy market predictors. The relationship between oil and the oil futures market and non-oil commodities proved to be quite promising. Unfortunately, none of the MIDAS models were able to beat the no-change forecast model at the first forecast horizon. However, it was shown that there are benefits of using univariate MIDAS models over multiple forecast horizons. Forecasting crude oil price movements with movements in the crude oil futures spread showed high predictive accuracy for the MIDAS models. The largest RMSE reduction of 57% was achieved at the horizon of 18 months by the PDL Almon and Beta MIDAS models. The gains in directional accuracy were rather modest at 55% and statistically insignificant. The exponential Almon model also produced promising results when forecasting changes in the price of crude oil with the help of global metal prices. With the exception of horizon 5, the RMSE reductions between horizons of 2 and 6 quarters were sustainable at 32%. While these forecasts achieved high directional accuracy of 80%, so did the no-change forecast model. In terms of individual model performance there is no reason to prefer one MIDAS specification over the other. However, achieving convergence for the non-linear least square optimization function using a PDL Almon MIDAS augmentation is easier. Finding the appropriate start values for the exponential Almon and Beta lag specifications, so that the optimization function can find a local minimum is very challenging and tedious. Hence, implementing MIDAS models successfully, is anything but straightforward.

Yet, the overall failure of MIDAS models can only partially be attributed to the unforeseen oil price crash of 2014. Although most of the MIDAS forecasts performed better than the no-change forecast over multiple forecast horizons, the major problem was the presence of parameter estimation uncertainty, which causes me to deny the validity of the forecasts for most of predictors. This was because most of the MIDAS regression estimates for the in-sample data lacked statistical significance and in many cases the MIDAS coefficient restrictions suffered from autocorrelation and heteroscedasticity. Similar challenges were faced whilst using a different forecast evaluation period. In the context of this research and looking at the one-step ahead predictive accuracy of MIDAS models, it can be concluded that high-frequency financial and energy market data do not appear to have a very significant bearing on improving forecast performance in the short run. Although the high-frequency data may contain information rich signals, it is not powerful enough to compensate for the additional noise, which is said to be the cost of using high-frequency data in forecasting.

However, this should not come as a surprise, since the OPEC till date remains the main influencer of oil prices. Even if the price fluctuations caused by changes in OPEC’s production decisions in oil markets are short lived, they cause significant unrest in the oil futures market by triggering panic among hedgers and worse, encouraging herd behavior among speculators. In fact, most of the trading in the oil futures market is driven by non-commercial trade on behalf of speculators. Hence, the price of oil is largely determined in the oil futures market, rather than by traditional macroeconomic demand and supply mechanisms. In the short run, the higher the noise is being made about an event, the higher will its effect be on the price of crude oil. But this effect may vanish quickly and result in inaccurate forecasts when relying on mixed-frequency data.

Moreover, although macroeconomic variables such as interest rates or global economic GDP growth play a role in determining the price of crude oil, they alone cannot be expected to explain the incessant oil price fluctuations. In recent years rising concerns about shortfalls in global oil supply, political conflicts and terrorism in the Middle East, environmental disasters and increased investment in alternative energy fuels have entered the picture as important factors. Hence, the determinants of global oil prices today extend beyond traditional macroeconomic fundamentals and OPEC's cartel behavior. Without doubt, establishing a clear link between these determinants and measuring their magnitude and implications for the future price of oil is complicated and continues to be a challenge.

Limitations and Scope for Future Research

The scope of the master's thesis was limited by certain aspects, the main one being restrictions in access to data. Initially, the intention was to use European diesel prices to measure product spreads. But the unavailability of this data, prior to 2005, would have meant comprising on the forecast evaluation period and risking spurious forecast results. While, the disappointing results for U.S. crude oil inventories did not come as a surprise, the fact remains that the inclusion of weekly U.S. crude oil inventory data in the MIDAS models documented the best improvements in forecast accuracy in the reference paper by Baumeister et al (2014). Confirming their finding by using global crude oil inventories would have been a reinforcing and an important contribution to existing literature on forecasting oil prices using MIDAS models. However, the lack of high-frequency data for global crude oil inventories hindered the attempt.

Another rather plausible limitation or weakness remains that a univariate model cannot be expected to outperform other models at each horizon. Furthermore, the variables differ in terms of importance in forecasting the changes in crude oil prices. Perhaps a combination of forecasts using multiple variables and models would yield more robust and accurate results. It is therefore inaccurate to state that the univariate models used in this paper are the best or worst models to forecast changes in oil prices, without having implemented a multivariate MIDAS model. Most researchers focus on one high-frequency predictor at a time while forecasting oil prices with MIDAS models. It seems multivariate MIDAS modeling for oil prices did not gain popularity for two reasons. First, the latter approach is considered to be more appealing in the context of forecasting macroeconomic variables. And second, looking at the literature on oil price forecasts using MIDAS models it becomes evident that only a small pool of potential high-frequency predictors is relevant enough to be included in the forecasts. Since the price of oil is determined in global oil markets, the majority of financial and energy market variables fail to show large, statistically significant and systematic improvements in predictive accuracy (Baumeister et al., 2014).

Nevertheless, the findings presented in this research can be used as a strong foundation for understanding the role of the individual high-frequency financial data in forecasting crude oil price movements in conjuncture with the mixed-data sampling approach. For future endeavours, the most promising high-frequency financial and energy market predictors such as the oil futures spread, diesel product spreads, oil stock returns, crude oil inventories and important non-oil commodity indices could be tested together in a multivariate mixed-data sampling model.

10. Appendix

Table 1: Forecasting the Monthly Real Price of Brent Crude Oil with Financial & Energy Market Data
Evaluation Period: January 1995 - December 2014

