Copenhagen Business School

Master's Thesis

Finance and Investments

Oslo Stock Exchange's reliance of oil

An empirical analysis of factors affecting Oslo Stock exchange over time

Authors:

Martin Øgreid Danielsen (125228)

Martin Norstrand Torgersen (102616)

Supervisor:

Michael E. Jacobsen

Hand-in date: 15th of May

No. of pages: 111 (Characters: 182,614)

Abstract

This thesis investigates the influence of the Brent crude oil price on the Oslo Stock Exchange, using daily data between January 1st, 1988, and December 31st, 2019. The influence is estimated by utilizing an OLS multiple regression, where variables correspondent with previous research and conceivable interactions are included. Further, the period examined is divided into two subparts, with the first covering the years from 1988 to 2008, and the second 2009-2019. Accordingly, the distinctive regression models are statistically processed to allow for more secure inferences about the true relationships, in addition to the respective changes in influence throughout time. Finally, an additional variable is introduced to the last model in order to assess the dynamics between the challenged petroleum industry and the rising renewable energy sector.

We find the Brent crude oil price positive and statistically significant in all regression models conducted. Moreover, its associated coefficient carries higher value in subperiod two, compared to both subperiod one, and the full period. Indicating that the Oslo Stock Exchange has been more exposed to fluctuations in the oil price. For the entire period, we find that a one percent increase in the oil price leads to a 0.092 increase at the Oslo Stock Exchange. However, exposure to global financial markets prevails as the most decisive factor. The inclusion of the renewable energy variable does not detriment the effect from the Brent crude oil price.

Table of Contents

1. Introduction

When Norway found oil off its coast back in 1969, it turned the country into one of the largest oil producers in the world. The petroleum industry has played a central part in the development of today's welfare society in Norway. Since the production started in the early 1970s, it is estimated that the petroleum activities have contributed more than 14 900 billion to Norway's GDP, determined in current NOK (Norwegian Petroleum, 2020). Moreover, the petroleum industry constitutes the country's largest industry measured in value-added, government revenues, export value, and investments. It is thereby considered the most important industry for Norway (Ibid.).

At Oslo Stock Exchange (OSE), the energy sector has for periods constituted by far the most considerable portion in terms of market capitalization. Mainly, from the beginning of the 2000s, increased energy demand from China and other emerging markets contributed to a rising oil price (Norges Bank, 2020). The extended effects of this were the development of the oil supply industry in Norway, where numerous new companies emerged and contributed to a growing energy sector at OSE. Oil production in the country saw a peak in 2001, and since then, the extracted volume has halved (Norwegian Petroleum, 2020). However, revenue from oil production continued to surge in conjunction with the oil price, which from the end of the 1990s to 2013, improved from around 10 USD to over 100 USD per barrel of Brent North Sea oil.

The Norwegian oil industry has, during recent decades, been affected by considerable fluctuations in the oil price. Before the financial crisis, the oil price reached a historic high, at 140 USD per barrel, while years later, subsequent to the oil price collapse, the price was lowered to only 28 USD at the start of 2016. The positive development preceding the massive price drop led to record high investments in the Norwegian oil sector, together with an unhealthy cost structure among the oil companies (Barstad, 2015). Moreover, this advancement contributed to additional strengthening of the negative price effect, which resulted in barely any oil fields were profitable (Fransvold, 2016).

Following the severe downturn between 2014 and 2016, the oil price has seen a rebound to a certain degree. However, nowhere near the levels seen in advance of the collapse. Previously, when the demand for oil was booming, many were worried about the world's remaining oil reserves, which were regarded as the "Peak Oil" theory (Trocmé & Kruse, 2016). Nevertheless, in recent years these concerns have seen a turnaround, where the long-term uncertainties have shifted from supply to demand. Greater energy efficiency within developed markets, increased interest, and investments in alternative energy sources, together with major political commitments to the environment, represent powerful aspects that diminish the future demand for oil (Ibid.).

Both historically and to this day, Norwegian newspapers and stock analysts reports on how the OSE fluctuate as a result of oil price changes. In this sense, there seems to be a consensus that the Norwegian equity market is sensitive to movements in the commodity. For a nation where extraction of oil has constituted the motor of economic growth, it appears to be intuitive to make predictions based on forecasts of the price of oil. Lower expected oil prices will affect the earnings of the oil producers, and hence their market value. Further, weaker outlooks for the market may lead the oil producers to reduce their investments and cut operational costs, which in turn curtail the expectations and market values for oil supply companies (Brander, Brekke, & Naug, 2013).

The Norwegian economy's oil dependence has mainly been a subject in retrospect of the strong deterioration in the oil price between 2014 and 2016. However, during the same period, OSE was hardly affected. Based on these foundations, the following research question is made:

How has the oil price affected the Oslo Stock Exchange (OSE), and has the influence changed over time?

In order to answer the research question, the following sub-questions are supplemented:

1. How strong is the influence of oil compared to other factors?
2. How will the inclusion of a renewable energy factor affect the influence of the oil price?

1.1. Delimitation of focus

Regarding the focus of this thesis, the only variable to be explained is the index representing the Oslo Stock Exchange. As sector indexes with an advantage could have been included, the purpose is to examine the overall development of the Norwegian equity market. Moreover, the division of companies into sector indices took place subsequent to the preferred start of the estimation period for this thesis. Another limitation worth emphasizing is the lack of ability to draw inferences about shorter time horizons. The reason behind this is the choice to investigate how the oil price has affected OSE over the longest possible time frame. Therefore, covariation of the variables in the long term is preferred over dynamics in the short term. Concerning the time span for the analysis in this thesis, the possibilities for extraction of available data for OSE and the oil price was decisive in order to determine the starting period. Data for the Brent Crude oil was not available before 1988.

Consequently, the analysis does not start before this year. Among other possibly relevant factors for the analysis, there were variables where daily data was not accessible for the chosen period. Of importance, Norwegian long term interest rates were not accessible.

2. Methodology

The following section will present the methodology employed to answer the research question. Firstly, the philosophy of science and the choice of paradigm will be presented. Thereafter, the research approach will be outlined, followed by the method of collecting feasible empirical data. These measures are taken in order to generate reliable and valid findings.

2.1. Philosophy of Science

Research philosophy relates to a system of beliefs and assumptions about the development of knowledge (Saunders, Lewis, & Thornhill, 2019). As the research philosophy adopted in this thesis contains some decisive assumptions about both the nature of reality and the nature of knowledge, these will accordingly be explained.

Questions of the nature of reality are related to ontology, where there exist two broad and contradictory positions. The first position is objectivism, which considers that there prevails an independent reality. On the other hand, constructionism holds that reality is the product of social processes (Tuli, 2010). Researchers with the orientation of objectivism are described as belonging to the positivist paradigm. These believe that reality is "out there", and that it needs to be found through scientific methodologies.

Furthermore, researchers with this viewpoint do not regard themselves as an essential component in their research, as they remain isolated from what they study. In contrast, constructivists, who are said to be a part of the interpretive paradigm, cannot acknowledge the concept of there being a reality out in the world, which exists without the interaction of people. In that way, they have the opinion that reality consists of people's subjective experiences, and thus, that reality is socially constructed by humans (Antwi & Hamza, 2015).

Another area of philosophy is epistemology, which centers around the creation of knowledge. Epistemology focuses on how we know what we know, and what are the most valid ways to attain the truth (Neuman, 2014). Also, within this area, there are two broad positions, positivism and interpretivism. Within the positivistic view, the goal of science is to work out the most objective methods attainable to get to the closest approximation of reality. As an investigator gather empirical evidence, he can either find that his pre-existing ideas about reality can be verified or found consistent with the evidence, or that the ideas are false due to lack of supporting empirical evidence. Thus, researchers with this orientation are striving to explain how variables interact, shape events, and cause outcomes (Tuli, 2010). Contrary, according to interpretivists, the purpose of a study is to understand a specific phenomenon in a particular context, rather than to generalize to a population. They, therefore, argue that knowledge about reality cannot be obtained only through observations, because interpretations and subjective aspects greatly affect all observations (Neuman, 2014).

Given the purpose of investigating how the oil price and other macroeconomic factors have affected the Oslo Stock Exchange in the previous three decades, it is implied that such factors exist irrespective of any interaction between the area of research and the researcher. This thesis will, moreover, operate with detectable and measurable circumstances. Consequently, the ontological orientation for this study is completed within the positivist paradigm.

The foundation for this dissertation is built upon available market data. By exploring the relationship between the OSE and the other factors, the purpose of the research is to explain the underlying causes and effects scientifically. These explanations will be tested experimentally with statistical analysis. Further, the researcher claim to be external to data collection, as it is limited to what can be done to alter the substance of the data collected for this thesis. It is therefore implied that reliable knowledge can be obtained based on direct observation of natural phenomena through empirical means. As such, the epistemological orientation for the dissertation is also achieved through the positivist paradigm.

2.2. Methodological choice and research approach

Methodology constitutes a research strategy that translates ontological and epistemological foundations into guidelines that display how research is to be managed (Tuli, 2010). The choice of the methodology is commonly distinguished between quantitative and qualitative research, where the first strategy gives priority to the use of numbers in the analysis, whereas qualitative research emphasizes the use of non-numerical approaches (Bryman & Bell, 2015). The positivist ontology and epistemology requires a research methodology that is detached or objective, where the attention is on measuring variables and conduct statistical tests that are connected to general causal explanations (Antwi & Hamza, 2015). The data collection technique for this dissertation will focus on gathering hard data for the chosen variables, in the form of numbers, to enable evidence to be presented in a quantitative form.

Furthermore, the research can either make use of a deductive or an inductive research approach. With a deductive approach, the researcher starts with abstract concepts or theoretical propositions which outlines the logical linkage among concepts. From here, the researcher moves next to evaluate the propositions and concepts against concrete evidence (Neuman, 2014). In other words, the deduction begins with an expected pattern that is tested against empirical observations, and thereafter either rejected or confirmed. On the other hand, an inductive approach aims to develop or confirm a theory that starts with concrete empirical evidence. With an inductive viewpoint, a theory is the outcome of research where the researcher seeks to draw generalizable inferences out of observations (Bryman & Bell, 2015). By initially looking at established theories and knowledge, and after that test and analyze how the factors have impacted the development of OSE, this thesis accepts the premises of a deductive methodological approach.

2.3. Data collection

The time horizon selected for this study is longitudinal. The main advantage of choosing a longitudinal research approach is the ability to study development and change over time (Neuman, 2014). Due to the scope behind the research problem for this thesis, there is required to capture and measure the different variables over a longer time horizon. Therefore, a set of panel data with multiple factors are collected and measured over the time frame from 1988 to 2019. In this way, it is implied that the different effects on OSE, together with their possible alterations of influence, will be captured.

Bloomberg has been used as the primary source to extract daily prices for each variable. Nevertheless, Bloomberg did not support all the data required. Hence, Datastream, Oslo Børs, and Norges Bank were used as complementary sources. Daily data has been chosen for the period 01.01.1988- 30.12.2019, as this was the oldest data we could find for the Brent crude oil.

The data has been adjusted for Norwegian holidays and other days with no trading. For dates where some variables have missing observations due to foreign holidays or other various reasons, linear interpolation has been used to estimate the prices. All data has been processed in Microsoft Excel, with all the following statistical tests and regression computed in STATA/SE 16.0.

3. Literature review

There is a vast scope of research that contributes to the investigation of the relationship between fluctuations in stock market returns and oil prices. However, the majority of these studies concentrate on stock markets in oil-importing countries, particularly in the U.S. market. Analysis of the relation in oil-exporting countries is, therefore, more limited. In the following section, previous research on the connection between oil price and stock indices within oil-exporting countries will be presented.

Hammoudeh and Li (2005) highlight the benchmark indexes within two oil sensitive countries, Norway, and Mexico. They study how these are influenced by a world index and the oil price, based on daily data between 1986 and 2003. With the utilization of an APT-model, they deduce that changes strongly influence the Norwegian market in the oil price and that the oil market leads OSE with one day. Still, they disclose that investors face greater risk associated with developments in the world index than the oil price. Their results additionally indicate that the Norwegian market moves asymmetrically with the world index, where the Norwegian market seems to be more sensitive to negative fluctuations rather than positive fluctuations.

Driesprong, Jacobsen, and Maat (2008) utilize daily and monthly data from 1973-2003 in order to analyze the relationship between stock indices in 18 countries, together with a world index, with the oil price. The examination reveals that stock markets have a delayed reaction to changes in the oil price. They consider that investors underreact to fluctuations in the oil price, which is in contrast to the market efficiency hypothesis. In the research paper, it appears that a lag of 6 trading days provides the strongest explanatory power for stock markets. Consequently, they argue that a trading strategy can be constructed, where one can utilize the anomaly. Eventually, they specify that the underreaction is lower within oil-related sectors than in other sectors.

Næs, Skjeltorp, and Ødegaard (2009) examine what factors are affecting OSE. Their study is based on monthly returns from 1980-2006. They demonstrate that the majority of the world's stock markets decrease when the oil price surge, while the OSE is positively correlated with the oil price. Further, they identify that changes in oil prices provide significant outcomes in cash flows to most of the companies listed on OSE. However, they emphasize that oil is not a priced risk factor in the Norwegian stock market.

Sørensen (2009) worked with daily, monthly, and quarterly data from 1973-2007, where the world index, indices for the G7 countries together with Norway is employed to examine whether the oil price can be used to predict stock returns. A considerable discovery from his paper is that the oil and gas sector represents the sector that is least predicted by oil. The explanation for this is that the oil price possesses a first-order effect on the oil and gas sector, and is thereby priced in immediately. This will explain why the lag-effect on oil price changes will be greater towards other sectors, where oil has a second-order effect. Exogenous components that cause strong movements in the oil price are isolated. Within these exists components, which not can be explained by macroeconomic terms like OPEC collapses or military conflicts in the Middle East. It is then illustrated that the oil price is worse to predict stock markets when these events are excluded. Further, the study claims that negative fluctuations in the oil price have a greater influence on stock markets than what positive movements have. However, the examination shows that positive and negative changes in the oil price are not significantly distinct from each other.

Bjørnland (2009) demonstrate that a 10 percent surge in oil price will bring an approximate increase of 2.5 percent in stock prices on the Norwegian market. This finding is in accordance with Park and Ratti (2008), who likewise discover that increases in oil prices bring positive effects on the Oslo Stock Exchange. Additionally, they find the opposite effect for oil-importing countries. Moreover, Jung and Park (2011) analyze the significance of reaction to oil supply and demand shocks by stock markets in both an oil-exporting country (Norway) and an oil-importing country (South Korea). The study finds that the response of the respective stock market returns differs considerably from each other. Further, they observe that a larger portion of stock return fluctuation in small open economies like Norway can be explained by the crude oil market, contrary to the U.S. market. According to this investigation, this implies that small open economies have greater oil-dependent technology and limited access to the global financial market.

As highlighted above, previous research has been conducted on the effect oil price have on stock exchanges. However, these researches examine shorter periods and will not able to capture recent changes and compare the results over the decades. Following the studies mentioned above, the oil market has experienced volatile alterations. It is leading to a possible change in the data generating process, not yet investigated. Additionally, the share of the energy sector at the Oslo Stock Exchange has simultaneously fluctuated. This thesis, therefore, provides additional research within this field.

4. Variable selection

The following section will present the variables selected in the analysis, and a discussion of its relevance for the purpose of this thesis. This thesis aims to examine the relationship between the oil price and OSE, as well as its development over time. By only considering the oil price and OSE, we would fail to capture external factors that could affect both variables, as well interactions between factors.

4.1. Oslo Stock Exchange

OSEAX is a market capitalization-weighted index that tracks the stock performance of all shares listed on OSE in its respective sectors. The index is adjusted for corporate actions daily, whereas the current outstanding number of shares is applied in the index. Additionally, the index is adjusted for dividend payments (Oslo Børs, 2020). As of 31/12/2019, the index consisted of 176 stocks, with a total market capitalization of NOK 2.596 Billion (Bloomberg, 2020).

4.1.1. Industry sectors

The industry sectors on the OSE is based on the Global Industry Classification Standard (GICS). Typical for industry indices is their designation with OSExxGI (Oslo Stock Exchange, industry code, the international standard). It is expected that the individual industries on OSE react differently to changes in the oil price and the other selected variables. Industries, where companies operate with extraction or sale of oil, will have increased profits with higher oil prices. Opposite, companies having oil as inputs of production will face increased costs with higher oil prices.

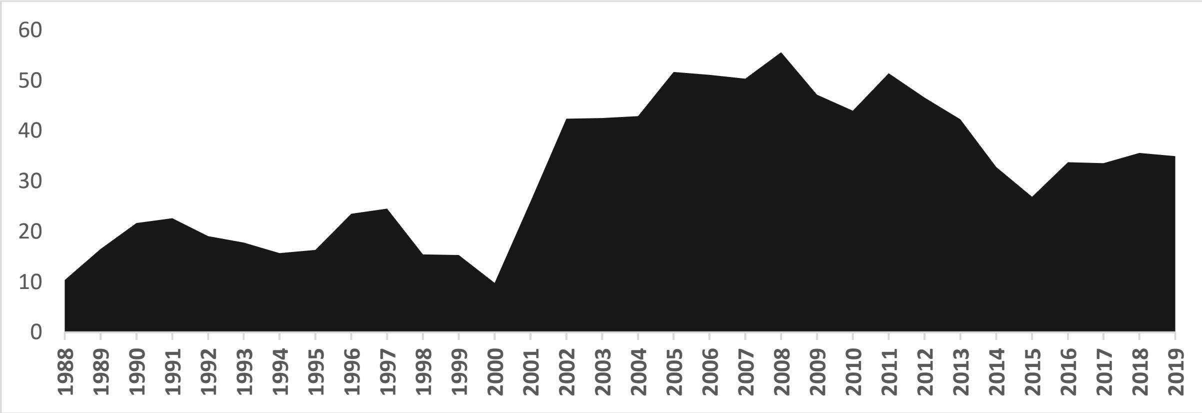


Figure 4.1 – The Energy Sector Market Cap Fraction of OSEAX (Source: Oslo Børs, 2020)

The energy sector OSE10GI, mainly consisting of oil-related companies, have clearly comprised the major industry sector during the last twenty years. Figure 2.1 above presents the development of the energy sector's market capitalization in comparison to OSEAX over the period 1988-2019. Huge fluctuations are detected, where especially the considerable increase from the start of the 2000s is remarkable. This upsurge can primarily be explained by the stock market listing of Statoil in 2001, and their subsequent merger with Norsk Hydro in 2007. Besides, the formation of numbers of oil-service companies, together with substantial growth in the oil price, expanded the size of the energy sector. Following the supply shock in 2014, the size of the energy sector seemingly reduced. At the same time, other sectors experienced growth during this period, particularly the consumer discretionary due to increases in the salmon price. Since then, the oil price and the energy sector have grown steadily, which again has given the industry a fraction above 40% of the total market capitalization at OSE (Oslo Børs, 2018).

4.1.2. Shareholder structure

It is conceivable that the shareholder structure has an impact on how responsive OSE is towards fluctuations in the oil price. Figure 2.2 below highlights that foreign investors increased their investments during the 2000s when the oil price was strong. Then, during the financial crisis, many foreign investors escaped from the OSE together with the decreasing oil price. From 2008 and onwards, overseas investors have gradually increased their positions in the Norwegian stock market. From the figure, it is visible that the oil price collapse in 2014 did not affect these investor's appetite for Norwegian stocks. Government and municipalities ownership, on the other hand, has fallen to its lowest level since before the financial crisis. This is mainly due to the fact that they have generally not changed their positions during the last decade (Oslo Børs, 2019). Their largest holding has been Eqionor (formerly Statoil), whose stock price together with other oil-related companies, has performed worse than other sectors at OSE in the last years.