MIDAS Summary Results

	Exponential Almon				**PDL Almon**				**Beta**			
Variable	Coefficient	Std. Error	t-stat	Prob(>|t|)	Coefficient	Std. Error	t-stat	Prob(>|t|)	Coefficient	Std. Error	t-stat	Prob(>|t|)
Independent					***Brent Futures Spread***							
Intercept	0.004	0.005	0.810	0.419	0.004	0.005	0.728	0.467	0.004	0.005	0.838	0.403
Brent(-1)	0.247	0.104	2.385	0.018**	0.245	0.105	2.339	0.020**	0.216	0.097	2.225	0.027**
Slope	-2.619	0.919	-2.851	0.005***					-6.309	1.814	-3.478	0.000****
θ_1	1.858	0.974	1.908	0.058*	0.381	0.496	0.768	0.443	1.026	0.008	131.115	0.000****
θ_2	-0.409	0.207	-1.982	0.049**	-1.125	0.332	-3.387	0.010***	3.793	1.453	2.611	0.001***
θ_3					0.230	0.067	3.419	0.010***				
Residual S.E.	0.079 on 233 degrees of freedom			4 lags	0.079 on 234 degrees of freedom			4 lags	0.078 on 231 degrees of freedom			16 lags
Independent					***Diesel Spread***							
Intercept	0.002	0.002	0.652	0.515	0.003	0.003	0.974	0.331	0.003	0.002	1.145	0.254
Brent(-1)	0.264	0.038	6.906	0.000****	0.346	0.054	6.456	0.000****	0.238	0.054	4.420	0.000****
Slope	-4.652	0.285	-16.331	0.000****					-4.248	0.272	-15.588	0.000****
θ_1	1.240	0.17	7.389	0.000****	0.416	0.288	1.445	0.150	2.522	0.280	8.993	0.000****
θ_2	-0.170	0.02	-8.044	0.000****	-0.665	0.290	-2.293	0.022**	6.623	0.943	7.020	0.000****
θ_3					0.060	0.066	0.905	0.366				
Residual S.E.	0.040 on 233 degrees of freedom			8 lags	0.053 on 234 degrees of freedom			4 lags	0.042 on 232 degrees of freedom			12 Lags
Independent					***ARCA Oil & Gas Returns on Stock***							
Intercept	-0.001	0.005	-0.162	0.872	-0.003	0.006	-0.518	0.605	-0.001	0.005	-0.210	0.834
Brent(-1)	0.102	0.070	1.447	0.149	0.034	0.069	0.499	0.618	0.117	0.076	1.542	0.124
Slope	3.748	0.608	6.166	0.000****					3.588	0.51	6.979	0.000****
θ_1	0.966	0.478	2.022	0.044**	0.039	0.163	0.237	0.813	2.991	0.78	3.841	0.000****
θ_2	-0.081	0.044	-1.817	0.071*	0.118	0.047	2.532	0.012**	4.347	1.27	3.425	0.000****
θ_3					-0.008	0.002	-3.395	0.000****				
Residual S.E.	0.075 on 233 degrees of freedom			8 lags	0.078 on 231 degrees of freedom			16 lags	0.075 on 232 degrees of freedom			12 lags
Independent					***U.S. Commericial & SPR Crude Oil Inventories***							
Intercept	0.002	0.006	0.295	0.769	0.003	0.006	0.565	0.573	0.004	0.006	0.696	0.487
Brent(-1)	0.187	0.116	1.621	0.11	0.184	0.110	1.664	0.097*	0.174	0.109	1.587	0.114
Slope	5.101	2.886	1.768	0.079*					-4.119	2.13	-1.934	0.054*
θ_1	1.156	2.74	0.422	0.674	-1.478	1.088	-1.359	0.176	1.091	0.09	11.810	0.000****
θ_2	-0.044	0.12	-0.384	0.701	0.307	0.408	0.753	0.452	11.923	11.06	1.078	0.282
θ_3					-0.009	0.032	-0.284	0.777				
Residual S.E.	0.085 on 231 degrees of freedom			16 lags	0.085 on 232 degrees of freedom			12 lags	0.085 on 232 degrees of freedom			12 lags

Using **(0%) ***(1%), **(5%) and *(10%) significance level**

Table 1: Forecasting the Monthly Real Price of Brent Crude Oil with Financial & Energy Market Data
Evaluation Period: January 1995 - December 2014

MIDAS Summary Results

	Exponential Almon				**PDL Almon**				**Beta**			
Variable	Coefficient	Std. Error	t-stat	Prob(>|t|)	Coefficient	Std. Error	t-stat	Prob(>|t|)	Coefficient	Std. Error	t-stat	Prob(>|t|)
Independent					***Global Metal Index***							
Intercept	-0.001	0.016	-0.054	0.957	0.014	0.015	0.931	0.355	0.007	0.016	0.457	0.649
Brent(-1)	0.183	0.067	2.740	0.008***	0.128	0.093	1.366	0.176	0.136	0.107	1.272	0.208
Slope	1.968	0.620	3.172	0.002***					2.782	1.15	2.421	0.018**
θ_1	8.919	4.358	2.046	0.044**	0.782	0.362	2.159	0.034 *	1.049	0.05	21.049	0.000****
θ_2	-1.501	0.741	-2.025	0.047**	-0.096	0.108	-0.891	0.376	4.329	1.74	2.484	0.015**
θ_3					0.001	0.008	0.096	0.924				
Residual S.E.	0.121 on 73 degrees of freedom			6 lags	0.126 on 71 degrees of freedom			12 lags	0.125 on 71 degrees of freedom			12 lags
Independent					***Global Agricultural Raw Materials Index***							
Intercept	0.020	0.013	1.508	0.136	0.016	0.013	1.219	0.227	0.021	0.013	1.669	0.010*
Brent(-1)	0.081	0.118	0.685	0.495	0.114	0.116	0.983	0.329	0.089	0.101	0.882	0.381
Slope	4.165	1.048	3.975	0.000****					3.890	1.05	3.708	0.000****
θ_1	-1.018	0.76	-1.348	0.182	1.686	0.853	1.976	0.052*	0.976	0.02	53.198	0.000****
θ_2	0.099	0.11	0.887	0.378	-0.364	0.286	-1.275	0.207	2.762	2.05	1.345	0.183
θ_3					0.018	0.021	0.863	0.391				
Residual S.E.	0.121 on 73 degrees of freedom			6 lags	0.123 on 71 degrees of freedom			12 lags	0.122 on 72 degrees of freedom			9 lags
Independent					***EURIBOR Interest Rates***							
Intercept	0.000	0.021	-0.017	0.987	0.013	0.011	1.172	0.245	0.000	0.022	0.005	0.996
Brent(-1)	0.290	0.116	2.499	0.015**	0.202	0.091	2.218	0.030**	0.280	0.114	2.457	0.016**
Slope	-0.323	0.184	-1.755	0.083*					-0.314	0.19	-1.626	0.108
θ_1	2.942	10.27	0.287	0.775	0.703	0.253	2.776	0.007***	4.446	8.59	0.518	0.606
θ_2	-0.193	0.70	-0.276	0.783	-0.334	0.143	-2.338	0.022**	1.037	0.14	7.240	0.000****
θ_3					0.032	0.017	1.893	0.063*				
Residual S.E.	0.140 on 72 degrees of freedom			9 lags	0.119 on 73 degrees of freedom			6 lags	0.140 on 72 degrees of freedom			9 lags

Using **(0%) ***(1%), **(5%) and *(10%) significance level**

Table 2: Forecasting the Monthly Real Price of Brent Crude Oil with Financial & Energy Market Data Evaluation Period: January 1995 - December 2014												
Testing the Adequacy of Restrictions												
	Exponential Almon				**PDL Almon**				**Beta**			
Test	value	df	p-value	Success	value	df	p-value	Success	value	df	p-value	Success
Variable	*Brent Futures Spread*											
Hah	1.344	1	0.246		2.377	1	0.123		10.712	13	0.635	
Hahr	0.982	1	0.322		2.529	1	0.112		13.304	13	0.425	
Deriv				**T**				**T**				**T**
Variable	*Diesel Spread*											
Hah	25.416	5	0.000*		0.486	1	0.486		60.343	9	0.000*	
Hahr	32.089	5	0.000*		0.405	1	0.525		50.095	9	0.000*	
Deriv				**T**				**T**				**T**
Variable	*ARCA Oil & Gas Returns on Stock*											
Hah	16.550	5	0.005*		67.001	13	0.000*		35.385	9	0.000*	
Hahr	15.508	5	0.008*		88.713	13	0.000*		24.791	9	0.003*	
Deriv				**T**				**T**				**T**
Variable	*U.S. Commericial & SPR Crude Oil Inventories*											
Hah	8.106	13	0.837		5.869	9	0.753		5.173	9	0.819	
Hahr	10.900	13	0.619		7.121	9	0.625		6.958	9	0.642	
Deriv				**T**				**T**				**F**
Variable	*Global Metal Index*											
Hah	2.168	3	0.538		13.024	9	0.162		12.143	9	0.205	
Hahr	2.487	3	0.478		12.053	9	0.210		11.004	9	0.276	
Deriv				**T**				**T**				**T**
Variable	*Global Agricultural Raw Materials Index*											
Hah	2.316	3	0.509		8.441	9	0.490		3.573	6	0.734	
Hahr	1.963	3	0.580		18.305	9	0.032		3.072	6	0.800	
Deriv				**T**				**T**				**T**
Variable	*EURIBOR Interest Rates*											
Hah	39.133	6	0.000*		3.827	3	0.281		39.192	6	0.000*	
Hahr	31.415	6	0.000*		7.758	3	0.051		27.520	6	0.000*	
Deriv				**T**				**T**				**T**
***Rejection of null hypothesis (H_0: functional constraint on MIDAS coefficients is adequate)**												
T: NLS optimization problem converged; F: NLS optimization problem did not converge												

Table 3: Forecasting the Monthly Real Price of Brent Crude Oil with Brent Futures Spead
Evaluation Period: January 1995 - December 2014
MIDAS Results

Horizon	Exp Almon		PDL Almon		Beta	
(Months)	RMSE Ratio	Success Ratio	RMSE Ratio	Success Ratio	RMSE Ratio	Success Ratio
1	5.480	1.000	5.568	1.000	4.894	1.000
3	**0.654**	0.333	**0.631**	0.333	**0.598**	0.333
6	**0.510**	0.333	**0.495**	0.333	**0.474**	0.333
9	**0.540**	0.333	**0.522**	0.333	**0.509**	0.333
12	**0.538**	0.417	**0.525**	0.417	**0.515**	0.417
15	**0.509**	0.467	**0.501**	0.467	**0.497**	0.467
18	**0.442**	**0.556**	**0.435**	**0.556**	**0.431**	**0.556**

Predictive Accuracy measured via RMSE Ratio = RMSE of MIDAS/ RMSE of No Change Forecast
Directional Accuracy measured via success ratio of Pesaran Timmerman test
Statistically significant improvements in directional accuracy marked using *(5% significance level) and **(10% significance level)
Boldface indicates improvements over the No Change Forecast

Table 4: Forecasting the Monthly Real Price of Brent Crude Oil with Diesel Spead
Evaluation Period: January 1995 - December 2014
MIDAS Results

Horizon	Exp Almon		PDL Almon		Beta	
(Months)	RMSE Ratio	Success Ratio	RMSE Ratio	Success Ratio	RMSE Ratio	Success Ratio
1	3.495	1.000	3.407	1.000	3.026	1.000
3	**0.191**	**1.000**	**0.181**	**1.000**	**0.164**	**1.000**
6	**0.159**	**1.000****	**0.166**	**0.833****	**0.135**	**1.000****
9	**0.157**	**1.000***	**0.162**	**0.778***	**0.140**	**0.889***
12	**0.160**	**1.000***	**0.160**	**0.833***	**0.149**	**0.917***
15	**0.147**	**1.000***	**0.176**	**0.867***	**0.138**	**0.933***
18	**0.130**	**1.000***	**0.159**	**0.889***	**0.122**	**0.944***

Predictive Accuracy measured via RMSE Ratio = RMSE of MIDAS/ RMSE of No Change Forecast
Directional Accuracy measured via success ratio of Pesaran Timmerman test
Statistically significant improvements in directional accuracy marked using *(5% significance level) and **(10% significance level)
Boldface indicates improvements over the No Change Forecast

Table 5: Forecasting the Monthly Real Price of Brent Crude Oil with Return on ARCA Oil & Gas Stocks
Evaluation Period: January 1995 - December 2014
MIDAS Results

Horizon	Exp Almon		PDL Almon		Beta	
(Months)	RMSE Ratio	Success Ratio	RMSE Ratio	Success Ratio	RMSE Ratio	Success Ratio
1	9.130	1.000	8.830	1.000	8.344	1.000
3	**0.658**	0.667	**0.659**	0.330	**0.603**	**1.000***
6	**0.475**	**0.833***	**0.487**	0.500	**0.438**	**1.000***
9	**0.480**	**0.778***	**0.490**	0.556	**0.460**	**0.889***
12	**0.471**	**0.750***	**0.499**	0.500	**0.456**	**0.833***
15	**0.440**	**0.733***	**0.485**	0.533	**0.427**	**0.800***
18	**0.390**	**0.778***	**0.426**	**0.611**	**0.381**	**0.833***

Predictive Accuracy measured via RMSE Ratio = RMSE of MIDAS/ RMSE of No Change Forecast
Directional Accuracy measured via success ratio of Pesaran Timmerman test
Statistically significant improvements in directional accuracy marked using *(5% significance level) and **(10% significance level)
Boldface indicates improvements over the No Change Forecast

Table 6: Forecasting the Monthly Real Price of Brent Crude Oil with U.S. Crude Oil Inventories
Evaluation Period: January 1995 - December 2014
MIDAS Results

Horizon	Exp Almon		PDL Almon		Beta	
(Months)	RMSE Ratio	Success Ratio	RMSE Ratio	Success Ratio	RMSE Ratio	Success Ratio
1	9.300	1.000	7.956	1.000	7.723	1.000
3	**0.734**	0.333	**0.718**	0.333	**0.706**	0.333
6	**0.538**	0.500	**0.527**	0.500	**0.531**	0.333
9	**0.558**	0.556	**0.551**	0.556	**0.564**	0.444
12	**0.560**	0.583	**0.558**	0.417	**0.566**	0.417
15	**0.555**	**0.600**	**0.550**	0.467	**0.561**	0.400
18	**0.487**	**0.667***	**0.483**	**0.556**	**0.494**	**0.500**

Predictive Accuracy measured via RMSE Ratio = RMSE of MIDAS/ RMSE of No Change Forecast
Directional Accuracy measured via success ratio of Pesaran Timmerman test
Statistically significant improvements in directional accuracy marked using *(5% significance level) and **(10% significance level)
Boldface indicates improvements over the No Change Forecast

Table 7: Forecasting the Quarterly Real Price of Brent Crude Oil with Global Metal Index
Evaluation Period: January 1995 - December 2014
MIDAS Results

Horizon	Exp Almon		PDL Almon		Beta	
(Quarters)	RMSE Ratio	Success Ratio	RMSE Ratio	Success Ratio	RMSE Ratio	Success Ratio
1	3.133	1.000	3.678	1.000	3.536	1.000
2	**0.677**	0.500	**0.773**	0.500	**0.702**	0.500
3	**0.710**	0.667	**0.882**	0.667	**0.771**	0.667
4	**0.681**	0.750	**0.840**	0.750	**0.730**	0.750
5	**0.842**	0.800	1.045	0.600	**0.926**	0.800
6	**0.678**	0.667	**0.746**	0.667	**0.715**	0.667

Predictive Accuracy measured via RMSE Ratio = RMSE of MIDAS/ RMSE of No Change Forecast
Directional Accuracy measured via success ratio of Pesaran Timmerman test
Statistically significant improvements in directional accuracy marked using *(5% significance level) and **(10% significance level)
Boldface indicates improvements over the No Change Forecast