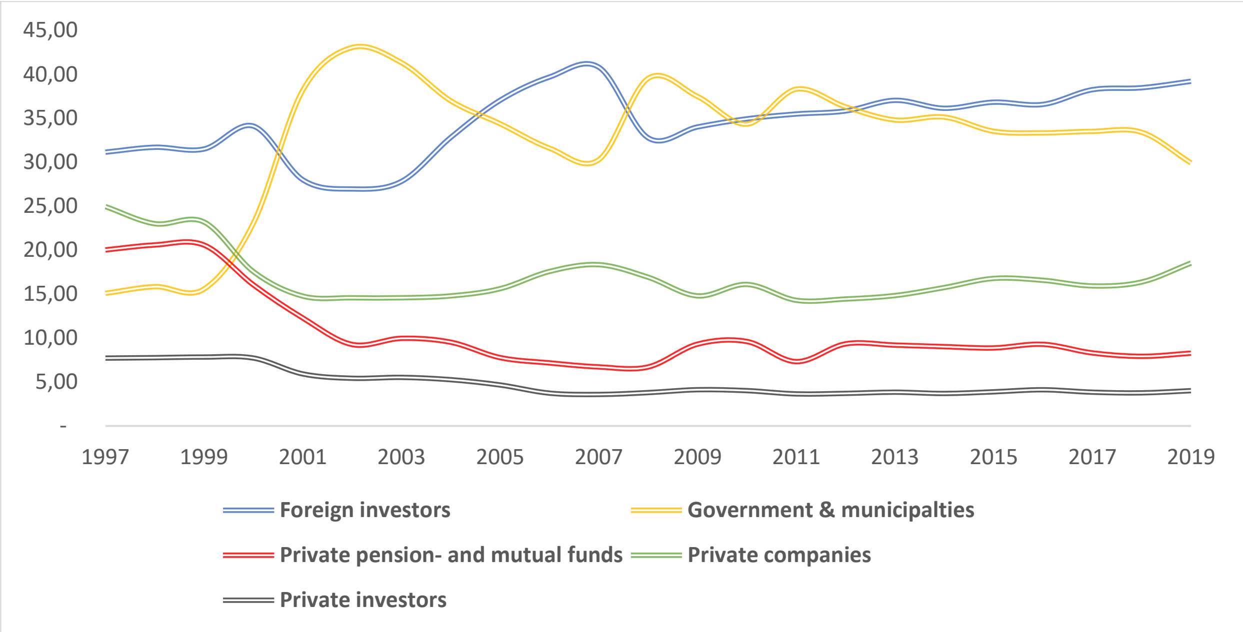


Figure 4.2 – Ownership structure OSE (Source: Oslo Børs, 2020)

4.2. Oil price

There are several petroleum products traded on commodity exchanges around the world. The two most popular traded grades include the Brent North Sea Crude (Brent Crude) and West Texas Intermediate (WTI). Brent crude attributes to oil originated from fields in the North Sea between the Shetland Islands and Norway, where WTI is sourced from the U.S., mainly in Texas, Louisiana, and North Dakota. WTI serves as the benchmark crude for North America and is traded on the New York Mercantile Exchange. Brent Crude, on the other hand, constitute the benchmark for African, European and Middle Eastern crude oil. Hence, the pricing mechanism for Brent crude precepts the value of approximately two-thirds of the crude oil production in the world (Hecht, 2020). Even though there exist regional differences in oil prices, there is found a great degree of covariation among them. Studies utilizing different benchmarks for oil prices in order to estimate stock returns have been completed without finding significant differences between the different oil prices (Driesprong, Jacobsen, & Maat, 2008). As this thesis focuses on the Norwegian stock market, it seems favorable to employ Brent Crude in the model as an independent variable.

4.3. Interest rate

NIBOR (Norwegian Interbank Offer Rate) refers to the collective term for Norwegian money market rates. NIBOR is destined to reflect the interest level of a lender demand for unsecured money market loans in Norwegian kroner. The money market rates are determined and published with maturities of one week, one month, two months, three months, and six months (Finans Norge, 2017). Based on a simple average of interest rates suggested by the NIBOR panel banks, the NIBOR interest rates are calculated upon what rates the banks charge from other banks who are active in the Norwegian money and foreign exchange market.

Three-month NIBOR rate is relevant to apply in this thesis, as this is frequently utilized as the risk-free rate (Morningstar, 2019). A higher interest rate will, according to the dividend model increase the required return for investors, and hence make stocks less attractive. Additionally, higher interest rates make companies' loans more expensive, which leads to a lower degree of investment. All things being equal, an interest rate increase will decrease the value of future cash flow and dividends, and therefore the stock prices will decline (Opland, 2017). Accordingly, the three-month NIBOR is included in the model as an independent variable.

4.4. USD/NOK

Norwegian companies are exposed to the U.S. dollar through several channels, like interest rates, savings, and earnings from goods and services sold abroad. Norway does not have its own money market, as there is a lack of float between banks in its small economy. Therefore, the Norwegian money market is linked to international money markets that are active in the currency swap market. In this way, increased revenue and liquidity is ensured. The dollar exchange rate consequently obtains a direct influence on the Norwegian interest rates (Bernhardsen, Kloster, & Syrstad, 2012).

Further, a great portion of savings within the Norwegian business industry takes place abroad, and especially in U.S. dollars (USD). Changes in this currency, therefore highly influence returns on financial investments placed in foreign countries, and thereby also influences the Norwegian economy. Another important factor is that the oil price is quoted in USD, which strengthens the effect on the oil-heavy Norwegian economy (Bernhardsen & Røisland, 2000). As oil companies receive most of their income in USD, they will need to change this back into Norwegian kroner (NOK). For exporting Norwegian companies, a depreciation of the NOK will serve as beneficial, as they receive more NOK for their products and services. For importing companies, the opposite effect will appear. As Norway is a net exporter, it is expected that the cross-currency between USD and NOK will have a positive impact on OSE. The currency pair will, therefore, be included in the model as an independent variable.

4.5. S&P 500

Standard and Poor 500 Index (S&P 500) represents an index covering the 500 largest American companies listed on the New York Stock Exchange (NYSE) or the National Association of Securities Dealers Automated Quotations (NASDAQ), weighted by float-adjusted market capitalization. As a stock index, the S&P 500 is regarded as the benchmark of the overall U.S market (Amadeo, 2020). With the United States as the world's largest economy, its leading index will consistently affect OSE as a small open economy. The S&P 500 is, therefore, included as an independent variable in the model to capture the international financial market.

4.6. London Stock Exchange

The Financial Times Stock Exchange (FTSE) All-Share Index follows the prices of companies listed on the London Stock Exchange's (LSE) predominant market. This index is an aggregation of FTSE 100, FTSE 250, and FTSE Small Cap Indices, where companies are screened to comply with minimum liquidity and size requirements. Further, the FTSE All-Share Index represent nearly 98 percent of the market capitalization of all listed shares in the United Kingdom. Hence, it is considered to be the most extensive price performance measure for the British equity market (FTSE Russell, 2020). LSE makes up the second-largest stock exchange in Europe, only behind Euronext, measured in market capitalization (Statista, 2019). Due to the considerably shorter lifetime of Euronext, the FTSE All-Share Index is chosen to serve as an indicator for the European economy.

Additionally, U.K represents the country that imports the largest part of the oil extracted on the Norwegian continental shelf (SSB, 2020). The country's economic outlook is, therefore, of great importance for the Norwegian economy. Based on these facts, the FTSE All-Share index will comprise an independent variable in the model to measure the effect of the European financial market.

4.7. Renewable Energy

During the last decade, investing in the renewable energy sector has gained considerable traction. From 2010 to 2019, the renewable capacity has quadrupled from 414 GW to about 1,650 GW (UNEP, 2019). Simultaneously, the decade has seen $2.6 trillion invested in renewables, where the two biggest contributions have been made in solar and wind systems. This development is caused by increases in both environmentally conscious consumers and eco-friendly government policies, together with expanded attention within the capital markets (Myrseth, 2020). The trend towards "green" investments have also taken place at OSE, where companies operating within renewable energy have seen substantial growth. In order to capture and investigate the effect of this movement, the Renewable Energy Industrial Index – World (RENIXX) is included in the model. RENIXX represents the world's first global index for the renewable energy industry and consists of 30 stocks with the highest free-float market capitalizations (IWR, 2020). RENIXX is handy for the purpose of this thesis, as it allows for daily observations.

4.8. Difference in trading time

A discussion should also be held regarding the difference in trading time and location across variables. Since the variables are given in different time zones, there are times where the variables are not open for trading simultaneously. The S&P 500 is listed in New York Stock exchange, Brent crude oil is listed on International Commodity Exchange (ICE) in London, OSEAX is listed on Oslo Stock exchange, FTSEAX is listed on the London Stock exchange, and

RENIXX is listed on the Swiss exchange. Since we are using closing prices for all variables, not all information will be captured the same day if a market closes earlier than other markets. All variables will thus have different opening hours and can make them challenging to compare. Below all variables are presented with the opening hours given in Greenwich Mean Time Zone (GMT).

- Oslo Stock exchange closing time – 15.20 GMT
- ICE Closing time – 17.30 GMT
- New York Stock exchange closing time – 21.00 GMT
- London Stock exchange closing time – 16.30 GMT
- Swiss Stock Exchange closing time – 16.00 GMT

As we are comparing the end of day returns on OSE with stock exchanges across the world closing later on the day, the returns on the OSE will only capture what happens before 16:20 GMT+1. If there should occur significant changes in the American market, this would not be reflected in the Norwegian stock market until the day after, when OSE opens again. This could, in theory, mean that the other exchanges are leading OSEAX by one day, due to the various trading times. For all exchanges except S&P 500, the difference in closing time is so little that, in reality, the effect would be modest. To examine the effect of time difference on the S&P 500, it is possible to adjust returns with a zero- and one-day lag. The results on correlation with OSEAX show a much higher correlation when using zero lag, and thus there is reason to believe that most of the changes are captured the same day, and no adjustments are made.

‘

5. Energy Market

As the energy market is a significant part of this thesis, a thorough review of the energy market and its dynamics will be presented. For the purpose of this thesis, and the variables included, the section will be split in two. One part will take a close look at the oil market and what drives the price. The other segment will dive into the renewable sector and look at how the industry has grown over recent years.

5.1. Oil Market

Oil is regarded as one of the most crucial commodities in the world, and its price is like any other commodity price driven by supply and demand. On the supply side, the usual distinction is divided between the Organization of the Petroleum Exporting Countries (OPEC) and non-OPEC countries. On the contrary, the demand side is split between the Organization of Economic Co-operation and Development (OECD) countries and non-OECD. Additionally, the level of global oil storages and dynamics in the world's financial market has a considerable impact on the price of oil (EIA, 2020a).

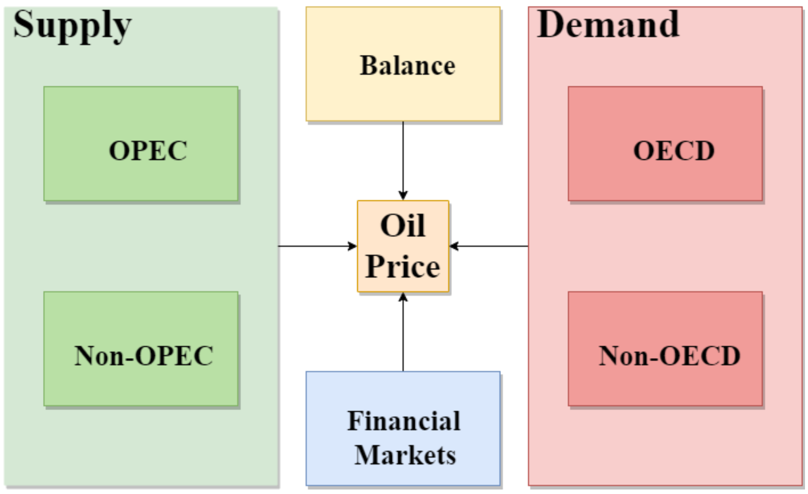


Figure 5.1: Oil price drivers (Source: EIA,2020)

5.1.1. Supply

OPEC countries stand for about 40 percent of today's oil production globally. With this market share, the production from OPEC possesses the ability to affect the oil price highly. OPEC seeks to manage oil production between its member countries by deciding upon production targets in order to optimize the balance volume and price (EIA, 2020). Its largest and most influential member is Saudi Arabia, which is by itself able to affect oil prices with its production. Historically, they have also had the largest spare capacity, having 1.5-2 million barrels per day of spare capacity on standby for price management (Ibid.). Moreover, crude oil prices tend to surge during times when OPEC countries agree to reduce their production targets. On the other hand, conflicts of interest also occur within OPEC countries, where disagreements over predetermined cuts typically lead to lower oil prices (Siripurapu & Chatzky, 2020).

Non-OPEC countries stand for the remainder 60 percent of production, with central production regions being in North America, Russia, and the North Sea. While OPEC countries are mostly government-controlled with national oil companies, non-OPEC countries are generally investor-owned companies. Thus, their objective is first to maximize shareholder's value and make investment decisions solely on economic factors. As a result, Non-OPEC countries have a more active and rapid response to oil operations and investments in order to react more quickly to changes in market conditions (EIA, 2020). Norway constitutes a particular exception in this manner, as the Norwegian government owns 67 percent of the shares in the country's largest oil-producing company, Equinor (Hovland & Lorentzen, 2019)

Country	**Million barrels per day**	**Share of world total**
United States	17.94	17.78%
Saudi Arabia	12.42	12.31%
Russia	11.40	11.30%
Canada	5.38	5.33%
China	4.81	4.77%
Iraq	4.62	4.58%
Iran	4.46	4.42%
UnitedArab Emirates	3.79	3.76%
Brazil	3.43	3.40%
Kuwait	2.91	2.88%

Total top 10	**71.15**	**70.52%**
Norway	**1.86**	**1.85%**
World Total	**100.89**	

Table 5.1: 10 largest oil producers and Norway, 2018 numbers (Source: EIA,2020)

While OPEC countries actively seek to affect oil prices, non-OPEC countries are generally viewed as price takers, where they are not attempting to influence market prices with setting production targets. This leads non-OPEC to produce at or near-maximum capacity usually, and thus rarely have spare capacity. All else being equal, lower levels of supply from non-OPEC countries will push oil prices upwards as global supply decreases. In parallel, this drives the oil price more sensitive towards OPEC. This circumstance is known as "call on OPEC". An increase in the call on OPEC increases the OPEC country's ability to influence oil prices (EIA, 2020c). Moreover, not only the current supply from non-OPEC will have an impact on the prices. Expectations of future non-OPEC production also affect current oil prices, as it forms an indicator of what influence OPEC will have in the coming periods.

5.1.2. Demand

Moving over to the demand side, the separation is made between OECD and non-OECD countries. OECD is an association that consists of 35 countries, where the majority is from Europe, four is from the Americas (United States, Canada, Mexico, and Chile), and four Pacific members (Australia, New Zealand, South Korea, and Japan) (Amadeo, 2020). The OECD countries thus represent advanced economies, with developed infrastructure and high vehicle ownership per capita. Consequently, the major portion of oil demand from the OECD countries is related to transport. There are few subsidies on end-use prices, which generally cause fluctuations in market prices of oil to be reflected in consumer prices (EIA, 2020). Sustained high or expectations of higher oil prices thus lead more consumers to consider more fuel-efficient vehicles or other transportation modes. In this way, oil price increases will often coincide with lower demand from OECD countries.

During the recent decade, the demand from OECD countries has seen a decreasing trend. In 2008, these countries composed 55.6 percent of global demand, while in 2018, this share was dropped to 47.5 (BP, 2019). This was a continuous movement, as it declined from 2000-2008 as well.

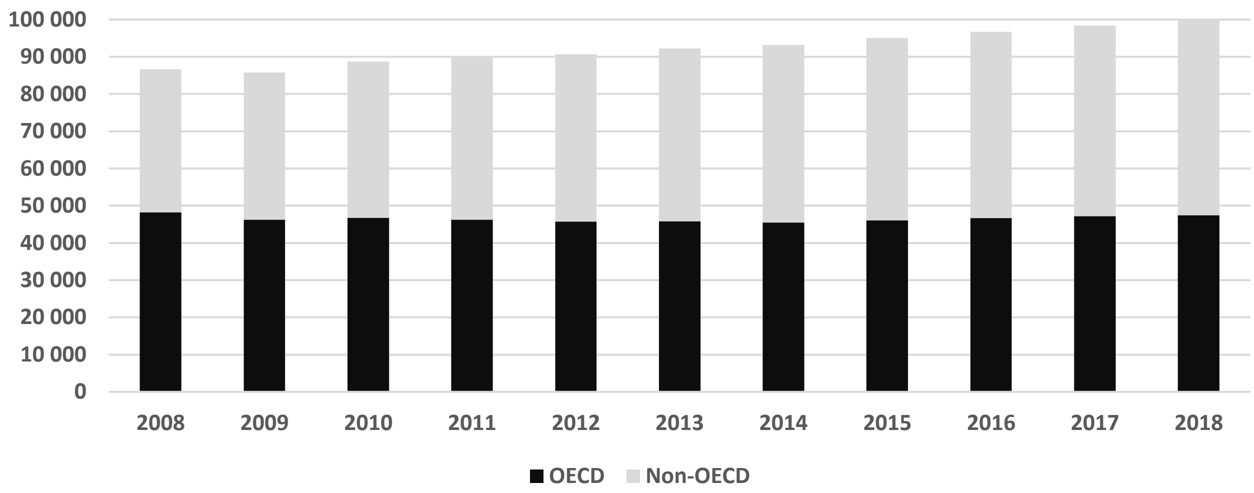


Figure 5.2: Daily demand for crude oil over recent years (Source: IEA, 2020)

On the other hand, non-OECD countries represent developing countries. Rising economic growth has been reflected in growing oil consumption by these countries. In addition to the oil used for transportation activities, non-OECD countries have a large share of manufacturing processes that uses oil as the primary energy source (EIA, 2020). When it comes to power generation, many of these countries also rely on oil. In contrast to OECD countries, many developing countries subsidize end-user prices, which suppress the effect of fluctuating oil prices for consumers. Together with experienced rapid growth in population, economic growth within developing countries has coincided with increased global oil demand (Ibid.).

While demand from OECD countries has diminished over the previous decades, demand from non-OECD has increased strongly. Between 2000 and 2010, oil consumption underwent a growth of 40 percent. Moreover, in 2018, these countries contributed 52.5 percent of global oil demand (BP, 2019). The global rise in demand for oil seen over prior periods is therefore directed towards the growth from non-OECD countries.

5.1.3. Oil Balance and financial markets

Despite OPEC's targeted production goals, global oil consumption and supply are generally hard to predict. Consequently, supply regularly does not match demand, by which oil inventories act as a balancing measure between the two. In order to avoid an overwhelming flood of oil in the market in periods when production exceeds consumption, oil can be stored in inventories and thus be held back for future use. The commodity can be stored as crude oil or finalized products like gasoline, heating oil, and diesel (EIA, 2020). Finalized products are often exposed to seasonal variances in demand, whereas inventories are being accumulated in periods of low demand, and opposite drew upon in high demand seasons (Ibid.). Oil inventories are also held by several countries as a strategic reserve. If an unforeseen event should appear, disrupting oil production, these countries can then use their oil reserves as an emergency response (Ibid.).

Financial markets represent the final force that affects the oil price. Oil is not just traded as physical products for delivery from producer to consumer, but also as contracts for future delivery or other derivatives. These trading contracts reveal expectations of future oil prices movements, and therefore also play an important role in affecting current oil prices (EIA, 2020)

Industries heavily exposed to oil and oil products such as oil producers and airlines, often seek to hedge themselves towards undesired price changes. Other participants that have no interest in the commodity itself are also active within commodity derivatives. Banks, hedge funds, traders, speculators, and arbitrageurs all try to profit from these derivatives. During recent decades, the interest in oil as an investment has increased substantially. In addition to profit-making and hedging purposes, these derivatives have progressively been traded for diversification objectives (Ibid.). Accordingly, the development in commodity trading has raised the impact of the financial markets on the determination of oil prices.

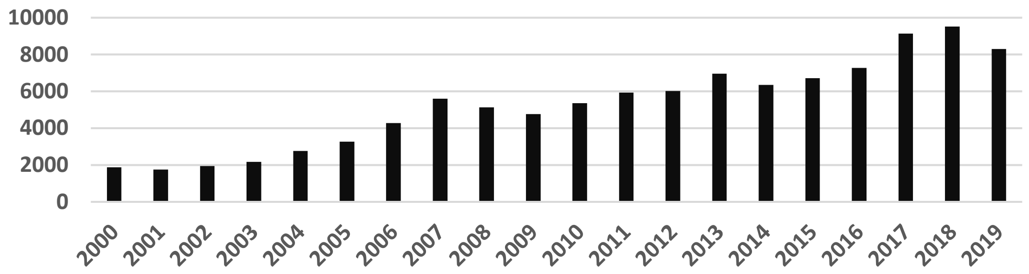


Figure 5.3: Numbers of crude oil future contracts annual average (Source: EIA 2020f)

5.1.4. Development of oil price and OSEAX

To get a better understanding of what drives the oil price, it is relevant to take a closer look at its development over recent decades. This will then be compared to the development of the Oslo all share index over the same period.

As the previous section seeks to explain, supply and demand are the main drivers for the oil price, along with oil balance (that is a consequence of consumption and production) and financial markets. These factors can, in turn, often be affected by exogenous factors at various significance, able to cause shocks in either supply or demand. Looking back at historical prices, some main events have had a massive impact on oil prices.

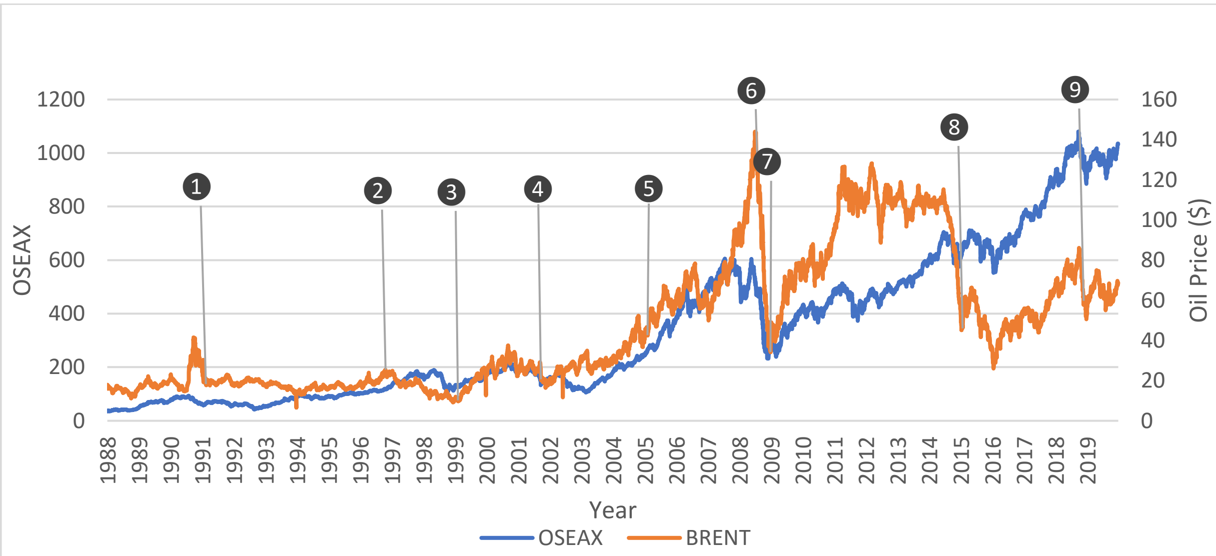


Figure 5.4: Historically price for OSEAX and Brent Crude with major events with a high impact on the oil price (Source: Bloomberg,2020; EIA,2020)

For the first decade, the oil price was steady, between 20-25 dollars, with no significant impacts. We can see that the Gulf war had some effect, but the oil price normalized afterward. Oil price was mostly affected by geopolitical factors and war in the middle east during this time, and thus affecting production and export from the middle east (EIA, 2020).

Point and year	**Event**
1: 1991	Gulf War
2: 1997	Asian financial crisis
3: 1999	OPEC cuts production targets
4: 2001	9/11 attack
5: 2005	Low spare capacity
6: 2008	Global financial crisis
7: 2009	OPEC cuts production targets
8: 2015	OPEC production quota unchanged
9: 2018	OPEC cuts production

Table 5.2: Events with a major effect on the oil price. (Source: EIA, 2020)

Oil price experienced a considerable increase from 2001, after the 9/11 terrorist attack, all the way towards the financial crisis that occurred in 2008. Much of this increase can be explained by new economies growing rapidly, such as China and India, creating a higher demand and increased economic activity globally. Following the downturn of the financial crisis, OPEC used its power by announcing cuts. Oil prices quickly recovered over the coming years, although behaving more volatile than previously.

During 2014, many of the previous emerging countries with high growth started to slow down on their growth rate. This slowly and steadily created an oversupply of oil. This then led to a substantial drop in oil price end of 2014 (Depersio, 2020). In more recent years after the drop, the oil price has experienced higher fluctuations for several reasons. There has been more uncertainty with the Trump administration, wars, trade wars, and political pressure to reduce emissions. Following this, it has been hard to predict the oil price, and it has not been on a steady level, but drifted between an interval ranging from 40-80 dollars.

Taking a closer look at the Oslo all shares index, it has followed the oil price to a high degree towards the financial crisis. Looking at point 6 in figure 5.4, the oil price started to exceed beyond OSEAX, before taking a downturn due to a demand shock. After the financial crisis, OSEAX has followed a positive trend, slowly increasing. It reached pre-financial crisis values already in 2015 before it accelerated in 2016 and went to an all-time high week after week towards the summer of 2018. During this same period, the oil price experienced a much more turbulent period, with higher volatility and a weaker link with OSEAX.

5.2. Renewable Energy

During the last decade, the renewable energy sector has gained considerable traction. Among the factors behind this prevail the increasing threat of exhaustion of oil reserves and the significant level of greenhouse gas emissions in the atmosphere. Carbon generated from the production of non-renewable energy has further been perceived as the dominant cause of the environmental crisis facing humanity (Singh, Nyuur, & Richmond, 2019). Consequently, there has been growing momentum in the development of clean energy sources, with an increasing number of developed and developing countries introducing green growth agendas. Following the Conference of the Parties 21 (COP21) Paris Agreement back in 2015, more than 170 countries agreed to try reducing global warming below two degrees Celsius, representing an effort that would demand considerable investments in renewable energy sources (Pickl, 2019).

From 2010 to 2019, global renewable capacity nearly quadrupled its figure, from 414 GW to 1,650 GW. The biggest contributors to this growth are represented by solar and wind farms, where global solar power capacity has seen an increase of 2,500% during the decade (Mathis, 2019). In terms of total added power capacity, renewable energy outpaced fossil fuels in 2012, and have maintained the dominance since. Moreover, there were more decommissioned fossil fuel power plants than built in Europe and the US in 2019 (UNEP, 2019). These developments have further increased the threat of an approaching stagnation in the oil demand, where EIA (2019) anticipates a downturn in consumption growth after 2025.

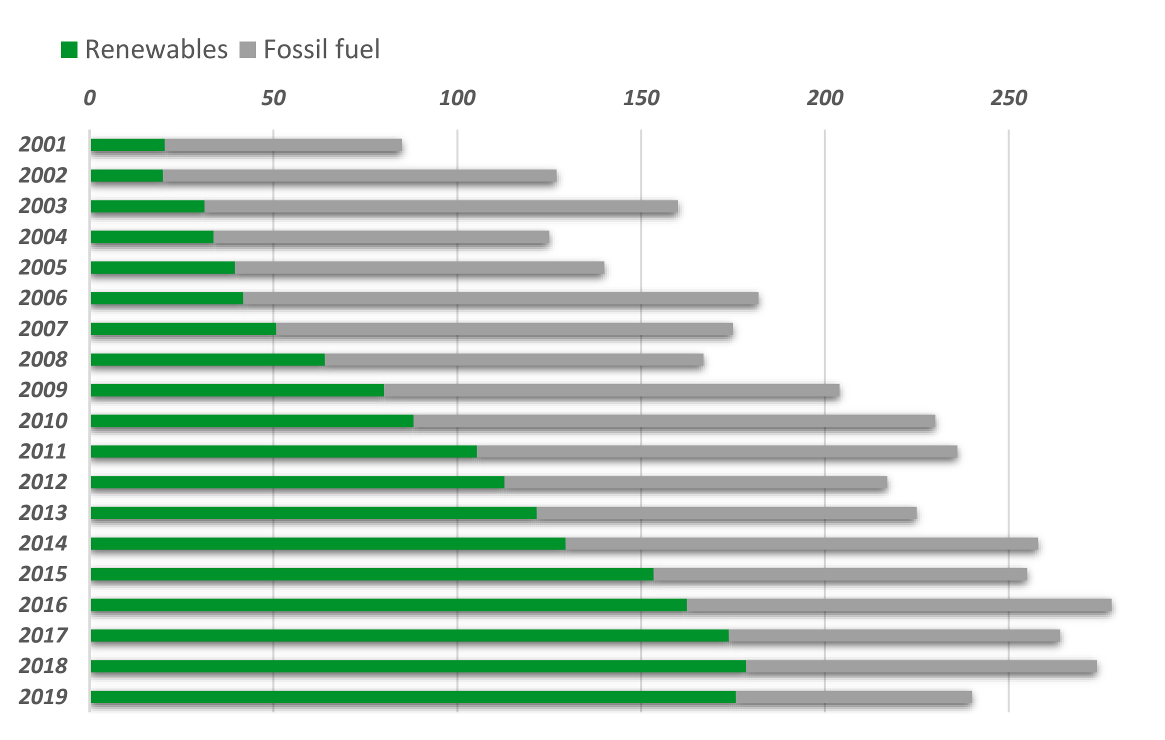


Figure 5.5 – New global electricity capacity added each year in GW (Source: Irena)

In addition to the policy initiatives that have served as a beneficial contribution to the renewable energy industry, decreasing costs have made renewables more competitive compared to fossil fuel-fired power plants (REN21, 2019). Measured by the Levelized cost of electricity (LCOE), which compiles the costs of development, construction, operations, finance, and maintenance from projects around the globe, the benchmark for solar has dropped by 81 percent from 2009. For wind, the comparable number have been a 46 percent decline for onshore, and a 44 percent fall for offshore wind (BloombergNEF, 2019). These considerable decreases in costs have made them more economical to build than new fossil fuel plants in two-thirds of the world, whereas some places are even more cost-effective than to continue operating existing fossil plants.

Behind these moves in the cost of renewable energy sources, there are several factors. There has been a considerable increase in auctions for renewable energy projects, which have risen competition and lowered overhead costs. Further, economies of scale and improved technology have also been in favor of renewable energy reducing the cost (Khashman, 2019). A strong downward trend in the cost of capital related to these projects has also contributed to the lowered LCOE. This is due to the significant upfront capital expenditure compared to the low operational costs, especially in the cases of solar and wind projects. The low-interest-rate environment during the 2010s and rising competition among investors and banks to engage in renewable energy projects have both committed to reduced cost of capital (BloombergNEF, 2020). Moreover, major investors have, over the decade, increasingly moved away from fossil fuel investments and towards renewable energy, wherein total the period has seen $2.6 trillion invested in renewables.

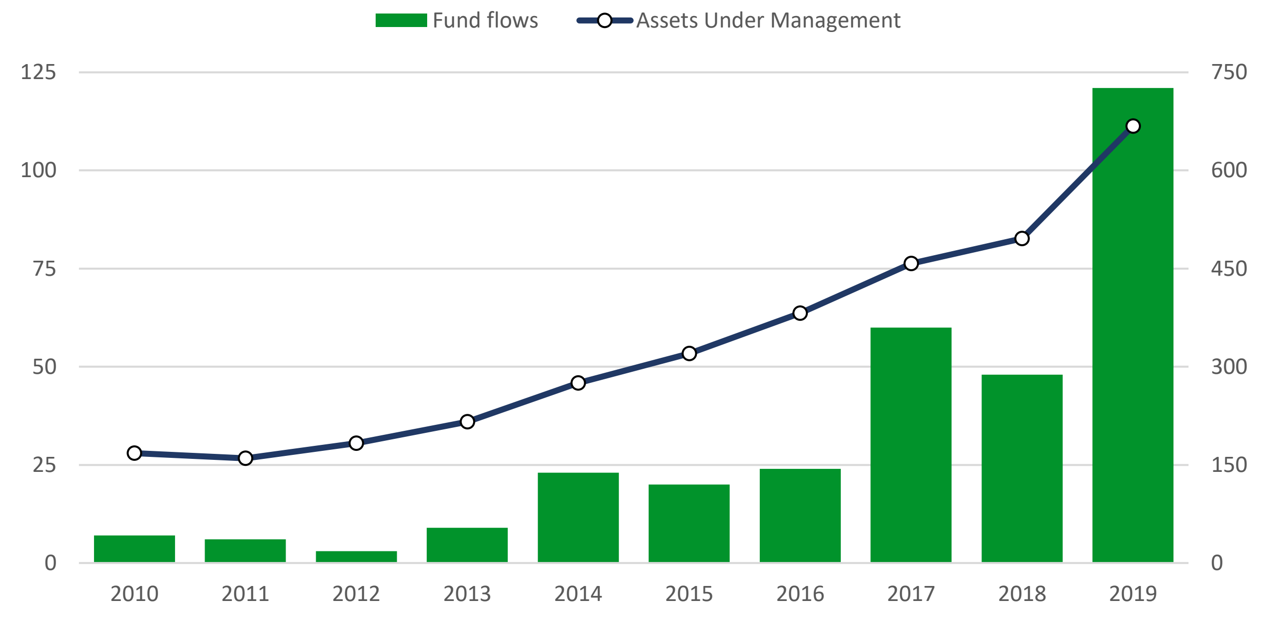


Figure 5.6 – Annual European fund flows (LHS) and assets under management (RHS) (Source: Morningstar)

The trend towards "green" investments have also taken place at OSE, where companies operating within renewable energy have seen substantial growth, especially in the last couple of years. The independent solar power producer Scatec Solar, the hydrogen company Nel and the wind power exposed company Bonheur are all examples of renewable energy stocks that have experienced considerable increases in both interest and returns (Nilsen, 2019). This has been part of the growing trend in the Norwegian equity market, where fossil fuel extracting companies have seen lowered interest. For Norwegian listed oil companies, these moves have placed them in a complicated position. However, back in 2018, the oil major Statoil changed its name to Equinor in a shift of strategy from a focused oil company into a broad energy company (Equinor, 2018).

6. Theoretical framework

This section will cover the theory that is relevant for understanding the methods used. The theoretical framework will be presented as two folded. The first section will cover more general understanding and theories relevant when using financial data as well as understanding how assets can be valued and what drives the price. The second part considers the econometric theory that will be applied when conducting the analysis. It will also cover important concepts when working with financial data, that are applied to the arguments for decisions made.

6.1. Financial Theory

6.1.1. Pricing of a stock

There are several ways to value stocks with fundamental analysis, but the most commonly used is the discounted cash flow approach. This method relies on discounting future cash flows to a present value, using a discount factor that reflects the relevant risk of the cash flows, together with the time value of money (Perersen & Plenborg, 2012). This implies that the present value of a stock exchange can be estimated by the excepted future cash flows of all listed stocks, discounted by an appropriate rate to reflect the associated risk. According to Petersen and Plenborg (2012), the following relation can be used to estimate the value of a stock:

$$Market\ Value\ Equity = \sum_{t=1}^{\infty} \frac{FCFE_t}{(1 + r_e)^t} \tag{1}$$

Where $FCFE_t$ is the free cash flow to equity for period t, and r_e is the required rate of return for investors. A change in either expected cash flow or required rate of returns would directly affect stock prices.

6.1.2. Market efficiency

Throughout years, there has been a vast amount of studies on developments in stock prices. A study by (Kendall & Hill, 1953) concluded that stock prices follow a random walk process. This postulate has later gained general acceptance, as it emphasizes that since new information is unpredictable, stock returns should accordingly behave uncertainly. Following the research by Fama (1970), the concept has been referred to as the efficient market hypothesis (EMH). The EHM states that stock prices fully reflect currently available information so that only new information can lead to changes in stock prices. Hence, an efficient market is characterized as a market in which prices consistently fully reflect accessible information (Fama, 1970).

It has become common to distinguish between three degrees of market efficiency, which is a weak, semi-strong, and strong degree (Bodie, Kane, & Marcus, 2014). Weak market efficiency assumes that stock prices reflect all historical information in the market. This includes past price movements, earnings data, and trade-volume, that is publicly available to all investors. Any technical analysis would, therefore, yield now excess return, as they are based on available information and already priced into stock prices.

A Semi-strong degree of market efficiency assumes that stock prices reflect both past and newly available information. This includes company announcements, quarterly reports, analyst estimates, and other information released. Further, it states that new information is immediately priced into stock prices. As a result, neither fundamental analysis or technical analysis would help an investor to gain excess return relative to the market (Berk and DeMarzo, 2014). In order to achieve a risk-adjusted excess return, one should possess inside information.

Finally, the strong degree of market efficiency supposes that security prices reflect all information. This contains historical information, new information, and also private or insider information. Hence, investors have access to all conceivable determinants that could affect stock prices. Consequently, this implies that no investors will be able to outperform the market over time (Berk and DeMarzo, 2014). This type of degree represents an extreme case, as it indicates that there should be no incentive for investors to manage a portfolio actively.

Moreover, there exists legislation and strict regulations against insider trading. A paradox affiliated with this hypothesis is that if all investors believe that the market is efficient, then all would prefer a passive investment strategy, in which trade would eventually stop. New, unforeseeable information would then not be captured and reflected in stock prices, and thus end up in an inefficient market (Grossman & Stiglitz, 1980). For a market to be efficient, there must be some investors that believe that the market is inefficient, which persistently captures new information and adjusts prices.

EMH has accordingly faced some critique, as the mentioned Grossmann-Stiglitz paradox. Another instance is the "Size-effect" anomaly by Fama and French (1996), where smaller companies tend to outperform larger companies on a risk-adjusted basis. Other anomalies consider momentum strategies, where one would buy winner stocks and sell losers due to a belief of positive correlation from the previous period (Jegadeesh and Titman, 1993). Moreover, the assumptions behind EHM have raised doubt whether they are valid in real life. For instance, EHM assumes that all investors will react in the same manner under new information. This assumption faces opposition in empirical studies showing that markets over- and under-react to new information (Malkiel, 2003).

6.1.3. The Capital Asset Pricing Model

The relationship between risk and return is an essential element of financial theory. In order to elucidate this connection, the Capital Asset Pricing Model (CAPM) is employed. CAPM is based on Markowitz's mean-variance portfolio theory, where a risk-averse investor with a one-period time horizon only considers expected return and standard deviation (Berk and DeMarzo, 2017). The model takes the following form:

$$E(r_i) = r_f + \beta_i(r_M - r_f) \tag{2}$$

- r_i = return on security i
- r_f = risk-free rate
- β_i = systematic risk factor for security i

- $\left(r_M - r_f\right)$ = market risk premium

The systematic risk factor indicates the amount of market risk incurred concerning the market by investing in a specific security. In that respect, the model describes the movement in a security, by the move in the underlying market. Moreover, the beta coefficient is defined as the covariance between the respective security and the market, divided by the variance of the market:

$$\beta_i = \frac{Cov(r_i, r_M)}{Var\ (r_M)} = \frac{\sigma_i \sigma_M \rho_{i,M}}{\sigma_M^2} \tag{3}$$

The variance and correlation with the market for individual security will thus determine the load of market risk associated with the investment. Moreover, it will decide the expected return. This can be illustrated more clearly by considering CAPM in the following form:

$$E(r_i) - r_f = \beta_i\left(r_M - r_f\right) \tag{4}$$

The expected risk premium for the security $E(r_i) - r_f$ reflects the exposure to the market risk $\beta_i\left(r_M - r_f\right)$. In case an investor does not expose himself against market risk, $\beta_i = 0$, then the expected return will equal the risk-free rate. However, CAPM does not address unsystematic risk, as it assumes that all investors possess a diversified portfolio that eliminates risk related to investments in singular assets. In this sense, CAPM affirms that the only way to increase expected return is to raise the exposure towards market risk. This assumption, among others, as the premise of no tax or transaction costs, has raised doubt whether the model is accurate in financial markets (Bodie et al., 2014). Nevertheless, CAPM holds a strong position in order to describe the relationship between risk and return. Additionally, the model forms the basis of other financial models, as the arbitrage pricing theory.

6.1.4. Arbitrage pricing theory

Arbitrage pricing theory (APT) represents a multifactor model used to describe the relationship between risk and expected return for financial security. The model was introduced by Ross (1976), and serve as an expansion of the CAPM model. With APT, the objective is to estimate the return of a security based on its sensitivity against different macro and financial factors (Bodie et al., 2014). The model can be defined as follows:

$$r_i = E(R_i) + \beta_{i1}F_1 + \beta_{i2} + \cdots + \beta_{in} + F_n + \varepsilon_i \tag{5}$$

where the return of the security is affected by the expected return $E(R_i)$, the sensitivity to risk factor F measured by β_i as well as the idiosyncratic risk ε_i.

The APT model intends to identify the correct price for financial security. In case the price deviates from the model's estimated price, and thereby appear as over or underpriced, market participants will react and converge the price back to the "fair price" by exploiting arbitrage. APT allows investors to examine whether a security is exposed to certain factors, and in that way, makes it more accessible to construct portfolios with desired exposures. In contrast to the one-factor model CAPM, this model will be a function of a complex set of macro factors, and will accordingly form a more precise and realistic model. On the other hand, the model will be more demanding to utilize since it requires additional analysis. For instance, it is more challenging to identify the factors that possibly affect the given security, together with how sensitive it is towards changes in the respective factors. By employing such models, one must consider a balancing between precision and complexity. A simple model will serve as more applicable and coherent but may fall short when it comes to estimating ability.

Similar to the CAMP model, APT estimates the factor specific beta coefficients from a linear regression of historical returns from the corresponding factors. Moreover, in contrast to CAPM, this model does not state which factors or how many that are appropriate, but declares that both class and number factors will differ across different markets and over time.

6.2. Econometric theory

In the following section, the econometric framework for the analysis will be presented. First, an introduction of regression analysis will be outlined, followed by how statistical inferences can be drawn based on a sample. Certain characteristics are necessary for time series variables to be included in a regression model. These characteristics will, therefore, be explained, with a focus on the assumptions underlying the multiple regression model and stationarity.

6.2.1. Regression analysis

With regression analysis, the purpose is to explain changes in one variable based on changes in one or several other variables. The variable to be explained is called the dependent variable, while one of the variables used to explain this is named independent or explanatory variables (Brooks, 2014). Linear regression will be the focus of attention for this thesis, where the aim is to analyze the linear relation between variable y and several independent variables of x. y then becomes a linear function of x, and is in its simplest form presented as:

$$y_t = \alpha + \beta x_t + u_t \tag{6}$$

Where α is the intercept, β represents the slope or the relation between x and y, and u serve as a random disturbance term. The dependent variable and the independent variable is treated differently in a regression. The y variable is assumed to be a random variable with an associated probability distribution, while the x variables are treated as fixed or non-stochastic. Additionally, the disturbance term u is included to capture deviations in y, which the explanatory variables fail to capture (Brooks, 2014).