Table 8: Forecasting the Quarterly Real Price of Brent Crude Oil with Global Agricultural Raw Materials Index
Evaluation Period: January 1995 - December 2014
MIDAS Results

Horizon	Exp Almon		PDL Almon		Beta	
(Quarters)	RMSE Ratio	Success Ratio	RMSE Ratio	Success Ratio	RMSE Ratio	Success Ratio
1	4.420	1.000	4.469	1.000	4.349	1.000
2	**0.811**	0.500	**0.835**	0.500	**0.801**	0.500
3	**0.892**	0.667	**0.892**	0.667	**0.901**	0.667
4	**0.868**	0.750	**0.872**	0.750	**0.889**	0.750
5	1.163	0.600	1.101	0.600	1.228	0.600
6	**0.814**	0.667	**0.764**	0.667	**0.844**	0.667

Predictive Accuracy measured via RMSE Ratio = RMSE of MIDAS/ RMSE of No Change Forecast
Directional Accuracy measured via success ratio of Pesaran Timmerman test
Statistically significant improvements in directional accuracy marked using *(5% significance level) and **(10% significance level)
Boldface indicates improvements over the No Change Forecast

Table 9: Forecasting the Quarterly Real Price of Brent Crude Oil with EURIBOR Interest Rates
Evaluation Period: January 1995 - December 2014

MIDAS Results

Horizon	Exp Almon		PDL Almon		Beta	
(Quarters)	RMSE Ratio	Success Ratio	RMSE Ratio	Success Ratio	RMSE Ratio	Success Ratio
1	4.376	1.000	4.749	1.000	4.385	1.000
2	**0.861**	0.500	**0.857**	0.500	**0.858**	0.500
3	1.040	0.333	1.035	0.333	1.036	0.333
4	1.001	0.500	1.000	0.500	**0.999**	0.500
5	1.119	0.600	1.128	0.600	1.119	0.600
6	**0.843**	0.500	**0.821**	0.500	**0.841**	0.500

Predictive Accuracy measured via RMSE Ratio = RMSE of MIDAS/ RMSE of No Change Forecast

Directional Accuracy measured via success ratio of Pesaran Timmerman test

Statistically significant improvements in directional accuracy marked using *(5% significance level) and **(10% significance level)

Boldface indicates improvements over the No Change Forecast

Table 10: Forecasting the Monthly Real Price of Brent Crude Oil with Financial & Energy Market Data
Evaluation Period: January 1995 - December 2012

MIDAS Summary Results

	Exponential Almon				PDL Almon				Beta			
Variable	Coefficient	Std. Error	t-stat	Prob(>|t|)	Coefficient	Std. Error	t-stat	Prob(>|t|)	Coefficient	Std. Error	t-stat	Prob(>|t|)
Independent					***Brent Futures Spread***							
Intercept	0.005	0.005	1.014	0.312	0.006	0.006	1.057	0.292	0.007	0.005	1.283	0.201
Brent(-1)	0.218	0.093	2.338	0.020**	0.224	0.110	2.039	0.043**	0.221	0.109	2.021	0.045**
Slope	-4.632	0.994	-4.659	0.000****					-2.988	0.731	-4.086	0.000****
θ_1	0.912	0.354	2.573	0.011**	0.305	0.520	0.586	0.558	1.414	0.566	2.496	0.013**
θ_2	-0.170	0.061	-2.797	0.006***	-1.050	0.343	-3.063	0.002***	2.819	1.364	2.067	0.040**
θ_3					0.211	0.070	3.005	0.003***				
Residual S.E.	0.080 on 208 degrees of freedom			12 Lags	0.081 on 209 degrees of freedom			4 lags	0.082 on 209 degrees of freedom			8 lags
Independent					***Diesel Spread***							
Intercept	0.003	0.003	0.915	0.362	0.004	0.003	1.038	0.300	0.003	0.003	1.158	0.248
Brent(-1)	0.345	0.057	6.054	0.000****	0.346	0.057	6.042	0.000****	0.244	0.055	4.470	0.000****
Slope	-3.197	0.265	-12.049	0.000****					-4.253	0.287	-14.822	0.000****
θ_1	1.685	0.57	2.978	0.003***	0.411	0.295	1.391	0.166	2.533	0.285	8.899	0.000****
θ_2	-0.223	0.10	-2.169	0.031**	-0.665	0.296	-2.246	0.026**	6.706	0.962	6.970 *	0.000****
θ_3					0.060	0.067	0.897	0.371				
Residual S.E.	0.055 on 209 degrees of freedom			4 lags	0.055 on 209 degrees of freedom			4 lags	0.043 on 208 degrees of freedom			12 Lags
Independent					***ARCA Oil & Gas Returns on Stock***							
Intercept	0.002	0.006	0.373	0.709	0.004	0.005	0.687	0.493	0.002	0.006	0.309	0.757
Brent(-1)	0.081	0.072	1.124	0.262	0.058	0.067	0.865	0.388	0.098	0.079	1.244	0.215
Slope	3.692	0.639	5.778	0.000****					3.562	0.53	6.666	0.000****
θ_1	0.977	0.520	1.880	0.0615*	-0.606	0.228	-2.663	0.008***	2.970	0.78	3.806	0.000****
θ_2	-0.080	0.048	-1.676	0.0953*	0.422	0.109	3.887	0.000****	4.239	1.25	3.398	0.000****
θ_3					-0.034	0.012	-2.771	0.006***				
Residual S.E.	0.077 on 209 degrees of freedom			8 lags	0.076 on 209 degrees of freedom			8 lags	0.078 on 208 degrees of freedom			12 lags
Independent					***U.S. Commericial & SPR Crude Oil Inventories***							
Intercept	0.006	0.006	0.959	0.338	0.006	0.006	0.998	0.320	0.006	0.007	0.846	0.399
Brent(-1)	0.172	0.069	2.493	0.013**	0.161	0.119	1.354	0.177	0.169	0.118	1.431	0.154
Slope	2.050	2.587	0.792	0.429					-0.787	3.16	-0.249	0.803
θ_1	-0.944	2.74	0.000	1.000	-0.801	1.857	-0.432	0.666	0.680	21.50	0.000	1.000
θ_2	0.519	1.06	0.000	1.000	-0.247	1.026	-0.240	0.810	2.021	17.47	0.000	1.000
θ_3					0.070	0.117	0.601	0.548				
Residual S.E.	0.087 on 209 degrees of freedom			8 lags	0.087 on 209 degrees of freedom			8 lags	0.088 on 207 degrees of freedom			16 lags

Using **(0%) ***(1%), **(5%) and *(10%) significance level**

Table 10: Forecasting the Monthly Real Price of Brent Crude Oil with Financial & Energy Market Data
Evaluation Period: January 1995 - December 2012