6.2.2. Ordinary Least Squares (OLS)

The ordinary least squares (OLS) estimator select the regression coefficients, which generates the regression line as close as possible to the observed data. Closeness in this regard is determined by the estimated values of α and β which minimize the squared distance between the model's estimated $\hat{y}$ and observed y (Stock and Watson, 2015). Hence, the values $\hat{\alpha}$ and $\hat{\beta}$

represent estimates for the true values of α and β. $\hat{u}_t$ express the difference between the actual observed value of Y and the value estimated by the model $\hat{Y}$.

Minimizing the sum of the squared residuals, rather than just the sum of the vertical residuals, is done to consider the absolute values of the errors. This provides an exact solution as opposed to minimization of total vertical residuals, where errors above and under the regression line would cancel each other out (Brooks, 2014). The sum of the squared errors is described as the residual sum of squares (RSS). Figure 5.1 below illustrates the relationship between the actual value y and the estimated $\hat{y}$, as well as the observable error term $\hat{u}$.

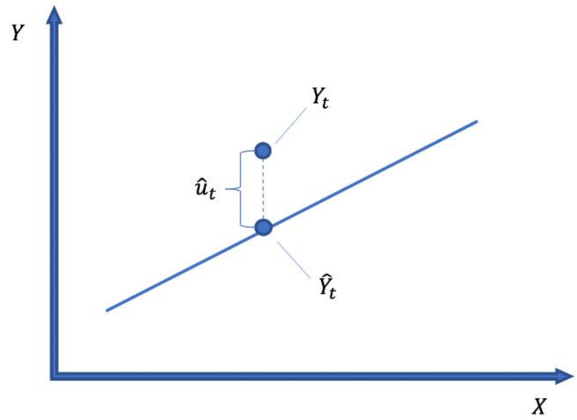


Figure 5.1 – Relationship between observed value, estimated value, and the error term

OLS selects the estimates of α and β that minimize RSS, which in turn is determined by $\sum \hat{u}_t^2$. From the above, one has that $\hat{u}_t = (y_t - \hat{y}_t)$ and $\hat{y}_t = \hat{\alpha} + \hat{\beta} x_t$. Therefore, the OLS minimizes:

$$RSS = \sum_{t=1}^{T} \left(y_t - \hat{\alpha} - \hat{\beta} x_t\right)^2 \qquad (7)$$

By minimizing with regard to $\hat{\alpha}$ and $\hat{\beta}$, the coefficient estimators are given by:

$$\hat{\beta} = \frac{\sum_{t=1}^{T}(x_t - \bar{x})(y_t - \bar{y})}{\sum_{t=1}^{T}(x_t - \bar{x})^2} \tag{8}$$

$$\hat{\alpha} = \bar{y} - \hat{\beta}\bar{x} \tag{9}$$

where $\bar{x}$ represent the mean of x, and $\bar{y}$ the mean of y.

6.2.3. Goodness of fit

To measure the precision of the regression model, R^2, or the coefficient of determination is the ratio of explained variation relative to the total variation. In other words, R^2 express the ratio of the explained sum of squares to the total sum of squares (Studenmund, 2017). The coefficient of determination can be defined by the following relations:

$$R^2 = \frac{ESS}{TSS} = 1 - \frac{RSS}{TSS} \tag{10}$$

Where ESS correspond to the explained sum of squares, and TSS to the total sum of squares. These can be explained by:

$$TSS = \sum_{i=1}^{n}(y_i - \bar{y})^2 \tag{11}$$

$$ESS = \sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2 \tag{12}$$

$$RSS = \sum_{i=1}^{n}\hat{u}_i^2 \tag{13}$$

R^2 will take values between 0 and 1. Values close to 1 indicates that the model explains a high portion of the variance in the dependent variable, where values towards zero indicate the opposite. The ratio serves as an intuitive and straightforward measure to interpret. In practice, however, the measure is not unproblematic. R^2 will never decrease when an explanatory variable is added to the model, making it a deficient tool to decide whether a variable should be added to the model or not (Wooldridge, 2012). Instead, we can use the adjusted r-squared. The adjusted r-squared looks at the trade off by including more variables to the model against the increased sum of squares of the model. Making it a better and more suitable measurement for the goodness of fit.

6.2.4. Hypothesis testing

Hypothesis tests serve as an important tool in econometrics in order to draw conclusions from the findings in regression analysis, with a certain level of confidence. The standard is to use a confidence interval of 95 percent when conducting the tests, implying that a hypothesis is kept if there is a 95 percent probability for that it is true. There are always two hypotheses, the null hypothesis (H_0) and the alternative hypothesis (H_A) (Wooldridge, 2012).

A hypothesis test can be carried out in two different ways, a one-sided test or a two-sided test. For a two-sided test, the null hypothesis could be that the true value of β is equal to zero, where the alternative hypothesis states that it is not zero. The hypothesis test can thereby be expressed as:

$$H_0: \beta = 0$$

$$H_A: \beta \neq 0$$

In this case, the alternative hypothesis declares that β could be either above or under zero. Alternatively, H_A could take the following form:

$$H_A: \beta > 0$$

From this, one has a one-sided test. It is worth noticing that the null hypothesis always considers a specific value.

Moreover, in order to be able to test the hypothesis, a test statistic is necessary. A typical test to use is the t-test, which relies on the t-distribution with $n - k - 1$ degrees of freedom, with n

being the number of observations and k the number of variables included. The t-test statistics can be formalized as follows:

$$t_{\beta_j} = \frac{\hat{\beta}_j - \beta_{H_0}}{se(\hat{\beta}_j)} \tag{14}$$

With the t-value has been calculated, it will be compared to a critical value from the t-distribution. If the t-statistics were higher than the critical value, one would reject the null hypothesis.

The t-test is suitable when testing one parameter. If one would go about to test several of the joint parameters, an F-test is used. The F-test checks if the model specified with restrictions is different from the model without, just the intercept. If the H_0 is rejected, it means that the model one has specified is a better fit then without the restrictions. With the restrictions being the specified variables. The F-test can be calculated as:

$$F = \frac{\frac{R^2}{(k-1)}}{\frac{1-R^2}{(n-k-1)}} \tag{15}$$

Where n is the number of observations, and k is the number of variables/restrictions.

Another test that will be used is the Z-test. The Z-test assumes a normal distribution and checks if the means are the same between two populations. To use the test, the sample size needs to be large enough due to the normal distribution. For this thesis, the z-test can be used to check for equality between two beta coefficients of the same regression. The formula for the Z-test can be stated as:

$$Z = \frac{b_1 - b_2}{\sqrt{(SEb_1^2 + SEb_1^2)}} \tag{16}$$

Where b_1 and b_2 represent the beta coefficients being tested. SE is the corresponding standard error of the beta estimates.

6.2.5. Lagged variables

A lagged variable is simply the value a variable took in the previous period. The variable x_t would thus have the lagged variable x_{t-1}. To obtain the lagged variable from the initial, all observations are pushed forward one period. Meaning that one loses an observation for each time step that is lagged, as the first observation will be pushed forward. The reason for using lagged variables in a model often has its root in time difference between observed value and publication of the value. It is also essential to understand the concept of lagged variables when testing for autocorrelation, as it looks into the relationship between the correlation between a variable and its previous values. It can also give insightful information when trying to predict tomorrow's value of a dependent value based on today's observed values.

There is, however, no definite answer to the choice of an optimal number of lags to include in equation x. A standard method is to choose the number of lags (p), which minimizes an information criterion (Brooks, 2014). The information criteria consist of two parts, where the first part employs a function of the sum of squared residuals (SSR), while the other is a function of parameters used in the model concerning the total number of observations. Therefore, adding an additional lag to the model will have two competing impacts on the information criteria. On the one hand, SSR will decrease, but opposite, the value of a penalty term related to the loss of degrees of freedom will increase (Ibid.). For this thesis, the Akaike's Information Criteria (AIC) will be utilized. AIC is given as

$$AIC = \ln(\hat{\sigma}^2) + \frac{2k}{T} \tag{17}$$

where $\hat{\sigma}^2$ is the estimated residual variance, k is the total number of parameters, and T represents the sample size.

6.2.6. Multiple linear regression

A regression analysis with more than one explanatory variable is considered a multiple regression. The purpose of multiple regression analysis is to examine the effect of the independent variable on the one independent variable. Regarding the simple model outlined earlier, this model is expanded to include additional variables:

$$Y_t = \beta_0 + \beta_1 x_{1t} + \beta_2 x_{2t} + \cdots + \beta_k x_{kt} + u_t, \qquad t = 1,2,\dots,T \tag{18}$$

Where k represents the number of independent variables in the selected model. The equation express that the dependent variable Y depends on the constant term β_0, together with the parameter coefficients $\beta_1, \beta_2, \dots, \beta_k$ which quantifies the effect from each of the independent variables (Brooks, 2014). Further, the beta coefficients measure the change in the dependent variable due to a one-unit increase in the respective independent variable, holding all other independent variables fixed. Thus, when estimating the effect of x_k on Y, the other variables are controlled for. The error term u captures the stochastic variation in Y, in which the explanatory variables fail to explain (Stock and Watson, 2015).

6.2.7. OLS assumptions

In order to secure that the output from the regression model is accurate and can be used as a foundation for concluding, the input data for the model must satisfy certain assumptions. Under OLS estimation, the number of assumptions utilized varies, where some sources may split the assumptions into components. Additionally, some assumptions are more trivial than others. Therefore, six important and unique assumptions will be presented, where it applies that if assumption 1 to 5 holds, the β estimators from the OLS is said to be the Best Linear Unbiased Estimators (BLUE). (Wooldridge, 2012). The consequences for violation of the assumptions, together with how to detect them, will also be outlined.

Assumption 1: Linear in Parameters

A decisive assumption for the regression model is that the dependent variable Y must be a linear function of the explanatory variables, which is reflected in equation xx above. In case this linearity assumption is violated, the model will attempt to force the data into a linear relation. In order to rectify problems with non-linearity, a transformation of the data may be suitable. For instance, employ the logarithm of the variable or use the squared values.

Assumption 2: No Perfect Collinearity

Another important assumption is that the independent variables should not be correlated. If there exists a high correlation between the explanatory variables, there will prevail in an internal cause-effect relationship that may lead to inaccuracy and noise in the analysis. The more highly correlated the explanatory variables are, the more challenging it will be to determine how much of the variation in Y each explanatory variable is responsible for. Hence, the beta coefficients will achieve higher variances, increased standard errors, and therefore, lower t-values (Baltagi, 2011). In other words, the presence of multicollinearity in a multiple regression carries the risk of receiving statistically insignificant variables, when they should be significant. Moreover, the problem will also make the regression model extremely sensitive to changes in the model specification. Inclusion or exclusion of variables or observations may, therefore, greatly change the estimated coefficients.

With the objective of revealing multicollinearity, there are two main approaches. One method involves studying the matrix of correlation between the independent variables. Within this approach, obtained correlations close to -1 or 1 might indicate problems with multicollinearity. However, the shortcoming with this method is that it only detects multicollinearity between two variables (Brooks, 2014). On the other hand, the second method does capture the effect of all the explanatory variables. This is known as the variance inflation factor (VIF), which measures the extent to which multicollinearity has raised the variance of an estimated beta coefficient. It looks at the degree to which an independent variable can be explained by all the other independent variables in the model (Wooldridge, 2013). VIF is given by the following equation:

$$VIF\ (X_k) = \frac{1}{(1 - R_k^2)}, \quad k = 1,2, \dots, N \tag{19}$$

Where R_k^2 is the R^2 for the auxiliary regression for variable X_k.

Regarding how high the VIF should be before it occurs, problems with collinearity for the regression is said to be arbitrary. Some states that problems arise when the factor surpasses 10, while others view it as a complication when it exceeds 5. The VIF method, therefore not contribute a definite answer but express that one should pay attention to VIF values greater than 5.

Assumption 3: Zero Conditional Mean

Given the explanatory variables for the regression model, the expected value of the error term u_t should equal zero:

$$E(u_t|X) = 0, \qquad t = 1,2, \dots, n. \tag{20}$$

This is an important assumption which states that the error term at any given time needs to be uncorrelated with each of the independent variables. For a small sample size, it is likely that the mean is not exactly zero, but as the sample size approaches infinity, the mean of the error term will correspondingly approach zero. In order to avoid a violation of this assumption, the inclusion of an intercept will absorb the potential non-zero mean of the error term (Studenmund, 2017). Particularly, the intercept equals the fixed portion of the dependent variable that cannot be explained by the explanatory variables. Hence, the error term equals the stochastic part of the unexplained value of Y.

Assumption 4: Homoscedasticity

Independent by the levels of the explanatory variables, the variance of the unobservable error term u must be constant over time. A violation of this assumption is not meet; the error terms are heteroscedastic, which may lead the OLS estimators to no longer contain the minimum variance among the unbiased estimators. Consequently, if the OLS estimation is conducted in the presence of heteroscedasticity, it will still provide unbiased beta coefficients, but the standard errors might be misleading (Brooks, 2014).

To detect heteroscedasticity, the first possibility is to study the plot of the estimated residuals against the estimated dependent variable. If a cone-shaped pattern is found, where the vertical range of the residuals increases with time, this can be an indication of heteroscedasticity. Different tests can be employed to check for heteroscedasticity, whereas one of the most prevalent is the White's test (Wooldridge, 2013). This test applies auxiliary regression, where the squared residuals from the original model are regressed on the explanatory variables together with their squared values and the cross products between them. One can then perform the White's test that can be formalized as

$$White's\ Test: TR^2 \sim \chi^2_m \tag{21}$$

Where T is the number of observations, m is the degrees of freedom and R^2 is the one obtained in the auxiliary regression.

If the regression model has problems with heteroscedasticity, there are various ways to solve the issue. One way could be to conduct the regression upon the natural logarithms. Another possibility could be to use heteroscedasticity-consistent (H.C.) standard error estimates, also known as robust standard errors. This remedy adjusts the estimation of the standard errors while still applying the OLS estimates of the slope coefficients (Studenmund, 2017). The intuition behind relies on the fact that heteroscedasticity cause complications with the standard errors in the regression model, but not for the beta coefficients. Therefore, utilization of robust standard errors improves the standard errors by considering heteroscedasticity, while simultaneously, the beta coefficients remain unaffected. Generally, the standard errors provided with this approach will be larger than the uncorrected standard errors from the original OLS regression. In this way, the generated t-statistics will be lower, which will decrease the probability that an estimated beta coefficient will be significantly different from zero (Ibid.).

Assumption 5: No serial correlation

Serial correlation, also known as autocorrelation implies that the value of the error term from one period depends on the value of the error term in other periods, in some systematic way.

This assumption states that there should be no autocorrelation between the error terms in the regression model:

$$Corr\ (u_t, u_s) = 0, for\ all\ t \neq s \tag{22}$$

The consequence of conducting a regression with the presence of autocorrelation in the error terms is that the OLS estimators will not be BLUE, meaning that the estimated standard errors will be misestimated and provide unreliable t- and F-tests (Wooldridge, 2012).

Since the actual error terms cannot be observed directly, the estimated residuals from the regression model are utilized for detection. It can be useful to plot the residuals against time as well as an ACF-plot of the residuals, to look for indications of autocorrelation. However, it is common to rely on more formal tests for autocorrelation, as the Durbin-Watson (D.W.) test and the Cumby-Huizinga (C.H.) test (Brooks, 2014).

The D.W. test is used to determine whether there prevails first-order serial correlation in the error term, i.e., it checks for autocorrelation with its first lagged error term. When performing a D.W. test, the regression model must fulfill certain underlying assumptions (Brooks, 2014):

1. The regression includes an intercept term.
2. The autocorrelation is first order in nature:

$$u_t = \ \rho u_{t-1} + v_t \tag{23}$$

Where ρ represents the autocorrelation coefficient, and u is a normally distributed error term.

3. The regression model does not involve a lagged dependent variable as an explanatory variable.

The equation for the D.W. statistic for T observations is given as:

$$DW = \frac{\sum_{t=2}^{T}(\hat{u}_t - \hat{u}_{t-1})^2}{\sum_{t=2}^{T}\hat{u}_t^2} \tag{24}$$

If the calculated value of d is close to 2, it indicates that the regression does not suffer from autocorrelation. Differently, in a case where the observed value is near zero, one can assume that there exists positive autocorrelation, and if d is close to 4, then one can presume negative serial correlation (Studenmund, 2017). Moreover, the D.W. test applies an upper (d_U) and a lower (d_L) bound for critical values, to determine whether autocorrelation is present or not. These critical values depend on the numbers of observations and explanatory variables in the regression model. The following figure illustrates the dynamics:

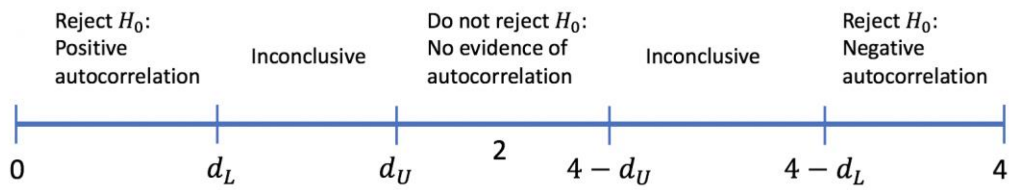


Figure 5.2 – Durbin-Watson test – Rejection and non-rejection regions for autocorrelation

The null-hypothesis states that there exists no first-order autocorrelation in the regression. As seen in figure 5.2, for some intervals of d, there will not be drawn any conclusion of whether there prevails autocorrelation in the model or not. Another drawback of the D.W. test is that it will not find any autocorrelation beyond the first lagged error term. Therefore, it is also desirable to investigate a joint test for autocorrelation. The CH test constitutes a more general test, which can check for autocorrelation between the error term and several of its lagged values simultaneously. Furthermore, there are several benefits to this test, which does not apply to other autocorrelation tests (Baum and Schaffer, 2013). If autocorrelation is expected at a specified lag order, many test statistics are incapable of testing for autocorrelation in the succeeding lags.

Similarly, these tests are only capable of determining whether there prevails autocorrelation within the predetermined range or not. The CH test, however, can determine which lags that suffer from autocorrelation. Another advantage with the usage of the C.H. test is that the test is compatible with heteroscedasticity-consistent (H.C.) standard error estimates (Cumby and Huizinga, 1990).

If autocorrelation is present in the regression model, it can be a consequence of several factors. One can be the problem of an omitted variable, where the error term will capture the effect of this, and thus cause autocorrelation. Other factors could be the inclusion of irrelevant explanatory variables or the use of an inappropriate functional form (Brooks, 2013). Problems with the presence of autocorrelation can be rectified by adjustments of the regression model, where the utilization of Newey-West standard errors represents a typical application (Studenmund, 2017). This method represents an augmented version of the robust standard errors, which takes both the presence of autocorrelation and heteroscedasticity into account, by merely estimating prioritized covariances between residuals. In this way, the Newey-West method simplifies the formula for the regression model's standard errors. The model will, in this case, no longer be efficient, but inferences of the model parameters will be valid due to the greater accuracy of the Newey-West standard errors compared to a model uncorrected for serial correlation (Ibid.).

Assumption 6: Normality

The final assumption states that the error term u in the model is independent of the explanatory variables, and is normally distributed with zero mean and variance equal to σ^2. This assumption can then be expressed as follows:

$$u \sim N(0, \sigma^2) \tag{25}$$

In case this assumption is violated, the t-tests and F-tests associated with the regression model will be unreliable. Similarly, the p-values for the regression coefficients will not be trustworthy (Wooldridge, 2013). To check whether the error terms are normally distributed, a Jarque-Bera (J.B.) test statistic can be performed. This test uses the fact that a normal distribution does not possess skewness, but holds a kurtosis coefficient of 3. Further, the JB-test examines if there is a joint coefficient that is zero. The skewness and kurtosis coefficient can be expressed as (Brooks, 2017):

$$Skewness: b_1 = \frac{E[u^3]}{(\sigma^2)^{\frac{3}{2}}} \tag{26}$$

$$Kurtosis: b_2 = \frac{E[u^4]}{(\sigma^2)^2} \tag{27}$$

Where the error term is denoted by u. As the kurtosis for a normal distribution is 3, the excess kurtosis $(b_2 - 3)$ is zero. The JB-test statistic can then be formalized as:

$$JB: W = T\left[\frac{b_1^2}{6} + \frac{(b_2 - 3)^2}{24}\right] \tag{28}$$

Where the sample size is denoted T. The test follows an asymptotical X^2-distribution under the null hypothesis that assumes a symmetric and mesokurtic distribution of the series, i.e., a normal distribution. The null hypothesis is rejected if the error terms of the model are significantly skewed and/or leptokurtic (Brooks, 2017).

If error terms are not normally distributed, it is often caused by a few residuals not fitting with the rest of the observations. The observations that cause residuals that do not fit are often found in the extreme tails of the distribution, leading to a high kurtosis. These observations that do not fit the remainder are known as outliers (Brooks, 2017). Outliers penalize an OLS model heavily, as it will cause a higher RSS for being far from the fitted line. It is common to define an outlier as an observation that is three or more standard deviations away from the mean, and this will be used in this thesis. To identify outliers, it is usual to plot the studentized residuals over the time horizon. This gives a clear overview of what observations that might cause a problem for the normality in the error term.

6.2.8. Dummy Variable

Dummy variables are a helpful tool when conduction regression analysis and is often used. A dummy variable is also called a quantitative variable, as it can be used as a numerical value for a quantitative specification (Brooks, 2017). A dummy variable is a variable that takes a specific value when a requirement is met, often values 0 or 1. Dummy variables are used in the same way that other explanatory variables are used, but it allows it to make distinctions between the same explanatory variable. The dummy variable then allows for more specific tests without altering the dataset.

6.2.9. Time series and stationarity

Working with time series involves gathering a set of observations over a specified period. Each value is generated from a statistical process, where it emerges from an assumption that all observations are not foreknown, and thus, the observations are considered random. This leads to a sequence of random variables indexed by time, which is called a stochastic process or a time series process (Wooldridge, 2012).

For time series analysis, it is of great importance that the data is stationary. A time series follows a strictly stationary process if the probability distribution for its values is constant over time. This means that the probability the value of a given variable end up inside a definite interval is the same now as for any given point in time, in the past or the future (Brooks, 2014). For a time series to be employed in a regression, weak stationarity, also called covariance stationarity, is sufficient. In order to assert a process as weak stationary, the covariance has to be independent of time. Particularly, as long as the time-frequency between the observations are the same, the covariance must take the same value. In the rest of this thesis, stationarity will correspond to weak stationarity. The following equations must be fulfilled for the time series to be considered stationary (Ibid.):

$$E(y_t) = \mu \tag{29}$$

$$2.\quad E(y_t - \mu)(y_t - \mu) = \sigma^2 < \infty \tag{30}$$

$$3.\quad E\left(y_{t_1} - \mu\right)\left(y_{t_2} - \mu\right) = \gamma_{t_2 - t_1} \quad \forall\, t_1, t_2 \tag{31}$$

This states that that the series should have a constant mean, constant variance, and constant autocovariance structure to be considered a stationary process. The autocovariance determines how y is related to its previous values. In a stationary time series process, the covariance between two observations only depends on the distance between $t_1$and t_2, and not by t itself (Brooks, 2014). Such a time series will tend to return to its mean and fluctuate around this with steady amplitudes. Accordingly, for a stationary time series, shocks will not have an enduring effect. The effect of a shock arisen at t, will have a smaller effect at point $t + 1$ and a further reduced effect at point $t + 2$. In this way, for a stationary process, shocks are transitory and ensures that the process does not explode or wander off. Contrary, for a non-stationary process, shocks will have a permanent or persistent effect (Ibid.).

A time series which do not fulfill the above three equations are considered non-stationary processes. If non-stationary variables are used in a regression model, it can be confirmed that the standard assumptions for asymptotic analysis will not be valid (Brooks, 2014). Hence, the usual t-ratios will not follow a t-distribution, and the F-statistics will not follow an F-distribution. Therefore, If the variables are non-stationary, it is not attainable to undertake valid hypothesis testing of the regression parameters. Furthermore, non-stationary variables can lead to regression analysis to produce a high explanatory power (R^2), even though there is no relationship between the variables in real life. This phenomenon is called a spurious regression, in which its estimators and test statistics can result in misleading conclusions (Wooldridge, 2013).

When applying non-stationary time series data for a given period, one cannot generalize the results to other periods. Thus, the results provide limited practical value. Many financial and economic time series are characterized as non-stationary, where the typical case is time series following a random walk.

6.2.10.Random walk and unit root

A model that has been frequently used to describe a non-stationary process is the random walk process. If a time series follows a random walk without drift, the best point estimate for the value of tomorrow will be today's value. This implies that the value of a variable y at time

t corresponds to the value at time $t - 1$, plus a random shock (Baltagi, 2011). A random walk without drift is defined as follows:

$$y_t = y_{t-1} + u_t \tag{32}$$

where the shock component u_t represent a stochastic error term with an expected mean of zero and constant variance. For a random walk without drift $E(y_t) = y_0$ and $Var\ (y_t) = t\sigma^2$, which indicates that the expected value is constant, while the variance increases proportionally with time t. Consequently, this stochastic process cannot be considered stationary, as the variance condition is broken (Baltagi, 2011). Some series have an apparent upward tendency, whereas the best model of the series should include an alteration for the increasing trend. This alteration leads to an extension of the random walk model, where a drift component is included. Random walk with drift is defined as follows:

$$y_t = \mu + y_{t-1} + u_t \tag{33}$$

For this model, it can be deduced that $E\ (y_t) = y_0 + t\mu$ and $Var\ (y_t) = t\sigma^2$. This relieves that both the mean and variance increases with time. Therefore, a random walk with drift violates the assumptions for stationarity and is therefore also considered non-stationary.

Equation 28 can be rewritten as:

$$Y_t = \mu + \rho Y_{t-1} + u_t, \qquad where -1 \leq \rho \leq 1 \tag{34}$$

Where ρ represents the autocorrelation coefficient. When $\rho = 1$, equation 33 is obtained, which represents a random walk with drift. The same is valid for the random walk without drift and describes what is characterized as a unit root process, which is non-stationary (Wooldridge, 2012). If $|\rho| < 1$, it implies that a shock occurring at time t will fade away correspondingly with the size of ρ, by which the process will be considered stationary. Accordingly, within time series analysis, it serves as essential to test for one or several unit roots in variables, to determine whether the time series is stationary or not.

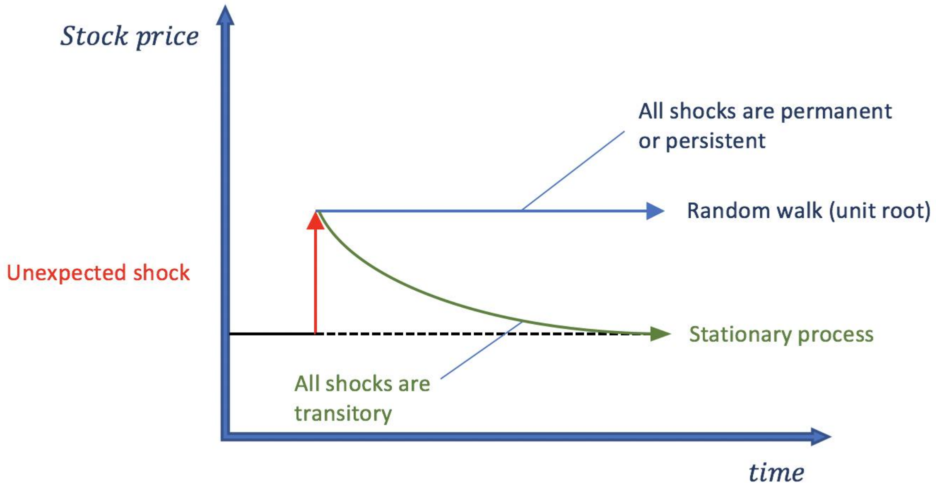


Figure 5.3 – Stationary and unit root process reaction to shocks

Stationarity and unit root testing

Before conducting formal tests for stationarity, it is beneficial to inspect the time series plot of the data. This can disclose properties for the time series examined, such as trends. Another plot that reveals useful information is an autocorrelation function (ACF) plot. A low first autocorrelation coefficient, in addition to a time series plot with no seeming trend, proposes that the series does not involve a trend (Stock and Watson, 2015). However, if doubt remains, formal statistical procedures can be used to test the hypothesis that the data contain a unit root against the alternative that there is no unit root.

The Dickey-Fuller test for a unit root represents one of the most commonly used test in practice, as well as one of the most reliable (Stock and Watson, 2015). A unit root will be present if the coefficient $\rho = 1$, which conveys that the time series is not stationary. Initially, equation 27 is rewritten to:

$$\Delta y_t = \delta y_{t-1} + u_t \tag{35}$$

where $\delta = (\rho - 1)$, and Δ indicates that the time series is first-differenced $(y_t - y_{t-1})$. Thus, the null hypothesis of the existence of a unit root is H_0: $\delta = 0$, i.e. non-stationarity. Where the alternative hypothesis is one-sided, H_1: $\delta < 0$. The test statistic is similar to an ordinary t-test, but since the null hypothesis implicates non-stationarity by the presence of a unit root, an alternative set of critical values from Dickey and Fuller (1981) is used. Within this test, in absolute values, the critical values are larger than the values employed in a standard t-distribution, because of more strict requirements related to the rejection of the null hypothesis (Brooks, 2014). If the null hypothesis is rejected, it indicates that the time series is stationary. When conducting this test, it must be adjusted to whether the series follows a random walk without drift (equation 35), a random walk with drift, or a random walk with drift around a stochastic trend. The remaining two equations are defined as follows:

$$\Delta y_t = \mu + \delta y_{t-1} + u_t \tag{36}$$

$$\Delta y_t = \mu + \beta_t + \delta y_{t-1} + u_t \tag{37}$$

The test is only valid if the residuals are serial uncorrelated (Brooks, 2014). To face the possible complication with autocorrelation in the residuals, one can include lagged changes in y. This extension of the test is called the Augmented Dickey-Fuller test (ADF) and is due to the above mentioned more applied in practice. For a random walk with drift around a stochastic trend, the ADF will take the following form:

$$\Delta y_t = \mu + \beta_t + \delta y_{t-1} + \sum_{i=1}^{p} \alpha_i \Delta y_{t-i} + u_t \tag{38}$$

Under the ADF test, the same null hypothesis is applicable, and the two tests have the same critical values.

Transformation of non-stationary time series

In order to avoid problems with spurious regressions that can occur with non-stationary time series, it is decisive to transform the time series to make them stationary. The method of transformation depends on whether the time series is a difference stationary process (DSP) or a trend stationary process (TSP). Most of the financial time series are characterized as DSP's (Baltagi, 2011).

Random walk processes with or without drift are defined as DSP, which implies that stationarity can be obtained by taking the first-order difference. A random walk process without drift (Equation 4), contain a unit root and can be described as an integrated process of order I (1) since the time series becomes stationary by differencing once. This can be shown as:

$$\Delta y_t = y_t - y_{t-1} = u_t \tag{39}$$

Here, u_t represent an error term with the expectation zero mean and constant variance. Similarly, it can be demonstrated that a random walk with drift (Equation 36), turn stationary with differencing:

$$\Delta y_t = y_t - y_{t-1} = \mu + u_t \tag{40}$$

A random walk with drift is also an integrated process of order I (1) since the non-stationarity in y_t is eliminated by differencing the series once (Brooks, 2012). In a case where a time series must be differenced further to become stationary, the process is said to be integrated of order I (d), where d represents the number of differentiations required to generate a stationary process.

Contrary, a trend stationary process is stationary around a linear trend. This type of process will require different treatment to induce stationarity. If the trend in a time series is predictable, it is characterized as a deterministic trend and will follow a trend stationary process. Such a process is described as follows.

$$y_t = \mu + \beta_t + u_t \tag{41}$$

In this case, a de-trending is required to generate a stationary process. If it is believed that only this class of non-stationarity is existing, a regression on the above form would be run, and the subsequent estimation would be performed on the associated residuals. Hence, the linear trend would be removed (Brooks, 2012).

7. Data Presentation

In this section of the thesis, the data for the selected variables will be presented and processed. The goal will be to differentiate until we are certain our variables are stationary. To get a better insight into the variables individually, descriptive statistics will be presented and further examined, as well as commented. Giving us a better understanding, before looking gin to the pairwise correlation between variables, that gives a good indication towards the regression analysis.

7.1. Structural break

When investigating financial time series over a long period, some major changes have likely occurred and affected the underlying data generating process. Such changes typically arise from different kinds of shocks disturbing the markets. Among these, there can appear supply, demand, or financial shocks. As this thesis investigates the relationship between OSE and other variables, including the oil price, over time, a division into subparts seems favorable.

As previously mentioned, several events have caused major fluctuations in the oil price. One of these is the financial crisis in 2008 that also had a considerable impact on several of the other variables investigated in this thesis. This is also consistent with the high volatility detected during this period. Consequently, we will first run the full regression model from 1988-2019, and thereafter divide the period into two subparts. The first period will thus start from January 1988 and run through September 2008, while the second period will run from the beginning of 2009 until the end of 2019. Moreover, in order to examine the oil price's influence on OSE both with and without the renewable energy index (RENIXX), a fourth regression will be supplemented.

The period from September 2008 to the end of 2008 has been left out of both subparts. This is done to reduce the abnormal volatility affecting the markets during this period. As a result, we will be conducting four regression in total. The regression will be defined as follows

1. Model 1: The whole period from 1988 to the end of 2019

$$OSEAX = \beta_0 + \beta_1 SP500 + \beta_2 Brent + \beta_3 FTSEAX + \beta_4 NIBOR + \beta_5 USDNOK$$

2. Model 2: Subpart 1, running from 1988 to end of August 2008

$$OSEAX = \beta_0 + \beta_1 SP500 + \beta_2 Brent + \beta_3 FTSEAX + \beta_4 NIBOR + \beta_5 USDNOK$$

3. Model 3: Subpart 2, running from the start of 2009 to the end of 2019

$$OSEAX = \beta_0 + \beta_1 SP500 + \beta_2 Brent + \beta_3 FTSEAX + \beta_4 NIBOR + \beta_5 USDNOK$$

4. Model 4: Subpart 2 (including RENIXX), running from the start of 2009 to the end of 2019

$$OSEAX = \beta_0 + \beta_1 SP500 + \beta_2 Brent + \beta_3 FTSEAX + \beta_4 NIBOR + \beta_5 USDNOK + \beta_6 RENIXX$$

The purpose of running regression model 2 and 3, is to be able to compare them both the each other, and to the full period. In this way, primarily the development of the oil prices' influence on OSE will be examined. The similar will be relevant for the remaining variables. With regression model 4, the main intention is to compare this to regression model 3.

We will split up the testing defined by our four regression models, and conduct tests in an orderly way to give a clear overview of the testing. This means that each regression model will be treated individually and may contain slight differences if the tests suggest it.

For all tests conducted, we use a 5 percent significance level. Thus, we will only reject the null hypothesis based on the 5 percent significance and be 95 percent certain that it is indeed true. This will sometimes lead to results where there is a strong indication of rejecting the null hypothesis and indicated graphically. However, we cannot conclude by rejecting the null hypothesis on the 5 percent level.

7.2. Testing for stationarity

As covered in the theory section, time series used for time series regressions must be stationary. It is expected that the observed stock index development will not be stationary in its raw data. This is due to the matter that following the close of trading, a stock or index does not reset and starts from zero. Instead, it will progress from its last closing price and thus apply its previous price as a benchmark for further development. This feature is illustrated in figure x.

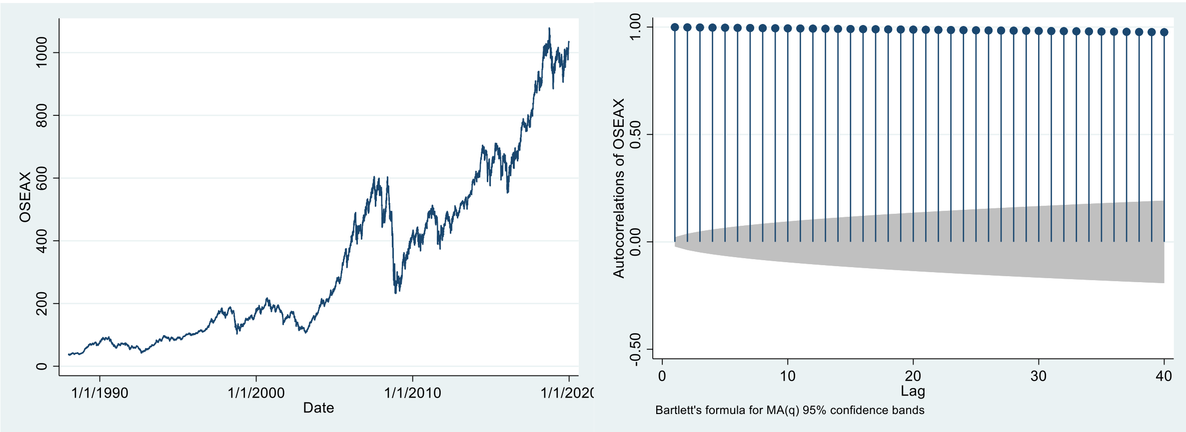


Figure 7.1 – Development of OSEAX, raw data (LHS)

Figure 7.2 – ACF plot OSEAX, raw data (RHS)

From figure x, the presence of a positive trend can be derived. OSEAX is treated with daily data, which makes the trend even more apparent, as longer intervals would, in some periods, smooth out the trend. Further, by looking at the ACF plot of OSEAX, this demonstrates larger systematic fluctuations outside the 95 percent interval. Based on these observations, it, therefore, appears evident that the variable is not stationarity in its original price form. For the remaining variables, price development over time, together with associated ACF plots can be found in appendix X. Similar to OSEAX, these variables also demonstrate problems concerning stationarity, at price form.

In order to confirm these visual findings, a numerical test is favorable. To test for stationarity, an ADF is employed to check for unit roots. With critical values of -3.43 at 1%-significance, -

2.86 at a 5%-significance and -2.57 at a 10%-significance level, the ADF test confirm what was suggested by the visual inspection. We cannot reject the null hypothesis for any of the variables, on any of the significance levels, except for NIBOR. The underlying reason why NIBOR makes an exception is that this variable is quoted as percentage points, rather than prices as other variables.

Variable	T-Statistic	Result
OSEAX	0.91	Non-Stationarity
SP500	1.014	Non-Stationarity
Brent	-1.586	Non-Stationarity
FTSEAX	-2.183	Non-Stationarity
NIBOR	**-3.482**	**Stationarity**
USDNOK	-1.841	Non-Stationarity
RENIXX	-1.987	Non-Stationarity

Table 7.1 – ADF test on variables (initial form)

7.3. Variable Transformation

As the variables appear non-stationary, except for NIBOR, we will use the natural logarithm to differentiate the time series. When working with financial data, it is beneficial to operate with log returns, as they possess favorable characteristics, making them suitable for time series regression. This is done in the following way:

$$X_t = \ln\left(\frac{x_t}{x_{t-1}}\right)$$

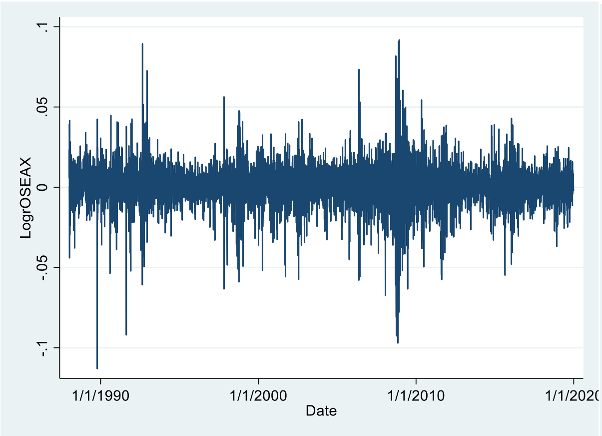
For all variables except NIBOR, this approach is used. However, as NIBOR is already presented in percentage terms, this variable will be differentiated by taking the change from the previous, to the current value:

$$X_t = x_t - x_{t-1}$$

With this regard, the variables now possess more desirable and appropriate attributes for further work with time series regression. However, these will also be tested for stationarity.

7.4. Test for Stationarity on log-returns

In order to test the differentiated variables for stationarity, a similar approach as with the initial data is utilized. First, we inspect the returns and the corresponding ACF plots.



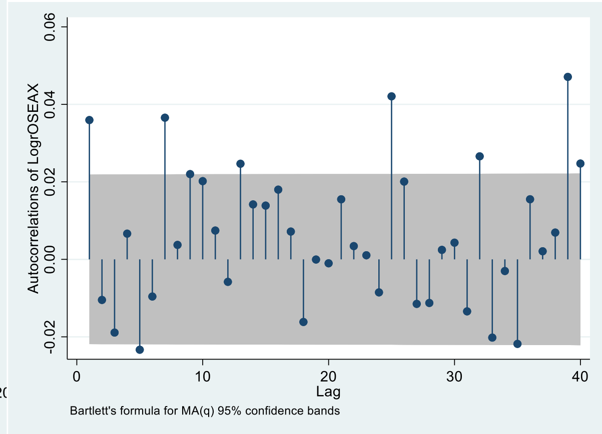


Figure 7.3 – Returns OSEAX (LHS)

Figure 7.4 – ACF plot OSEAX, returns (RHS)

In figure 7.3 above, it can be seen that the returns fluctuate more around zero. Moreover, there are no signs of any apparent trends in the return data. The ACF-plot confirms this, where there seem to be no systematic changes, and additionally, the lags are predominantly within the confidence interval. This evidence implies indications of stationarity in the return data for OSEAX. However, it is worth noting the more extreme variations that cluster before 2010 in figure 7.3, during the financial crisis. For the other variables, comparable plots can be found in Appendix D.