MIDAS Summary Results

	Exponential Almon				**PDL Almon**				**Beta**			
Variable	Coefficient	Std. Error	t-stat	Prob(>|t|)	Coefficient	Std. Error	t-stat	Prob(>|t|)	Coefficient	Std. Error	t-stat	Prob(>|t|)
Independent	***Global Metal Index***											
Intercept	0.008	0.018	0.447	0.656	0.012	0.016	0.744	0.460	0.010	0.017	0.586	0.560
Brent(-1)	0.176	0.065	2.717	0.008***	0.164	0.102	1.605	0.113	0.164	0.072	2.281	0.026**
Slope	2.040	0.705	2.895	0.005***					2.287	0.79	2.912	0.005***
θ_1	7.977	4.380	1.821	0.073*	-0.421	0.620	-0.679	0.500	5.070	2.89	1.753	0.084*
θ_2	-1.331	0.724	-1.839	0.071*	0.733	0.396	1.851	0.069*	13.739	9.02	1.524	0.132
θ_3					-0.113	0.055	-2.054	0.0440 *				
Residual S.E.	0.122 on 64 degrees of freedom			9 lags	0.125 on 65 degrees of freedom			6 lags	0.123 on 64 degrees of freedom			9 lags
Independent	***Global Agricultural Raw Materials Index***											
Intercept	0.029	0.014	2.012	0.048**	0.028	0.014	1.927	0.058*	0.030	0.014	2.099	0.040**
Brent(-1)	0.069	0.110	0.624	0.535	0.083	0.104	0.800	0.427	0.071	0.100	0.703	0.484
Slope	4.242	1.099	3.860	0.000****					3.991	0.94	4.236	0.000****
θ_1	-1.028	0.76	-1.357	0.179	2.435	1.590	1.531	0.131	0.972	0.03	33.646	0.000****
θ_2	0.095	0.12	0.789	0.433	-0.830	0.867	-0.958	0.342	1.529	1.41	1.086	0.282
θ_3					0.078	0.106	0.739	0.463				
Residual S.E.	0.121 on 65 degrees of freedom			6 lags	0.121 on 65 degrees of freedom			6 lags	0.120 on 65 degrees of freedom			6 lags
Independent	***EURIBOR Interest Rates***											
Intercept	0.006	0.024	0.240	0.811	0.021	0.012	1.778	0.080*	0.006	0.024	0.262	0.794
Brent(-1)	0.282	0.120	2.355	0.022**	0.180	0.087	2.071	0.042**	0.271	0.118	2.290	0.025**
Slope	-0.318	0.183	-1.737	0.087*					-0.308	0.20	-1.556	0.125
θ_1	3.059	10.83	282 0.778	0.779	0.710	0.258	2.753	0.008***	4.602	9.64	0.477	0.635
θ_2	-0.203	0.74	-0.275	0.784	-0.334	0.143	-2.339	0.022**	1.043	0.17	6.211	0.000****
θ_3					0.032	0.017	1.917	0.060*				
Residual S.E.	0.141 on 64 degrees of freedom			9 lags	0.119 on 65 degrees of freedom			6 lags	0.141 on 64 degrees of freedom			9 lags

Using **(0%) ***(1%), **(5%) and *(10%) significance level**

Table 11: Forecasting the Monthly Real Price of Brent Crude Oil with Financial & Energy Market Data Evaluation Period: January 1995 - December 2012												
Testing the Adequacy of Restrictions												
	Exponential Almon				**PDL Almon**				**Beta**			
Test	value	df	p-value	Success	value	df	p-value	Success	value	df	p-value	Success
Variable	*Brent Futures Spread*											
Hah	11.093	9	0.269		1.710	1	0.191		15.596	5	0.008*	
Hahr	17.361	9	0.043*		1.997	1	0.158		11.793	5	0.038*	
Deriv				**T**				**T**				**T**
Variable	*Diesel Spread*											
Hah	0.048	1	0.827		0.528	1	0.467		58.554	9	0.000*	
Hahr	0.050	1	0.824		0.468	1	0.494		55.529	9	0.000*	
Deriv				**T**				**T**				**T**
Variable	*ARCA Oil & Gas Returns on Stock*											
Hah	15.461	5	0.009*		9.552	5	0.089		31.932	9	0.000*	
Hahr	15.124	5	0.001*		7.612	5	0.179		23.850	9	0.005*	
Deriv				**F**				**T**				**T**
Variable	*U.S. Commericial & SPR Crude Oil Inventories*											
Hah	4.544	5	0.474		2.253	5	0.813		9.111	13	0.765	
Hahr	5.532	5	0.354		3.191	5	0.671		12.135	13	0.517	
Deriv				**F**				**T**				**F**
Variable	*Global Metal Index*											
Hah	3.916	6	0.688		3.555	3	0.314		4.263	6	0.641	
Hahr	3.388	6	0.759		3.260	3	0.353		4.305	6	0.636	
Deriv				**F**				**T**				**F**
Variable	*Global Agricultural Raw Materials Index*											
Hah	3.066	3	0.382		3.572	3	0.312		2.332	3	0.507	
Hahr	2.951	3	0.399		3.517	3	0.319		2.564	3	0.464	
Deriv				**T**				**T**				**T**
Variable	*EURIBOR Interest Rates*											
Hah	40.501	6	0.000*		3.227	3	0.358		40.541	6	0.000*	
Hahr	32.234	6	0.000*		6.682	3	0.083		28.371	6	0.000*	
Deriv				**T**				**T**				**T**

***Rejection of null hypothesis at 5% sig. level (H_0: functional constraint on MIDAS coefficients is adequate)**

T: NLS optimization problem converged; F: NLS optimization problem did not converge

Table 12: Forecasting the Monthly Real Price of Brent Crude Oil with Brent Futures Spead
Evaluation Period: January 1995 - December 2012
MIDAS Results

Horizon	Exp Almon		PDL Almon		Beta	
(Months)	RMSE Ratio	Success Ratio	RMSE Ratio	Success Ratio	RMSE Ratio	Success Ratio
1	1.663	0.000	1.612	0.000	1.438	0.000
3	**0.831**	0.667	**0.801**	0.667	**0.915**	0.667
6	**0.702**	**0.833***	**0.678**	0.667**	**0.807**	0.500
9	**0.699**	**0.778***	**0.691**	0.667**	**0.783**	0.556
12	**0.829**	0.667	**0.791**	0.583	**0.815**	0.583
15	**0.917**	0.600	**0.884**	0.533	**0.871**	0.533
18	**0.937**	0.556	**0.899**	0.556	**0.871**	0.556

Predictive Accuracy measured via RMSE Ratio = RMSE of MIDAS/ RMSE of No Change Forecast
Directional Accuracy measured via success ratio of Pesaran Timmerman test
Statistically significant improvements in directional accuracy marked using *(5% significance level) and **(10% significance level)
Boldface indicates improvements over the No Change Forecast

Table 13: Forecasting the Monthly Real Price of Brent Crude Oil with Diesel Spead
Evaluation Period: January 1995 - December 2012
MIDAS Results

Horizon	Exp Almon		PDL Almon		Beta	
(Months)	RMSE Ratio	Success Ratio	RMSE Ratio	Success Ratio	RMSE Ratio	Success Ratio
1	**0.172**	1.000	**0.184**	1.000	**0.501**	1.000
3	**0.740**	**0.833***	**0.740**	**0.833***	**0.398**	**1.000***
6	**0.653**	**0.833***	**0.658**	**0.833***	**0.481**	**1.000***
9	**0.615**	**0.778***	**0.619**	**0.778***	**0.450**	**0.889***
12	**0.770**	0.667	**0.763**	0.667	**0.528**	**0.833***
15	**0.767**	**0.667****	**0.763**	**0.667****	**0.527**	**0.800***
18	**0.796**	**0.667****	**0.795**	**0.667****	**0.625**	**0.778***