To ensure what the plot suggests, we again run the ADF test to check for stationarity numerically. The t-statistics are now well below the critical values for all variables at all significance levels. Consequently, we are now able to reject the null hypothesis of a unit root, and thus conclude that the variables at the differentiated functional form are stationary. Furthermore, we are now confident that the data is suitable to be used in a time series regression.

Variable	T-Statistic	Result
OSEAX	-86.458	Stationarity
SP500	-93.786	Stationarity
Brent	-88.237	Stationarity

FTSEAX	-98.94	Stationarity
NIBOR	-86.656	Stationarity
USDNOK	-91.184	Stationarity
RENIXX	-60.178	Stationarity

Table 7.2: Result from ADF-test

7.5. Descriptive Statistics

Now that it has been concluded to make use of log-returns for all variables except for NIBOR, we will take a closer look at the characteristics and the variables descriptive statistics.

	OSEAX	***SP500***	***BRENT***	***FTSEAX***	***NIBOR***	***USDNOK***	***RENIXX***
Mean	0.00042	0.00032	0.00017	0.00019	-0.00166	0.00004	-0.00011
Std. Dev.	0.01271	0.01084	0.02272	0.01000	0.23113	0.00710	0.01973
Kurtosis	7.20378	9.08817	13.72839	6.80959	1104.69009	3.57072	8.36941
Skewness	-0.57572	-0.30700	-0.49529	-0.20333	16.40086	0.02291	-0.27680
Minimum	-0.11309	-0.09470	-0.36121	-0.08710	-5.79000	-0.05951	-0.15429
Maximum	0.09186	0.10957	0.19819	0.08811	12.19000	0.04415	0.17071
Count	8026	8026	8026	8026	8026	8026	4516

Table 7.3 – Descriptive statistics (1988-2019)

Over the full period, OSEAX has seen the greatest increase, with a daily mean of 0.00042 percent. The remaining variables also show positive daily means over the period, except for NIBOR and RENIXX. Most of the variables also exhibit negative skewness, indicating that the returns are not normally distributed. This is additionally backed up by the high kurtosis for most of the variables. If a variable is normally distributed, a kurtosis of 3 is expected. Anything above this is considered excess kurtosis. Thus, the variables above can be considered leptokurtic, meaning that there is a positive excess kurtosis. Combined with the negative skewness applicable to the majority of the variables, this implies negative outliers. In financial terms, the variables would be considered riskier due to this high excess kurtosis. Moreover, it should be noted that variable with the highest standard deviation, except for NIBOR, is the oil price.

7.6. Correlation Matrix

Before further analysis and testing, it is essential to investigate the relationship between the dependent and independent variables as well as the relationship between all independent variables. A correlation matrix indicates the interaction between the variables and is used to obtain an overview of relations worth further investigation.

	OSEAX	*SP500*	*BRENT*	*FTSEAX*	*NIBOR*	*USDNOK*	*RENIXX*
OSEAX	1						
SP500	0.3699	1					
BRENT	0.1914	0.0568	1				
FTSEAX	0.6535	0.4990	0.1156	1			
NIBOR	-0.0334	0.0041	-0.0036	-0.0162	1		
USDNOK	-0.1080	-0.0776	-0.0974	-0.1131	-0.0576	1	
RENIXX	0.6039	0.3821	0.1761	0.5964	0.0285	-0.1467	1

Table 7.4: Correlation matrix 1988-2019

The correlation matrix above is based in the full period, computed by the log-returns previously found to be stationary. A key take away from the correlation matrix is that all stock indexes are positively correlated with each other, as well as they are all positively correlated with Brent crude.

The highest correlation is also between OSEAX and RENIXX. This is an interesting finding as RENIXX has been presented to represent a new industry to influence the stock exchange in Norway. Furthermore, none of the correlations presented suggests that there will be any problem with near-perfect collinearity. In many ways, the correlations above is an average between the two periods. The most interesting and important points to note from the different periods is that OSEAX and Brent correlation has increased from 0.06 to 0.51, while FTSEAX had a high correlation with OSEAX for the first period, and vice versa for SP500 and OSEAX. It appears that the correlation has shifted from London to New York.

It is also vital that multicollinearity is controlled for. A new correlation matrix is then presented with only independent variables.

	SP500	***BRENT***	***FTSEAX***	***NIBOR***	***USDNOK***	***RENIXX***
SP500	1					
BRENT	0.0631	1				
FTSEAX	0.4990	0.1506	1			
NIBOR	0.0041	-0.0026	-0.0162	1		
USDNOK	-0.0776	-0.1243	-0.1131	-0.0576	1	
RENIXX	0.3821	0.2177	0.5964	0.0285	-0.1467	1

Table 7.5: Correlation matrix between all independent variables.

The highest correlation between independent variables can be found between RENIXX and SP500 at 0.3821. This is a bit high, but it can still be concluded that there will not be any problems with multicollinearity. This is also backed from a VIF test, where all variables take a value that can be rounded 1 one. As a rule of thumb, anything above a value of 5 from the VIF test suggests that a problem with multicollinearity exists. Correlation matrixes for each subpart can be found in Appendix F and G.

8. Analysis and Results

Following structural breaks, the data is split amongst two subperiods. To ensure transparency in the analysis, they are conducted in an orderly way, looking at one period at a time. Each model will first be presented with all initial variables before finding the optimal lag appropriate for each model. To ensure the validity of the models, all OLS assumptions will be controlled before presenting the model and check whether they are BLUE.

We will repeat the steps mentioned above for all models, to ensure each model is treated individually. As a result, there might be slight differences across the models, in terms of variables statistically significant, and the methods applied.

8.1. Analysis One: 1988 to 2019

Analysis one, running over the full period, is conducted first. Giving an overview and laying a baseline for further analysis.

Model Summary			
R	**R-square**	**Adjusted R-square**	**Std.Error of estimate**
0.6764	0.4575	0.4572	.0093

Table 8.1: Summary original model, whole period.

ANOVA					
	df	**Sum Of Squares**	**Mean Square**	**F**	**Sig.**
Model	5	0.5932	0.1186	1352.79	0
Residual	8020	0.7033	0.0001		
Total	8025	1.2965			

Table 8.2: ANOVA output from the original period, the whole period.

Coefficients					
	Coefficients	**Standard Error**	**t Stat**	**P-value**	**VIF**
Intercept	0.0002	0.0001	2.24	0.025	
SP500	0.0715	0.0111	6.42	0.0000	1.33
Brent	0.0921	0.0046	19.67	0.0000	1.02
FTSEAX	0.7571	0.0122	62.02	0.0000	1.35
NIBOR	-0.0013	0.0005	-2.98	0.044	1
USDNOK	-0.03	0.0149	-2.01	0.025	1.02

Table 8.3: Coefficient table from original regression, whole period.

Taking a brief look at the tables above, it is worth noting that there is a 45.72 percent explanatory degree over this period. We would consider this a high explanatory degree, considering the long period with over 8000 observations. F-statistic is also significant, as with the p-values for all individual variables. All VIF values are close to 1, indicating that there is no problem of multicollinearity between our explanatory variables.

8.1.1. Optimal Lag selection

Furthermore, we are investigating if the model can be improved by including any lagged variables. In the theory chapter, the market efficiency hypothesis was presented, stating that the market should always price in the relevant information available. For variables where there exists a time difference in opening and trading time, we suspect that the market efficiency hypothesis might not hold as expected due to the dynamics between markets difference in opening hours. There exists a possibility that some markets take longer to interpret the information available before it is reflected in the market.

We believe it is worth looking into the concerns mentioned above, as it might present us with relevant information as well as improving the model's explanatory power. The original model is extended to include two lags for all variables. This will be the basis for our lag selection, and then by applying the AIC-tests, we will find the model with the lowest AIC value to indicate the optimal number of lags to be included in the model.

After trial and error, it becomes evident that a lagged SP500 has the greatest improvement on the original model. By adding two lags of all variables, we get a model with a higher AIC score, but several of the lagged variables posses insignificant coefficients. Then adjusting the model, by removing insignificant variables, we were able to create a model with the best AIC-score and highest explanatory power. An effort was made towards stripping down the models to include as few as possible variables, to reduce complexity. It was evident that a lagged S&P 500 variable had the highest impact on the model.

	Adj-R^2	**AIC**
No lags	0.4572	-52193.12
SP500 lagged	0.4701	-52387.9
2 lags all variables	0.4719	-52403.97
2 lags all variables (ex. Insignificant)	0.4716	-52409.37
2 lags all variables (ex. Insignificant and SP500)	0.4586	-52214.93

Table 8.4: Goodness of fit summary by including lagged variables.

The lagged S&P 500 is assigned a coefficient value of 0.1442, while the remainder of lagged, significant variables all have coefficients under 0.005. Only the first-order lag of Brent is deemed significant, and the AIC-score suggests the model is overall a better fit opposed to including other lagged variables. Except for the Brent lagged variable, the $USDNOK_{t-1}$ variable has been added as well, resulting in a removal of $USDNOK_t$ due to insignificance at a 95 percent confidence interval.

After the inclusion of lagged variables, the revised model can be summarized as:

$$OSEAX = \beta_0 + \beta_1 SP500 + \beta_2 SP500_{t-1} + \beta_3 Brent + \beta_4 Brent_{t-1} + \beta_5 FTSEAX + \beta_6 NIBOR + \beta_7 USDNOK_{t-1}$$

Model Summary			
R	R-square	Adjusted R-square	Std.Error of estimate
0.6792	0.4721	0.4716	0.0092

Table 8.5: Summary of model one with lagged variables included.

Coefficients					
	Coefficients	**Standard Error**	**t Stat**	**P-value**	**VIF**
Intercept	0.0002	0.0001	1.7	0.089	
SP500	0. 1085	0. 0114	9.34	0	1.40
LagSP500	0.1442	0. 0103	14.17	0	1.16
Brent	0.0917	0. 0036	14.77	0	1.06
LagBrent	0.0203	0.0035	5.23	0	1.01
FTSEAX	0.6952	0 .0129	54.97	0	1.54
NIBOR	-0.0013	0.0005	-2.83	0.004	1
LagUSDNOK	0.0411	0.0148	2.94	0.005	1.02

Table 8.6: Coefficient table from regression with lagged variables.

8.1.2. Controlling OLS Assumptions

1. **Linearity in parameters**

The model has been selected and specified to only include variables of the first order, and we can be assured that the parameters are linear, causing the parameters to be linear.

2. **No Perfect Collinearity**

By looking at the correlation matrix along with the VIF tests, it is suggested that no perfect collinearity is present between independent variables. From the data descriptions, it is evident that none of the independent have a standard equal to zero. Excluding the possibility of a constant mean.

3. **Zero Conditional Mean**

With a large sample of over 8000 observations will naturally reduce the chance of problems regarding a zero conditional mean. By including an intercept in the model, it is implied that error terms have an expected value equal to zero. Thus, we can assume that this assumption is satisfied.

4. **Homoskedasticity**

To control for the assumption of constant variance in the error term, we plot the studentized residual values against the predicted values of OSEAX.

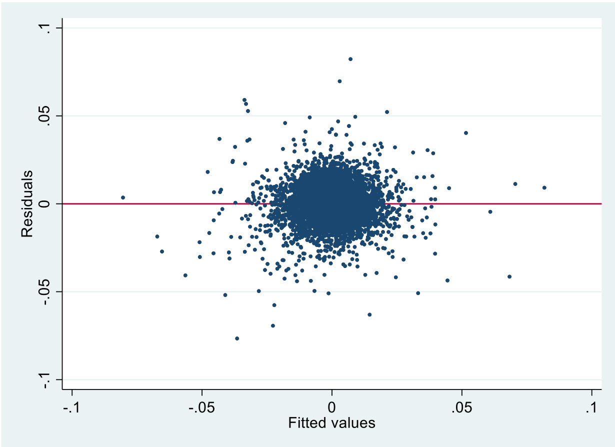


Figure 8.1: Residual vs. Fitted values

In the presence of heteroscedasticity, the plot in figure 8.1 would show a clear pattern. Most of the observations are centered around the middle, except for some outliers scattered across all directions. Having such a high amount of observation, it can be challenging to draw any conclusion by merely looking at the plot.

To formalize the analysis concerning heteroscedasticity, a White's test is conducted as a numerical measure. By performing the test, we obtain a Chi-squared value well above any critical value, resulting in a p-value equal to zero. Due to the low p-value, H_0 of homoscedasticity is rejected, suggesting that a heteroscedasticity problem is present.

Statistical Test	**H_0**	**Chi-squared**	**df**	**Critical Value**	**Conclusion**
White's test	Homoscedasticity	1724.17	44	60.48	Reject H_0

Table 8.7: Results from a white's test for homoskedasticity.

Having such a long-time horizon, it is not surprising that the variance has changed over time. Especially if one considers the fluctuations, and how they have increased over the recent years. With heteroscedasticity present, we know from the theory section that coefficients are unaffected. Heteroscedasticity arises a problem conserving the standard errors being too broad. Resulting in misleading p-values, as it will accept too large area within the confidence area. P-values will then, in fact, be too low, with heteroscedasticity, and insignificant values could be included. To solve this problem, we re-run the regression using heteroskedasticity robust standard errors. Since we do not know the form of heteroscedasticity, using the robust standard error is considered a safe way to solve this problem, especially with such a large sample size. Furthermore, as the HC-standard errors do not alter the model or coefficient, it would not be harmful towards the model. Applying HC-standard errors gives us more confidence regarding the significance of independent variables. Results from the regression run with robust error terms can be seen in table 8.8.

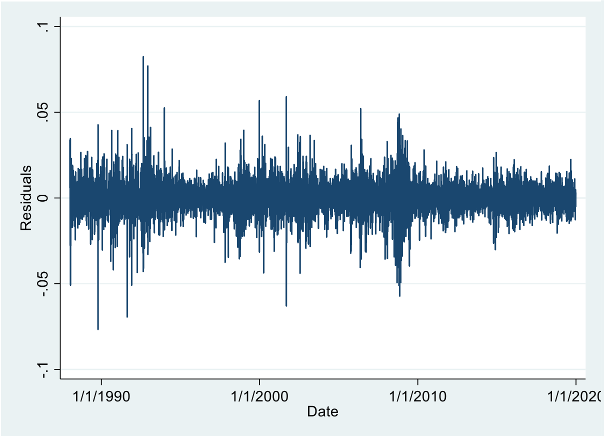
Coefficients					
	Coefficients	**Standard Error**	**t Stat**	**P-value**	**VIF**
Intercept	0.0002	0.0001	1.69	0.092	
SP500	0. 1085	0.0189	5.73	0	1.40
LagSP500	0.1442	0.0157	9.2	0	1.16
Brent	0.0917	0.0079	11.56	0	1.06
LagBrent	0.0203	0.0055	3.72	0	1.01
FTSEAX	0.6952	0.0226	30.81	0	1.54
NIBOR	-0.0013	0.0018	-0.72	0.473	1
LagUSDNOK	0.0411	0.0183	2.24	0.025	1.02

Table 8.8: Regression run with robust standard errors

With the HC-standard errors, all variables have a new t-statistics and its corresponding P-value. It is resulting in NIBOR becoming insignificant with a p-value of 0.473. As one would expect with an insignificant variable, the NIBOR coefficient is close to zero, indicating that it does not improve the model significantly. Using the HC-standard errors by running the model as robust, we regard the assumption of homoscedasticity satisfied.

5. Autocorrelation

To get an impression of the error term and its statistical properties, we plot the residuals over time along with the ACF-plot. This will give a better understanding of how the error term behaves and indicate possible autocorrelation problems.



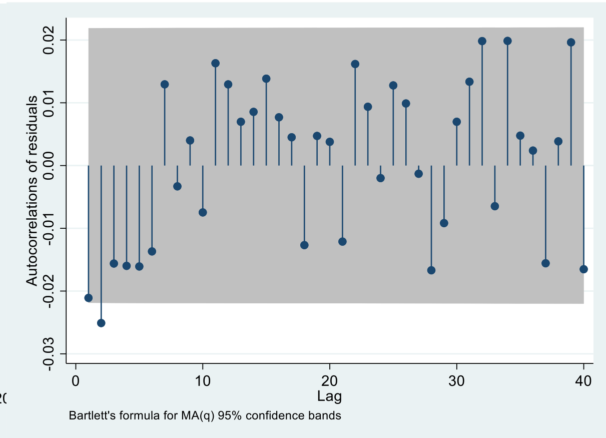


Figure 8.2: Residuals over time (LHS).

Figure 8.3: ACF-plot of residuals (RHS).

The time plot of residuals checks if residuals are independent of time. From figure 8.2, residuals appear to be scattered around randomly, with no clear patterns. Indicating that autocorrelation is not present. Figure 8.3 shows various lags revolving around zero, with no clear pattern in the development of the error term. All for all lags, the value stays within the 95 percent confidence interval, with lag two slightly exceeding. The ACF-plot gives a slight indication that the error term follows an AR(2) process.

As with the previous assumption, a numerical test is conducted to check for autocorrelation. To check for autocorrelation, we run both the DW-test and the CH-test. DW-test allows checking for autocorrelation of the first order, while the CH-test is a more general test, allowing to check for autocorrelation at higher lags, in the presence of heteroscedasticity.

DW-test gives a test statistic of 2.0421. We use the Savin and Whites DW significance table to find dL and dU using the appropriate significance, number of observations, and independent variables in the model, and find dL of 1.697 and dU of 1.841

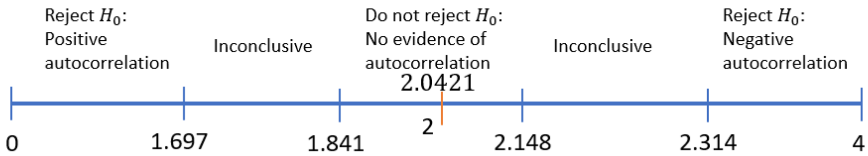


Figure 8.4: D.W. test statistic from regression.

Figure 8.4 illustrates the result of the test, showing the test falls in the area of not rejecting H_0 of no autocorrelation. According to the DW test, no autocorrelation is present in the first order.

To check for up to several lags, a general and often used method is the Breusch-Godfrey (BG) test for serial correlation. However, we have found the model to be heteroscedastic, and BG-test is then not applicable to the model. Instead, we use the presented CH-test. When having high-frequency financial data, it is not uncommon to observe a slight autocorrelation at higher lags, which can be hard to explain. We are, therefore, more concerned about the shorter lags and test for the first three lags using the CH-test. Allowing us to check for the second lagg, where it was indicated from the ACF-plot, that error terms where correlated.

Cumby-Huizinga test			
Lag	**Chi2**	**df**	**p-value**
1	1.384	1	0.2394
2	2.375	1	0.1233
3	0.839	1	0.3598

Table 8.9: Results from the Cumby-Huizinga test with 3 lags.

From table 8.9, it is evident from the CH-test, due to the high p-values, we can not reject the null hypothesis at a 5 percent significance level for any lags. From the plots, DW-test and CH-test, we assume there is no autocorrelation in the error term of the model. Thus, we make no alterations to the model and keep the HC-standard errors, and regard assumption five as satisfied.

6. Normality in error terms

To control the assumption of normality in error terms, we start by looking at a quantile-quantile (QQ)- plot.

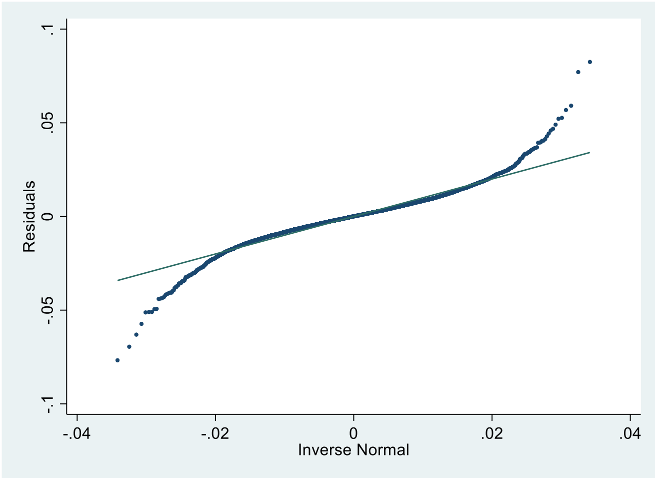


Figure 8.5: Results from a QQ-plot of residuals

A QQ-plot compares residuals values to a theoretical distribution, to see how well it fits the normal distribution of error terms. Figure 8.5 serves as a visual check and is meant to give an indication as well as a visual presentation. Residuals from the model have large deviations in both ends, indicating more extreme values one would find in the theoretical distribution. Contradicting the assumption of normality in error terms. We supplement the plot with a numerical test, the Jarque-Bera (JB) test. Using both the skewness and kurtosis to calculated the JB-statistic, we get a value of 10,000, well above any critical value. Based on the JB-test, we reject the null hypothesis of normality in error terms, and assumption 6 is not satisfied.

The problem regarding non-normality is likely to have its root in the non-normality of independent variables. It is causing them to be harder to fit in a model, leading to residual with higher deviations. In most cases, the assumption of normally distributed error terms are justified with the central limit theorem (CLT), especially if the sample size is large (R. Neu, 2019). We will thus regard this assumption as satisfied for now while examining other possibilities.

Having a sample size above 8000 observations, there is a possibility that there are subsets within the data, having different statistical properties. If this is the case, then models should be built separately for each subset, to fit the data better. Analysis two and three are split according to subsets into the data and will analyze this possibility. If will be interesting to see if it has any effect on the statistical properties across the models.

8.1.3. Outliers

From the QQ-plot in figure 8.5, we learned that we have more extreme values in the residuals than what can be expected from a normal distribution.

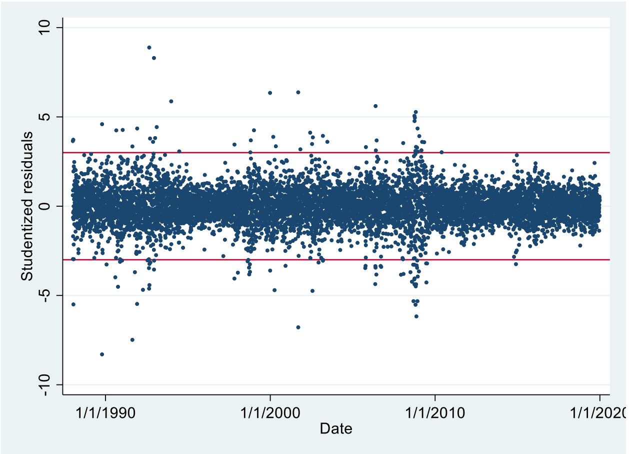


Figure 8.6: Scatter plot of the Studentized residual over time, and values indicating outlier over time.

We want to find out if there are any periods where outliers are clustered, indicating an event causing outliers. Figure 8.6 plots the studentized residuals across time, indicating when outliers occur. Outliers are typically defined as observations more than three standard deviations away from the mean. The red lines in figure 8.6 is a visual aid to identify where the outliers take place. In total, the model contains 115 outliers. Outliers are not a bad thing itself, as any datasets are expected to have a few extreme values. With over 8000 observations, we would expect to find 24 outliers, as 99.7 percent should theoretically be within three and a half standard deviation from the mean.

We have to careful when dealing with outliers, as we do not want to manipulate the data, reducing the variance, and making it seem more stable than in reality. There should be no patterns across the outliers, being spread randomly across the time series. From figure 8.6, there is evidence of clustering around 2008, when the financial crisis occurred. We thus remove the financial crisis between 15. September 2008 to 9. March 2009 from the data and create a new model excluding it.

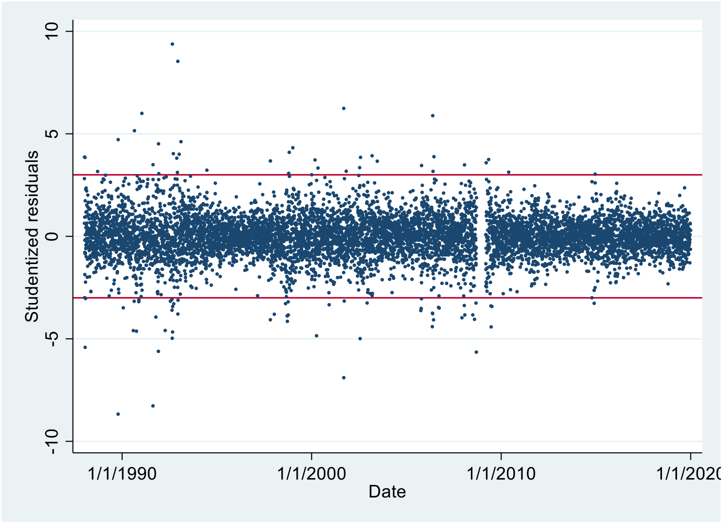


Figure 8.7: Scatter plot of residuals from the model excluding the financial crisis

A new regression is run with a dummy variable to indicate if the financial crisis is present. The same tests are conducted with the same findings. Heteroscedasticity is present, error terms are autoregressive, and error terms are not normally distributed. The beta coefficient is approximately the same, with all variables taking the same p-value as well. Full results can be found in appendix I. The effect of removing these outliers are minimal and seem not to change the significance nor the coefficients of included variables. We, therefore, continue to use the model with the financial crisis included to make a few alterations as possible.

8.1.4. Summary Analysis One

In the initial model, running over the whole period, all independent variables were significant. The time issue has been mentioned previously, due to the difference in trading hours. This got confirmed when analyzed further by including lagged variables. The one-time lagged SP500 variable became significant, taking an even higher beta coefficient than the coetaneous SP500. For Crude oil, we found the one-time lagged variable to be significant, but with a relatively low beta coefficient. Besides, the USDNOK currency was also added as a one day lagged variable while the USDNOK itself was dropped from the model, due to becoming insignificant.

When controlling the model's statistic properties, it was found to be heteroscedastic in its error terms, resulting in NIBOR being dropped when running the regression robust. There was no normality in error terms either. This was further looked in to, with no effect regarding dropping the financial crisis from the model. We also concluded that the assumption of normality in the error term would not impact the model or its relevance. In the end, the model that was best in can be summarized in the following equation:

$$OSEAX = 0.0002 + 0.1081SP500 + 0.1442SP500_{t-1} + 0.0916Brent + 0.0204Brent_{t-1} + 0.6959FTSEAX + 0.0416USDNOK_{t-1}$$

8.2. Analysis Two: 1988 to 2008

We repeat most of the steps from analysis one and apply them for this analysis. We re-start the process from the beginning, looking at the initial model again, but only for the first subpart.

Model Summary			
R	**R-square**	**Adjusted R-square**	**Std.Error of estimate**
0.5462	0.3071	0.3064	0.0100

Table 8.10: Summary of the initial model, first subpart

ANOVA					
	df	**Sum Of Squares**	**Mean Square**	**F**	**Sig.**
Model	5	0.2244	0.0448	458.87	0
Residual	5177	0.5064	0.0001		
Total	5182	0.7308			

Table 8.11: ANOVA output from the initial model, subpart one.

Coefficients					
	Coefficients	**Standard Error**	**t Stat**	**P-value**	**VIF**
Intercept	0.0003	0.0001	2.41	0.016	
SP500	0.0100	0.0151	0.6600	0.5110	1.21
Brent	0.0596	0.0059	10.0400	0.0000	1.00
FTSEAX	0.6748	0.0161	41.9100	0.0000	1.21
NIBOR	-0.0015	0.0005	-3.0500	0.0020	1.00
USDNOK	0.0899	0.0210	4.2800	0.0000	1.01

Table 8.12: Coefficient table from the regression of the initial model, subpart one.

Taking a brief look at the output, we note that the model has an adjusted r-squared of 30.64 percent. Lower than we found in analysis one, but still a respectable explanatory power. The lower explanatory power for subpart one indicates a much higher in subpart two. An indication that the theory of subsets within the data is true, making the model a better fit over some subperiods. As with analysis one, the F-test is significant, indicating the joint variables make a model that can explain changes in OSE. In contrast to the initial model for analysis one, all variables are not significant for analysis two with S&P 500 having a p-value of 0.51. The Brent crude coefficient is almost half then what we saw in the full period. At the same time, the VIF tests give slightly lower values, indicating there is less correlation across independent variables at this period.

8.2.1. Optimal Lag Selection

We learned from analysis one that including lagged variables, can improve the model. The insignificant S&P 500 variable will still be included when adding lags to the model, as we saw from analysis one that the inclusion of lagged S&P 500 was significant. Once again, we extend the model to include two lags for all independent variables, and by trial and error, find the optimal model.

After a trial and error approach, much of what was found in analysis one is evident. A lagged S&P 500 is once again significant, resulting in the coetaneous becoming significant as well, but with a smaller coefficient. For the remainder of variables, only a lagged Brent crude oil is significant, with a coefficient approximately half of the coetaneous.

	Adj-R^2	AIC
No lags	0.3064	-33136.72
SP500 lagged	0.3401	-33396.86
2 lags all variables	0.3444	-33421.29
2 lags all variables (ex. Insignificant)	0.3444	-33429.64
2 lags all variables (ex. Insignificant and SP500)	0.3084	-33153.86

Table 8.13: How lagged variables affect the model for the first period.

Table 8.13 presents some of the adjusted r-squared and AIC scores from the trial and error approach. As with analysis one, the lagged S&P 500 is the most influential variable for the model. Based on the figures in table 8.13, the best model consists of including up to two lags and then stripping away insignificant values. The model is presented below and will be used as presented when controlling the OLS assumptions.

The new model can then be formalized and summed up as:

$$OSEAX = \beta_0 + \beta_1 SP500 + \beta_2 SP500_{t-1} + \beta_3 Brent + \beta_4 Brent_{t-1} + \beta_5 FTSEAX + \beta_6 NIBOR + \beta_7 USDNOK$$

Model Summary			
R	**R-square**	**Adjusted R-square**	**Std.Error of estimate**
0.5802	0.3376	0.3366	0.0097

Table 8.14: Summary of model extended with lagged variables, period 1

Coefficients					
	Coefficients	**Standard Error**	**t Stat**	**P-value**	**VIF**
Intercept	0.0003	0.0001	1.86	0.0630	
SP500	0.0497	0.0149	3.3300	0.0010	1.24
LagSP500	0.2401	0.0142	16.8800	0.0000	1.13
Brent	0.0588	0.0058	10.1800	0.0000	1.08
LagBrent	0.0341	0.0058	5.9000	0.0000	1.00
FTSEAX	0.5796	0.0166	34.8800	0.0000	1.36
NIBOR	-0.0015	0.0005	-3.1100	0.0020	1.00
USDNOK	0.0948	0.0205	4.6300	0.0000	1.01

Table 8.15: Coefficient table from the extended model, including lagged variables, period one.

8.2.2. Controlling OLS Assumptions

We can no longer assume that the OLS assumption still holds, as we have a new model, including different variables. The number of observations is also reduced by 3000 observations and might affect the validity of the assumptions. Due to the nature of the first three assumptions, we can still assume they hold, and will further focus on the last three assumptions.

4. Homoscedasticity

The relationship between predicted values and studentized residuals is presented graphicly in figure 8.8.

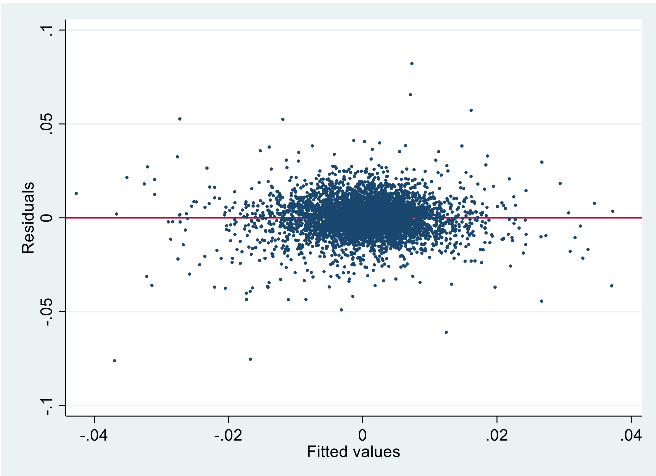


Figure 8.8: Residual vs. Fitted values.

From the plot, there seems to be more spread now, relative to analysis one. There are no clear and apparent patterns that can be pointed out immediately. As stated earlier, we supplement the blot with a numerical test and conduct the Whites test for homoscedasticity.

Statistical Test	H_0	Chi-squared	df	Critical Value	Conclusion
White's test	Homoscedasticity	1060.64	35	60.481	Reject H_0

Table 8.16: Results from White's test for homoscedasticity.

From table 8.16 with the results from a Whites test, we reject the null of heteroscedasticity as the Chi-squared value is well above the critical value. As before, we apply the HC standard error by running the regression as robust and get new standard errors.

Coefficients					
	Coefficients	**Standard Error**	**t Stat**	**P-value**	**VIF**
Intercept	0.0003	0.0001	1.85	0.0650	
SP500	0.0497	0.0211	2.3500	0.0190	1.24
LagSP500	0.2401	0.0184	13.0800	0.0000	1.13
Brent	0.0588	0.0092	6.4200	0.0000	1.08
LagBrent	0.0341	0.0059	5.7700	0.0000	1.00
FTSEAX	0.5796	0.0278	20.8400	0.0000	1.36
NIBOR	-0.0015	0.0018	-0.7900	0.4290	1.00
USDNOK	0.0948	0.0284	3.3400	0.0010	1.01

Table 8.17: Coefficient table with robust standard errors.

The model is not much affected by running with robust standard error. Just as in analysis one, we get insignificant result for the NIBOR variable, with a p-value 0.429. Based on this, the NIBOR is dropped from the model. All other variables are unaffected, but with a slightly higher p-value for both lagged S&P 500 and USDNOK currency. Full results with NIBOR excluded can be found in Appendix I.

5. Autocorrelation

We repeat the steps from analysis on, and plot the residuals in both a time plot and the ACF-plot. Giving us a petter view for any patterns regarding the error term, and checking for consistency over the period.

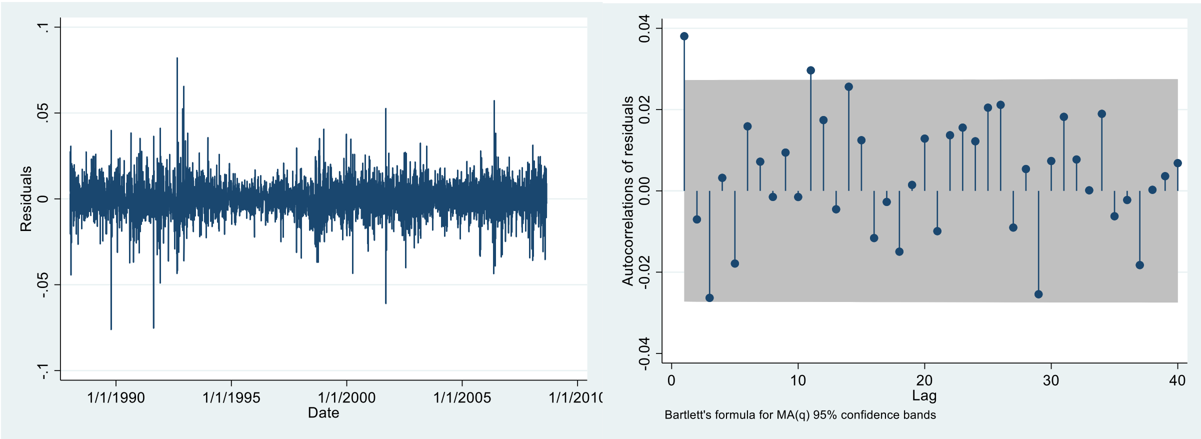


Figure 8.9: Residuals over time for subpart one (LHS).

Figure 8.10: ACF-plot of residuals for subpart one (RHS).

From the time plot of residuals, we can check if residuals are consistent over the period. From figure 8.9, residuals emerge close to zero, with some periods with higher residuals, but no clear patterns. Implying no presence of autocorrelation for subpart one. Figure 8.10 demonstrates the autocorrelation effect over different lags. As with analysis one, there are no obvious patterns, while staying in the range of the 95 percent confidence interval, except for the first-order lag. From the ACF-plot, we suspect that the residuals follow an AR(1) process.

We supplement the plots with the numerical Durbin Watson test, to check for autocorrelation of the first order. The DW-test gives us a test statistic of 1.924. We look at the Savin and White table to find dL and dU.

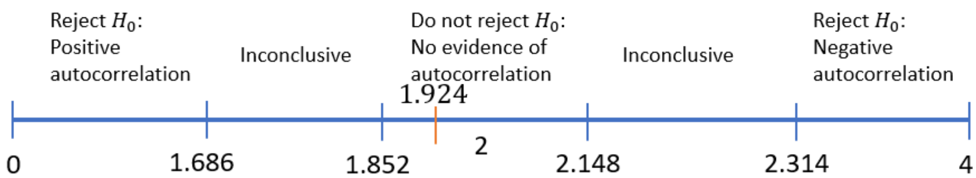


Figure 8.11: Results from the Durbin Watson test, with dL and dU values.

As within analysis one, DW-test suggests that there is no autocorrelation at lag 1, as can be seen in figure 8.11. The test is slightly leaned towards positive autocorrelation, but still well within the range of not rejecting the null hypothesis.

Due to the heteroscedasticity found in the previous assumption, we want to supplement with the CH-test, as was done in analysis one. Allowing us to control for the presence of heteroscedasticity and check for further lags. To be consistent in testing, we check for up til three lags in this analysis as well.

Cumby-Huizinga test			
Lag	**Chi2**	**df**	**p-value**
1	2.893	1	0.0890
2	0.012	1	0.9127
3	1.805	1	0.1792

Table 8.18: Results from CH-test for autocorrelation, subpart one.

As can be seen from table 8.18, an AR (1) process just falls short from being significant. We cannot reject the null hypothesis of any autocorrelation for any of the three lags. We see that there are signs of autocorrelation, but it does not meet the 95 percent confidence interval. To be extra cautious, using robust standard errors will not affect anything other than standard errors. We then changed the White's robust standard error to HAC standard error to see if this would have any effect on the significance of the variables. Applying the HAC robust standard error did not have any effect if there should be any case of autocorrelation. Although, we conclude that no autocorrelation is present, based on the plots and test above.

Moving onwards for analysis two, we keep running the standard OLS regression as robust with Whites standard errors. The model still looks the same after NIBOR was omitted due to the heteroscedasticity test.

6. Normality in error terms

We plot the QQ-plot as a visual representation of the residuals versus the theoretical distribution of residuals.

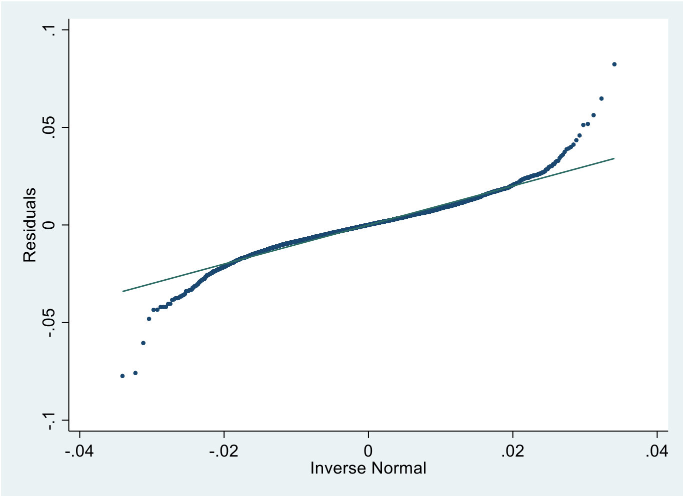


Figure 8.12: Quantile plot of residual predicted errors.

Most of the predicted error terms fit well with the theoretical distribution. However, some observations are deviating far from the fitted line, both positive and negative values. Suggesting that we have fat tails among the error term distribution and may, in fact, not be normally distributed. Conducting the Jaque-Bera tests confirms this, giving a chi-squared statistic of 6076, well above any critical values. Thus we reject the null hypothesis of normality in the error terms. Using the same arguments presented in analysis one, we are not particularly concerned with the rejection of the normality of error terms. As with previous analysis, we appeal to the CLM and further assume that this assumption is satisfied in analysis two.

8.2.3. Outliers

We build upon the findings from normality in the error term. Revealing that some values were far from the theoretical distribution. To explore further, we plot the studentized residuals to learn more about each observation to find out when they occure and if there are any clustering patterns.

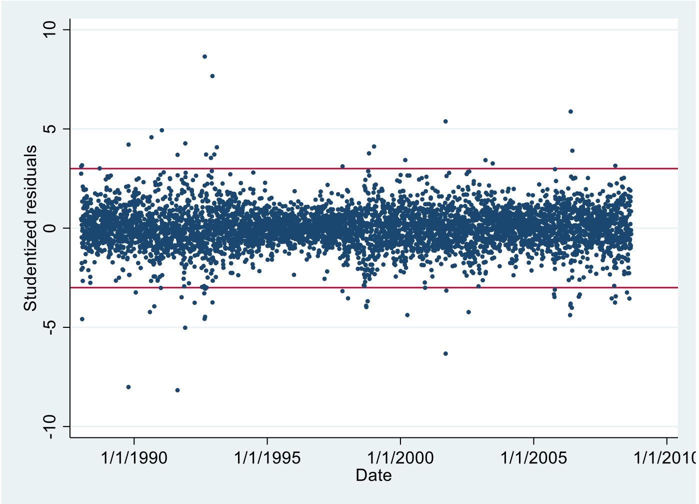


Figure 8.13: Scatter plot of the Studentized residual over time

In total, we have 68 observations that fall within our definition as an outlier, having a standard deviation above three times the mean. Theoretically, one could expect to find about 15 extreme values in a sample equal to this size. Looking at the plots, there are no apparent clustering of extreme values, as we had in the analysis one. We want to be careful about removing any single outlier, as they are scattered randomly across the period. We know from the variables used in the model, that they often take on extreme values, due to the high kurtosis. By removing outliers, we would manipulate the data by reducing the variance below what is true. We will thus not make any adjustments regarding outliers.