Predictive Accuracy measured via RMSE Ratio = RMSE of MIDAS/ RMSE of No Change Forecast
Directional Accuracy measured via success ratio of Pesaran Timmerman test
Statistically significant improvements in directional accuracy marked using *(5% significance level) and **(10% significance level)
Boldface indicates improvements over the No Change Forecast

Table 14: Forecasting the Monthly Real Price of Brent Crude Oil with Return on ARCA Oil & Gas Stocks
Evaluation Period: January 1995 - December 2012
MIDAS Results

Horizon	Exp Almon		PDL Almon		Beta	
(Months)	RMSE Ratio	Success Ratio	RMSE Ratio	Success Ratio	RMSE Ratio	Success Ratio
1	**0.604**	1.000	**0.354**	1.000	**0.574**	1.000
3	**0.883**	**1.000***	**0.925**	**1.000***	**0.970**	0.667
6	**0.860**	**0.833***	**0.840**	**0.833***	**0.881**	**0.833***
9	**0.930**	**0.778***	**0.899**	**0.778***	**0.931**	0.778
12	1.049	0.667	1.065	0.667	1.056	0.583
15	1.117	0.600	1.116	0.533	1.080	0.467
18	1.149	0.611	1.143	0.556	1.123	0.500

Predictive Accuracy measured via RMSE Ratio = RMSE of MIDAS/ RMSE of No Change Forecast
Directional Accuracy measured via success ratio of Pesaran Timmerman test
Statistically significant improvements in directional accuracy marked using *(5% significance level) and **(10% significance level)
Boldface indicates improvements over the No Change Forecast

Table 15: Forecasting the Monthly Real Price of Brent Crude Oil with U.S. Crude Oil Inventories
Evaluation Period: January 1995 - December 2012

MIDAS Results

Horizon	Exp Almon		PDL Almon		Beta	
(Months)	RMSE Ratio	Success Ratio	RMSE Ratio	Success Ratio	RMSE Ratio	Success Ratio
1	**0.483**	1.000	**0.639**	1.000	**0.600**	1.000
3	**0.685**	0.667	**0.672**	0.333	**0.651**	0.667
6	**0.549**	0.667	**0.536**	0.500	**0.513**	0.667
9	**0.467**	0.778	**0.438**	0.667	**0.433**	0.778
12	**0.488**	0.667	**0.441**	0.667	**0.447**	0.667
15	**0.482**	0.600	**0.436**	**0.667****	**0.455**	0.533
18	**0.478**	0.611	**0.433**	**0.667****	**0.449**	0.556

Predictive Accuracy measured via RMSE Ratio = RMSE of MIDAS/ RMSE of No Change Forecast

Directional Accuracy measured via success ratio of Pesaran Timmerman test

Statistically significant improvements in directional accuracy marked using *(5% significance level) and **(10% significance level)

Boldface indicates improvements over the No Change Forecast

Table 16: Forecasting the Quarterly Real Price of Brent Crude Oil with Global Metal Index
Evaluation Period: January 1995 - December 2012

MIDAS Results

Horizon	Exp Almon		PDL Almon		Beta	
(Quarters)	RMSE Ratio	Success Ratio	RMSE Ratio	Success Ratio	RMSE Ratio	Success Ratio
1	6.197	1.000	4.944	1.000	7.052	1.000
2	**0.750**	**1.000**	**0.568**	**1.000**	**0.812**	**1.000**
3	**0.774**	**1.000***	**0.806**	0.667	**0.844**	**1.000***
4	**0.818**	**0.750****	**0.868**	0.500	**0.878**	**0.750****
5	**0.810**	**0.800***	**0.859**	**0.600**	**0.866**	**0.800***
6	**0.817**	**0.833***	**0.922**	0.500	**0.869**	**0.667**

Predictive Accuracy measured via RMSE Ratio = RMSE of MIDAS/ RMSE of No Change Forecast

Directional Accuracy measured via success ratio of Pesaran Timmerman test

Statistically significant improvements in directional accuracy marked using *(5% significance level) and **(10% significance level)

Boldface indicates improvements over the No Change Forecast

Table 17: Forecasting the Quarterly Real Price of Brent Crude Oil with Global Agricultural Raw Materials Index
Evaluation Period: January 1995 - December 2012

MIDAS Results

Horizon	Exp Almon		PDL Almon		Beta	
(Quarters)	RMSE Ratio	Success Ratio	RMSE Ratio	Success Ratio	RMSE Ratio	Success Ratio
1	**0.665**	1.000	**0.726**	1.000	**0.209**	1.000
2	1.953	0.500	1.961	0.500	1.907	0.500
3	1.721	0.667	1.760	0.333	1.671	0.667
4	1.865	0.500	1.921	0.250	1.852	0.500
5	2.012	0.400	2.030	0.200	1.990	0.400
6	2.016	0.333	2.034	0.167	1.990	0.333

Predictive Accuracy measured via RMSE Ratio = RMSE of MIDAS/ RMSE of No Change Forecast

Directional Accuracy measured via success ratio of Pesaran Timmerman test

Statistically significant improvements in directional accuracy marked using *(5% significance level) and **(10% significance level)

Boldface indicates improvements over the No Change Forecast

Table 18: Forecasting the Quarterly Real Price of Brent Crude Oil with EURIBOR Interest Rates
Evaluation Period: January 1995 - December 2012
MIDAS Results

Horizon	Exp Almon		PDL Almon		Beta	
(Quarters)	RMSE Ratio	Success Ratio	RMSE Ratio	Success Ratio	RMSE Ratio	Success Ratio
1	2.350	1.000	1.174	1.000	3.065	1.000
2	1.134	0.500	1.122	0.500	1.166	0.500
3	1.225	0.333	1.091	0.667	1.248	0.333
4	1.246	0.250	1.137	0.500	1.270	0.250
5	1.238	0.200	1.156	0.400	1.261	0.200
6	1.238	0.167	1.156	0.500	1.261	0.167

Predictive Accuracy measured via RMSE Ratio = RMSE of MIDAS/ RMSE of No Change Forecast

Directional Accuracy measured via success ratio of Pesaran Timmerman test

Statistically significant improvements in directional accuracy marked using *(5% significance level) and **(10% significance level)

Boldface indicates improvements over the No Change Forecast

11.References

Books

[1]. Alquist, R., Kilian, L. and Vigfusson, R.J. (2013). Forecasting the Price of Oil. In *Handbook of Economic Forecasting*, ed. G. Elliott and A. Timmermann, Amsterdam: North-Holland, pp. 427-507.

[2]. Andersen, T., Bollerslev, T., Christoffersen, P. and Diebold, F. (2006). Volatility and Correlation Forecasting. In *Handbook of Economic Forecasting,* ed. E. Graham, C. Granger, and A. Timmermann, Amsterdam: North Holland, pp. 778-878.

[3]. Andreou, E., Ghysels, E. and Kourtellos, A. (2011). Forecasting with Mixed-Frequency Data. In *Oxford Handbook of Economic Forecasting,* ed. M. Clements and D. Hendry, D, pp. 225-245.

[4]. Bhansali, R. (1999). Parameter Estimation and Model Selection for Multistep Prediction of a Time Series: A Review. In *Asymptotics, Nonparametrics and Time Series*, ed. S. Ghosh. New York: Marcel Dekker. CRC Press, Boca Raton, pp. 201- 258.

[5]. Findley, D. (1985). Model Selection for Multi-Step-Ahead Forecasting. In *Proceedings of the 7th Symposium on Identification and System Parameter Estimation*, ed. S. Ghosh, S. and P. Young. Oxford: Pergamon, pp. 1039-1044.

[6]. Findley, D. (1983). On the Use of Multiple Models of Multi-period Forecasting. In *Proceedings of the Business and Statistics Section,* Americal Statistical Association, pp. 528-531.

[7]. Harvey, A. (1989). Forecasting, Structural Time Series Models and the Kalman Filter. Cambridge: Cambridge University Press.

[8]. Simmons, M.R. (2005). Twilight in the desert- The Coming Saudi Oil Shock and the World Economy. New Jersey: John Wiley & Sons.

Journals and Articles

[9]. Acquah de Graft, H. (2010). Comparison of Akaike Information Criterion and Bayesian Information Criterion in Selection of an Asymmetric Price Relationship. *Journal of Development and Agricultural Economics*, 2(1), 1-6.

[10]. Adelman, M.A. (1980). The Clumsy Cartel. *The Energy Journal*, 1(1), 43-54.

[11]. Alquist, R. and Kilian, L. (2010). What Do We Learn from the Price of Crude Oil Futures? *Journal of Applied Econometrics,* 25, 539-573.

[12]. Andreou, E., Ghysels, E. and Kourtellos, A. (2010). Regression Models with Mixed Sampling Frequencies. *Journal of Econometrics*, 158, 246-261.

[13]. Armesto, M.T., Engemann, K.M. and Owyang, M.T. (2010). Forecasting with Mixed Frequencies. *Federal Reserve Bank of St. Louis Review*, 92(6), 521-36.

[14]. Asimakopoulos, S., Paredes, J. and Warmedinger, T. (2013). Forecasting Fiscal Time Series using Mixed-Frequency Data. *European Central Bank*, Working Paper Series No. 1550.

[15]. Bai, J., Ghysels, E. and Wright, J. (2013). State Space Models and MIDAS Regressions. *Econometric Reviews*, 32(7), 779-813.

[16]. Barsky, R.B. and Kilian, L. (2004), Oil and the Macro Economy since the 1970s. *Journal of Economic Perspectives*, 18, 115–134.

[17]. Baumeister, C., Guérin, P. and Kilian, L. (2014). Do High- Frequency Financial Data Help Forecast Oil Prices? The MIDAS Touch at Work. *International Journal of Forecasting,* 31(2), 238-252.

[18]. Baumeister, C., Kilian, L. and Zhou, X. (2013). Are Product Spreads Useful for Forecasting? An Empirical Evaluation of the Verleger Hypothesis. Bank of Canada, Working Paper No. 2013-25.

[19]. Behmiri, N.B. and Pires Manso, J.R. (2013). Crude Oil Price Forecasting Techniques: A comprehensive Review of Literature. *SSRN Electronic Journal,* 1-35

[20]. Bernanke, B., Gertler, M. and Watson, M. (1997). Systematic Monetary Policy and the Effects of Oil Price Shocks. *Brookings Papers on Economic Activity*, 1, 91-157.

[21]. Degiannakis, S., Filis, G. and Floros, C. (2011). Dynamic Correlation Between Stock Market and Oil Prices: The case of oil-importing and oil-exporting countries. *International Review of Financial Analysis*, 20(3), 152-164.

[22]. De Bondt Werner, F. M. and Thaler, R.H. (1990). Do Security Analysts Overreact? *The American Economic Review*, 80(2), 52-57.

[23]. Chen, S.S. (2014). Forecasting Oil Price Movements with Oil- Sensitive Stocks. *Economic Inquiry*, 52(2), 830-844.

[24]. Chen, X. and Ghysels, E. (2011). News- Good or Bad- and its Impact on Volatility Predictions over Multiple Horizons. Oxford Journals, *The Review of Financial Studies*, 24(1), 46-81.

[25]. Chevillon, G. and Hendry, D. (2005). Non-parametric Direct Multi-step Estimation for Forecasting Economic Processes. *International Journal of Forecasting*, 21, 201-218.

[26]. Chevillon, G. and Rifflart, C. (2009). Physical Market Determinants of the Price of Crude Oil and the Market Premium. *Energy Economics*, 31(4), 537-549.

[27]. Clements, M. and Hendry, D. (1996). Mutli-Step Estimation for Forecasting. *Oxford Bulletin of Economics and Statistics*, 58, 657-684.

[28]. Foroni, C., Marcellino M. and Schumacher C. (2014). U-MIDAS: MIDAS Regressions with Unrestricted Lag Polynomials. *Journal of Royal Statistical Society Series A*, 178(1), 57-82.

[29]. Ghysels, E., Kvedaras, V. and Zemlys, V. (2016). Mixed-Frequency Data Sampling Regression Models: The R Package midasr. *Journal of Statistical Software*, 72(4), 1-35.

[30]. Ghysels, E., Santa-Clara, P. and Valkanov, R. (2004). The MIDAS Touch: Mixed Data Sampling Regression Models. Working paper, UNC and UCLA.

[31]. Ghysels, E., Sinko, A. and Valkanov, R. (2006a). Predicting Volatility: Getting the most out of Return Data Sampled at Different Frequencies. *Econometric Reviews*, 26(1),53-90.

[32]. Ghysels, E., Sinko, A. and Valkanov, R. (2006b). MIDAS Regressions: Further Results and New Directions. *Econometric Reviews*, 26(1), 53-90.

[33]. Götz, B.T., Hecq, A. and Urbain, J.P. (2014). Forecasting Mixed-Frequency Time Series with ECM-MIDAS Models. *Journal of Forecasting*, 33(3), 198-213.

[34]. Hamilton, J.D. (2009a). Understanding Crude Oil Prices. *The Energy Journal*, 30(2), 179-206.

[35]. Hamilton, J.D. (2009b). Causes and Consequences of the Oil Shock of 2007-08. *Brookings Papers on Economic Activity*, 40(1), 215-283.

[36]. Happonen, J. (2009). A Review of Factors Determining Oil Prices. Working paper.

[37]. Harvey, A. C., and Pierse, R.G. (1984). Estimating Missing Observations in Economic Time Series. *Journal of the American Statistical Association*, 79, 125-131.

[38]. Hotelling, H. (1931). The Economics of Exhaustible Resources. *The Journal of Political Economy*, 39(2), 137–175.

[39]. Huntington, H. (1994). Oil Price Forecasting in the 1980s: what went wrong? *Energy Journal*, 15, 1-22.

[40]. Kilian, L. (2009). Not all Oil Price Shocks are alike: Disentangling Demand and Supply Shocks in the Crude Oil Market. *American Economic Review*, 99(3), 1053-1069.

[41]. Kuzin, V., Marcellino, M. and Schumacher, C. (2011). MIDAS versus Mixed-Frequency VAR: Nowcasting GDP in the Euro Area. *International Journal of Forecasting*, 27(2), 529-542.

[42]. Kvedaras, V. and Zemlys, V. (2012). Testing the Functional Constraints on Parameters in Regressions with Variables of Different Frequency. *Economics Letters*, 116(2), 250-254.

[43]. Lin, J. and Granger, C. (1994). Model Selection for Multi-period Forecasts. *Journal of Forecasting*, 13, 1-9.

[44]. Manescu, C. and Robays I.V. (2014). Forecasting the Brent Oil Price: Addressing Time-variation in Forecast Performance", European Central Bank. Working Paper Series, No. 1735.

[45]. Marcellino, M., Stock, J. and Watson, M. (2006). A Comparison of Direct and Iterated multistep AR Methods for Forecasting Macroeconomic Time Series. *Journal of Econometrics*, 135, 499-526.

[46]. Mariano, R. S., and Murasawa, Y. (2003). A New Coincident Index of Business Cycles based on Monthly and Quarterly Series. *Journal of Applied Econometrics*, 18, 427-443.

[47]. Merino, A. and Ortiz, A. (2005). Explaining the so-called "Price Premium" in Oil Markets. *OPEC Energy Review*, 29(2), 133-52.

[48]. Pindyck, R.S. (1994). Inventories and the Short-run Dynamics of Commodity Prices. *The RAND Journal of Economics*, 25(1), 141-159.

[49]. Pindyck, R.S. (2001). The Dynamics of Commodity Spot and Futures Markets: A Primer. International Association for Energy Economics, *The Energy Journal,* 22(3), 1-29.

[50]. Rafailidis, P. and Katrakilidis, C. (2014). The Relationship between Oil Prices and Stock Prices: A Nonlinear Asymmetric Cointegration Approach. *Applied Financial Economics*, 24(12), 793-800.

[51]. Ready, R.C. (2016). Oil Prices and the Stock Market. The Simon School of Business - University of Rochester

[52]. Shleifer, A. and Vishny, R.W. (1997). The Limits of Arbitrage. *The Journal of Finance*, 52(1), 35-55.

[53]. Wu, T. and McCallum, A.H. (2005). Do Oil Futures Prices Help Predict Future Oil Prices? *Federal Reserve Bank of San Francisco Economic Letter*, No. 2005-38.

[54]. Ye, M., Zyren, J. and Shore, J. (2002). Forecasting Crude Oil Spot Price using OECD Petroleum Inventory Levels. *International Advances in Economic Research*, 8(4), 324-333.

[55]. Verleger, P.K. (2011). The Margin, Currency, and the Price of Oil. *Business Economics*, 46, 71-82.

[56]. Zadrozny, P. A. (1990). Forecasting US GNP at Monthly Intervals with an Estimated Bivariate Time Series Model. *Federal Reserve Bank of Atlanta Economic Review*, 75, 2-15.

Online Articles and Websites Consulted

[57]. BBC News. (2016). Oil rallies after OPEC Ministers announce Output Cut [BBC Business News]. Retrieved on September 29th, 2016 from: http://www.bbc.com/news/business-37502538

[58]. Bernanke, B.S. (2016). The Relationship between Stocks and Oil Prices [The Brooking Institution]. Retrieved on August 24th, 2016 from: https://www.brookings.edu/blog/ben-bernanke/2016/02/19/the-relationship-between-stocks-and-oil-prices/

[59]. Burstein, J. (2015). Copper and Oil Prices: A Look at the Correlation [Admiral Metals]. Retrieved on September 17th, 2016 from: http://www.admiralmetals.com/admiral-metals/copper-oil-prices-a-look-at-the-correlation/

[60]. Chavdarov, P. (2015). Impact of Low Oil Prices on Food and Other Commodities [CEIC-Euromoney Institutional Investor Company Data]. Retrieved on September 10th, 2016 from: https://www.ceicdata.com/en/blog/impact-low-oil-prices-food-and-other-commodities

[61]. Hoyos, C. (2007). The New Seven Sisters: Oil and Gas Giants dwarf Western Rivals [Financial Times]. Retrieved on August 23th, 2016 from: https://www.ft.com/content/471ae1b8-d001-11db-94cb-000b5df10621

[62]. Intercontinental Exchange (2016). ICE Crude & Refined Oil Products. [ICE Report Center]. Retrieved on September 3rd, 2016 from: https://www.theice.com/publicdocs/ICE_Crude_Refined_Oil_Products.pdf

[63]. McMohan, T. (2016). Historical Oil Prices Chart [Inflation Data]. Retrieved on September 1st, 2016 from: http://inflationdata.com/Inflation/Inflation_Rate/Historical_Oil_Prices_Chart.asp

[64]. Organization of Petroleum Exporting Countries (2016). Review of Developments in the World Economy [OPEC Monthly Oil Market Report, September 2016]. Retrieved on October 2nd, 2016 from: http://www.opec.org/opec_web/static_files_project/media/downloads/publications/MOMR%20September%202016.pdf

[65]. Piotrowski, M. (2015). How will a U.S. Fed Interest Rate Hike Impact Oil Prices? [The Fuse]. Retrieved on September 25th, 2016 from: http://energyfuse.org/how-will-a-u-s-fed-interest-rate-hike-impact-oil-prices/

[66]. United States Department of States. (2016). The 1928 Red Line Agreement [Office of the Historian]. Retrieved on August 27th, 2016 from: https://history.state.gov/milestones/1921-1936/red-line

[67]. U.S. Energy Administration Information. (2016). What drives Crude Oil Prices? An Analysis of 7 Factors that Influence Oil Markets [EIA Independent Statistics and Analysis]. Retrieved on July 30th, 2016 from: http://www.eia.gov/finance/markets/reports_presentations/eia_what_drives_crude_oil_prices.pdf

[68]. Watts, W. (2016). Why Crude Oil and Stocks keep plunging together [Market Watch]. Retrieved on August 24th, 2016 from: http://www.marketwatch.com/story/why-crude-oil-and-stocks-keep-plunging-together-2016-01-19

[69]. Wingfield, B. (2014). Oil Prices [Bloomberg]. Retrieved on August 25th, 2016 from: https://www.bloomberg.com/quicktake/oil-prices

Databases

[70]. European Commission Energy Database. (2016). European Commission, Weekly Oil Bulletin.

URL: https://ec.europa.eu/energy/en/data-analysis/weekly-oil-bulletin

[71]. European Money Markets Institute (EMMI) Database. (2016).

URL: http://www.emmi-benchmarks.eu/euribor-org/about-euribor.html

[72]. Federal Reserve Economic Data (FRED) Database. (2016). Federal Reserve Bank of St. Louis.

URL: https://fred.stlouisfed.org/

[73]. Quandl Database. (2016). Quandl Financial and Economic Data.

URL: https://www.quandl.com/data/CHRIS/ICE_B1-Brent-Crude-Futures-Continuous-Contract-1-B1-Front-Month

[74]. U.S. Energy Information Administration (EIA) Database. (2016). International Energy Statistics.

URL: www.eia.gov

[75]. Yahoo Finance Database. (2016).

URL: https://finance.yahoo.com/quote/%5EXOI/history?p=%5EXOI