8.2.4. Summary Analysis Two

In the starting model, all variables were significant except for the S&P 500. The time difference became even more evident in this analysis, relative to analysis one. As S&P 500 was not significant in the original model for this analysis but became significant when including a lag. The only additional adjustments that were made when finding optimal lag was the inclusion of the lagged brent. In total, the model was extended with two lagged variables before checking the OLS assumptions.

When controlling the OLS assumption, much of the same issues as in analysis one was present. The error terms were heteroscedastic, resulting in dropping NIBOR from the model. However, there were indications of an AR (1) process, but the null hypothesis could not be rejected on the 5 percent significance level. Additionally, as with analysis one, the assumption of normality of error terms was rejected. Due to the rejection of normality, we further reviewed the high amount of outliers causing non-normality. No clustering was found, leading to no adjustments being made. Due to the large sample size, the assumption was regarded as satisfied, and OLS control overall was fulfilled.

As a result of the analysis, we end up with a more fitting model that is satisfied as BLUE. The model can be summarized in the equation below:

$$OSEAX = 0.0003 + 0.0492SP500 + 0.2401SP500_{t-1} + 0.0589Brent + 0.0342Brent_{t-1} + 0.5811FTSEAX + 0.0986USDNOK$$

8.3. Analysis Three: 2009 to 2019

We now take a look at the model covering the second subpart, starting after the financial crisis. The initial model can be seen as summarized below:

Model Summary			
R	**R-square**	**Adjusted R-square**	**Std.Error of estimate**
0.6861	0.6564	0.6557	0.0069

Table 8.19: Summary of the initial model, period two.

ANOVA					
	df	**Sum Of Squares**	**Mean Square**	**F**	**Sig.**
Model	5	0.2542	0.0508	1051.27	0
Residual	2752	0.1330	0.00005		
Total	2757	0.3872			

Table 8.20: ANOVA output from initial mode, subpart two.

Coefficients					
	Coefficients	**Standard Error**	**t Stat**	**P-value**	**VIF**
Intercept	0.0002	0.0001	1.68	0.092	
SP500	0.1092	0.0167	6.52	0.0000	1.65
Brent	0.1218	0.0073	16.78	0.0000	1.24
FTSEAX	0.7608	0.0191	39.86	0.0000	1.89
NIBOR	-0.0161	0.0047	-3.42	0.001	1.07
USDNOK	-0.1115	0.0192	-5.81	0.0000	1.22

Table 8.21: Coefficient table from the initial model, sub-period two.

By taking a quick look at the summary, we note that the model has an adjusted r-squared above twice what we saw in the previous analysis. Indicating that overall, the independent variables are more relevant in this later period over the recent decade. Splitting the full period in two has affected positively on fitting the data. With the high explanatory power, it follows that all variables included are significant. All the beta coefficients take a higher absolute value, thus having a greater effect on OSEAX. At the same time, the VIF value is also increased, indicating that there is a higher pairwise correlation across the independent variables.

8.3.1. Optimal Lag Selection

We follow the same procedure as in previous analysis and examine the effect of including lagged variables. As a starting point, we include two lags of all independent variables and build on this by trial and error to find the optimal combination,

After running several regressions, we can find the optimal combination by excluding insignificant variables and comparing the AIC score along with adjusted r-squared. The lagged S&P 500 effect is not as indicative as in previous analysis. However, it is still significant, along with the second lag. Apart from the S&P 500, USDNOK lagged variable also become significant, but with the opposite sign of the coetaneous.

	Adj-R^2	**AIC**
No lags	0.6557	-19573.4
SP500 lagged	0.6556	-19567.3
2 lags all variables	0.6572	-19569.6
2 lags all variables (ex. Insignificant)	0.6565	-19570.8
2 lags all variables (ex. Insignificant and SP500)	0.6558	-19568.5

Table 8.22: Summary trial and error for optimal lag.

A selection of the regression run is presented in table 8.22, highlighting the effect of the S&P 500 lagged. By comparing to table 8.13, it becomes evident that lagging the S&P 500 does not have as much impact on subpart two. However, USDNOK has become more impactful, and with it, a significant one time lag. For the other variables, we found no significant lagged variables. The model has thus been extended to include three more variables in the lagged form. A summary of the extended model is presented below and will be tested when controlling OLS assumptions.

$$OSEAX = \beta_0 + \beta_1 SP500 + \beta_2 SP500_{t-1} + \beta_3 SP500_{t-2} + \beta_4 Brent + \beta_5 FTSEAX + \beta_6 NIBOR + \beta_7 USDNOK + \beta_8 USDNOK_{t-1}$$

Model Summary			
R	**R-square**	**Adjusted R-square**	**Std.Error of estimate**
0.6981	0.6575	0.6565	0. 0069

Table 8.23: Summary of model extended with lagged variables, period two.

Coefficients					
	Coefficients	**Standard Error**	**t Stat**	**P-value**	**VIF**
Intercept	0.0002	0.0001	1.35	0.178	
SP500	0.1195	0.0172	6.94	0.000	1.75
LagSP500	0.0343	0.0143	2.40	0.017	1.21
Lag2SP500	0.0313	0.0133	2.34	0.019	1.06
Brent	0.1197	0.0072	16.46	0.000	1.24
FTSEAX	0.7499	0.0196	38.24	0.000	2.00
NIBOR	-0.0151	0.0047	-3.18	0.001	1.08
USDNOK	-0.1038	0.0195	-5.31	0.000	1.27
LagUSDNOK	0.0534	0.0184	2.91	0.004	1.12

Table 8.24: Coefficient table from the extended model, including lagged variables, period one.

8.3.2. Controlling OLS Assumptions

As with analysis two, we have to conduct the last three assumptions. Due to the nature of the data, variables, and model specification, we can assume that the first three assumption still holds, and move on straight to assumption four.

4. Homoskedasticity

We use the same approach as in analysis one and two, and plot the residuals against the fitted values.

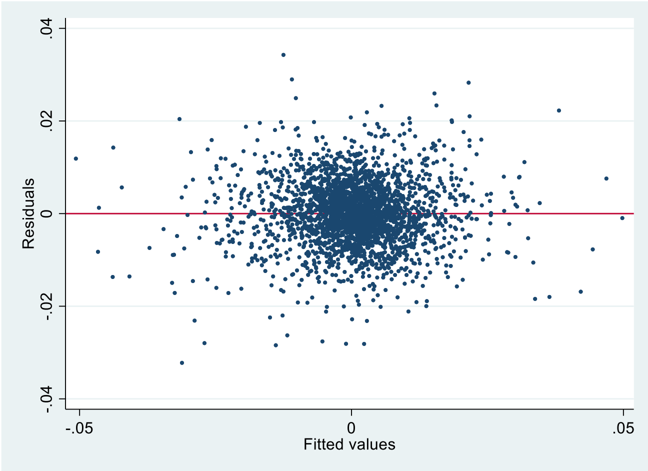


Figure 8.14: Residuals vs. Fitted values

The plot is similar to plots in the previous analysis, with most observations centered. Different from the previous plots, there appears to be even more spread between residuals and fitted values in this model. Before we make any conclusions, we supplement with a numerical test.

Statistical Test	**H_0**	**Chi-squared**	**df**	**Critical Value**	**Conclusion**
White's test	Homoscedasticity	411.11	44	60.48	Reject H_0

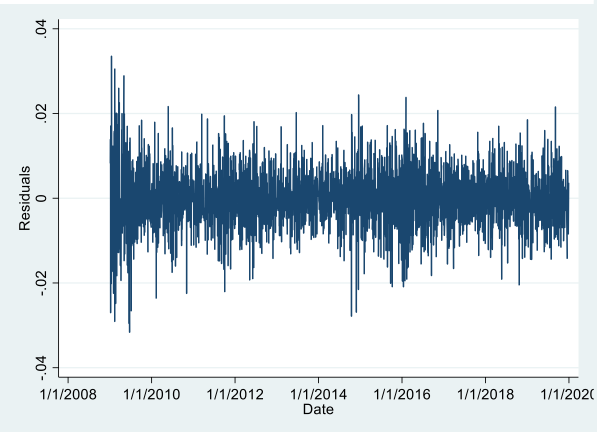
Table 8.25: Results from Whites test for homoscedasticity.

From table 8.16 with the results from a Whites test, we reject the null of heteroscedasticity as the Chi-squared value is well above the critical value. As before, we apply the HC standard error by running the regression as robust and get new standard errors.

As with the other analysis, we can see from the results in table 8.25 that the null hypothesis of homoscedasticity is rejected, and the error terms do not have constant variance. Applying the HC-standard error due to heteroscedasticity affects several of the variables. All three lagged variables that were added when finding optimal lag falls outside the 95 percent confidence interval. Resulting in dropping all three, and we are back to the initial model. Full results of including and excluding the lagged variables with HC-standard errors can be found in Appendix I.

5. Autocorrelation

We follow the prosses from the above analysis and look for autocorrelation in the model by plotting both residuals over time and the ACF-plot.



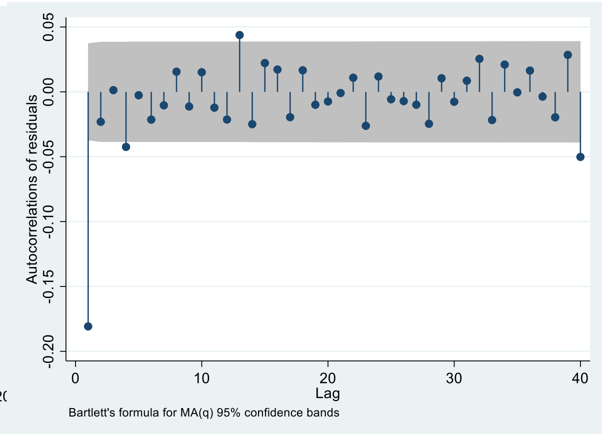


Figure 8.15: Residuals over time, subpart two (LHS).

Figure 8.16: ACF-plot of residuals, subpart two (RHS).

From the residual plot in figure 8.15, we see that the residuals fluctuate in this model, compared to model two. The residuals both fluctuate further away from zero and is more uneven than what we have observed in previous models. From the ACF-plot in figure 8.16, there is a strong indication of a negative autocorrelation of the first order, suggesting that we have an AR (1) process.

A DW-test is supplemented with, as a numerical test to formalize the null and alternative hypothesis. From the DW-test, we obtain a test statistic of 2.3609. We obtain the dL and dU from the Savin and Whites table.

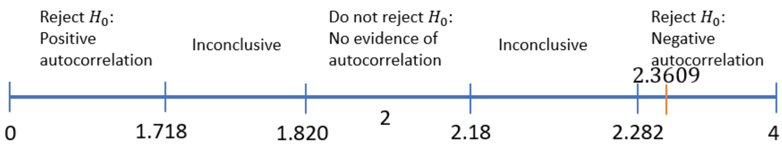


Figure 8.17: Results from DB-test illustrated.

As opposed to the previous analysis, the DW-test can confirm what was suggested from the ACF-plot, a negative correlation of the first order. Since we again in this model have heteroscedasticity present, we supplement with the CH-test, both to account for the HC-robust standard errors and examine autocorrelation at higher orders.

Cumby-Huizinga test			
Lag	**Chi2**	**df**	**p-value**
1	54.208	1	0
2	1.162	1	0.2810
3	0.003	1	0.9555

Table 8.26: Results from CH-test for autocorrelation, subpart two.

The CH-test can significantly reject the null hypothesis of no autocorrelation of the first order for this model, as can be seen in table 8.26, with a p-value equal to zero. Now that both heteroscedasticity and autocorrelation is present, an autoregressive conditional heteroscedasticity (ARCH) model would have been a fitting model to apply. As we have previously used the standard OLS model, with the robust standard errors, we will continue to use this method to ensure consistency across the models. We can no longer use the HC-standard errors, as we have to account for the autocorrelation, and apply HAC-robust standard error instead. It has also been argued that if the sample size is sufficiently large, using HEC robust standard errors will give apporxomatly the same estimates as an ARCH-model (Engle, 2011).

The CH-test serves as a good supplement, as it confirms the AR (1) process shown in both DW-test and the negative autocorrelation we found in figure 8.16. The null hypothesis of no serial correlation is rejected, and we conclude that the error terms follow an AR (1) process.

Since we assume the presence of an AR (1) process with heteroscedasticity present, we must change the robust standard errors. Previously we used the White's robust standard errors to account for heteroscedasticity. We will now use the Newey-West standard error, that not only corrects for heteroscedasticity but also autocorrelation for a specified lag. Using the Newey-West standard errors does not affect the significance of any of the variables in becoming insignificant. All independent variables are still significant for the 95 percent confidence interval. Thus, our model is unaffected by controlling for the presence of an AR (1) process. The full result using Newey-West standard errors can be found in appendix I.

6. Normality in error terms

To control the assumption of normality in error terms, we plot the QQ-plot comparing residuals of the model to the theoretical distribution of residuals.

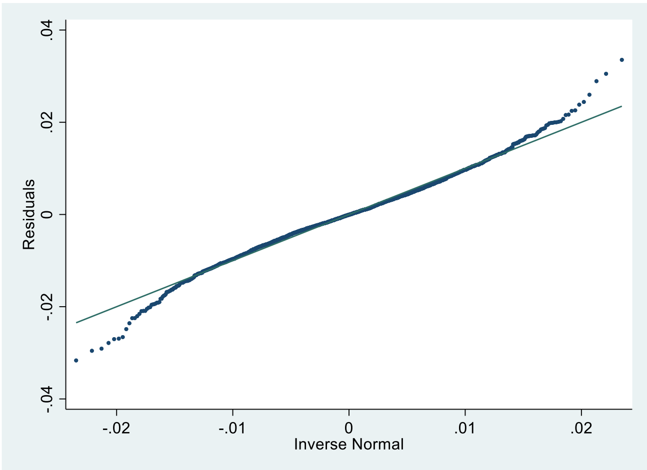


Figure 8.18: Quantile plot of predicted errors from the model for subpart two.

Most of the predicted error terms fit well with the theoretical distribution. However, some observations are deviating far from the fitted line, both positive and negative values. Suggesting that we have fat tails among the error term distribution and may, in fact, not be normally distributed. Conducting the Jaque-Bera tests confirms this, giving a chi-squared statistic of 6076, well above any critical values. Thus we reject the null hypothesis of normality in the error terms. Using the same arguments presented in analysis one, we are not particularly concerned with the rejection of the normality of error terms. As with previous analysis, we appeal to the CLM and further assume that this assumption is satisfied in analysis two.

The plot in figure 8.18 shows what we have seen in the previous analysis so far. Most of the observation fir well with the theoretical distributions, although, we have increased extreme values in both ends. Indicating that the error terms are not normally distributed. A JB-test is conducted to formalize the null hypothesis and alternative hypothesis. We acquire a chi-squared test statistic of 221.4, well above the critical value for any standard significance. We can constate that the error terms are indeed not normally distributed, by rejecting the null hypothesis. As with the previous analysis, we regard this assumption satisfied by appealing to the CLT. We thus end up with a model that satisfies BLUE.

8.3.3. Outliers

Due to the rejection of normality in error terms, we further examine the outliers in the model. We Plot the studentized residuals from the model across time, to gain an overview of when outliers occur.

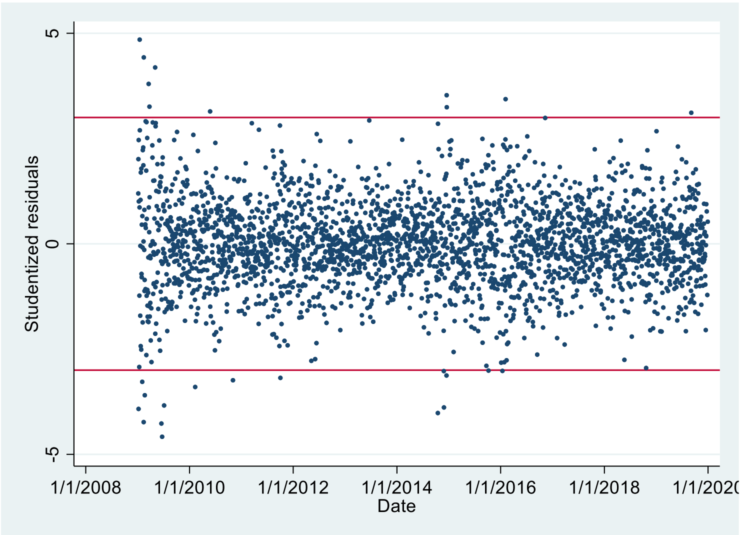


Figure 8.19: Scatter plot of the Studentized residual over time.

For subpart two, there are 26 observations defined as outliers with a standard deviation at or above three times the mean. The theoretically expected value would be around eight observations. The outliers are distributed randomly across the period, with no clear pattern, except for a slight clustering at the start, right after the financial crisis. Figure 8.19 indicates that subpart two has been more volatile than the preceding, with more fluctuations of the residuals. Making any alterations would give a false representation by reducing the variance of period two. No changes will be made towards the outliers in subpart two.

8.3.4. Introducing RENIXX

For subpart two, we will include RENIXX to the model. This will serve as the fourth model of the thesis and be an extension from model three. RENIXX is used as a benchmark towards the renewable energy sector presented earlier in this thesis. We will then be able to compare the renewable industry against the oil price, affecting the oil industry.

We build upon the model in analysis three after the OLS assumptions were controlled for, and add RENIXX. The model, including RENIXX, has undergone the same OLS controlling procedure as previous models. To avoid repetition, the full results from the test can be found in appendix I. As a result of the test, the model was found best to explain OSEAX when including $RENIXX_{t-1}$. The model was also found to be heteroscedastic and follow an AR (1) process. A comparison of the model, including and excluding RENIXX, can be seen in table 8.27.

	Int.	SP500	Brent	FTSEAX	NIBOR	USDNOK	RENIXX	LagRENIXX	Adj R^2	Std.Error	F
Pre RENIXX	0.0001	0.1092	0.1218	0.7608	-0.0161	-0.1115			0.6557	0.0070	613.49
RENIXX inc.	0.0003	0.0860	0.1193	0.6819	-0.0164	-0.1024	0.1140	-0.0322	0.6691	0.0068	488.09

Table 8.27: A comparison between regression with and without RENIXX included in the model

As can be seen in the table above, RENIXX appears to be a valuable addition to the model, capturing relevant information. The adjusted R-squared has increased, as well as RENIXX has a beta coefficient equal to the Brent crude. The Brent crude beta coefficient remains almost unaffected by the inclusion of RENIXX. Indication no correlation between the two, and they are no overlapping to what the variables offer to the model.

8.3.5. Summary Analysis Three

For the initial model running over subpart two, all variables were found to be significant. Three more additional lagged variables were found significant and added. When controlling OLS assumption, due to heteroscedasticity and autocorrelation being present, the same lagged variables were stripped from the model, due to being insignificant when using HAC robust standard errors. The model fulfilling BLUE, where the model including the original variables.

When further examining outliers after the rejection of the normality in error term assumption, we found more outliers compared to the theoretical normal distribution. However, the financial crisis has already been excluded from the subpart, as the time series starting in January 2009. No alterations were made regarding the outliers, as we want it to reflect the true variance of the period. The final model for analysis three can be seen as below.

$$OSEAX = 0.0001 + 0.1092SP500 + 0.1218Brent + 0.7608FTSEAX - 0.0161NIBOR - 0.1115USDNOK$$

The model was then extended to include RENIXX as a variable. This will not be added to the model found above, instead presented as a new model. The model was found to have the same statistical properties as in analysis three before the inclusion of RENIXX. By correcting for heteroscedasticity and autocorrelation by using the HAC robust standard errors, we regarded the OLS assumptions to be satisfied and the model to be BLUE. The inclusion of RENIXX leads us to believe that it is a valuable addition to the model and has a surprisingly high effect on OSEAX. Further discussion will be held in the next subpart of the thesis. The model, including RENIXX, can be summarized as:

$$OSEAX = 0.0003 + 0.0860SP500 + 0.1193Brent + 0.6819FTSEAX - 0.0164NIBOR - 0.1024USDNOK + 0.1140RENIXX - 0.0322RENIXX_{t-1}$$

8.4. Model Prediction

This section will present a graphical illustration of how the different models are able to predict OSEAX, using the beta coefficient found in the analysis, and compare them with the actual development of OSEAX during the same period. For each period, the model uses predicted yesterday's closing price multiplied with today's return based on the model. Except for the first observation, this must be an actual value of OSEAX.

$$\widehat{OSEAX} = \widehat{OSEAX}_{t-1} * e^{\sum_0^S (\beta_s x_s)}$$

$$S \in [0, S]$$

Where S represents the individual variables for each model, this is to avoid complex notation for each variable within each model.

For models estimating the second subpart, they will take a take the observed value from 02.01.2009, and then predict from this point on using the models. They will also be backtested by dividing today's estimated value by yesterday's increase to get the value from the day before. This will give a better visual representation of the models, showing deviations both before and after the starting point making it easy to compare with the actual development of OSEAX. Predicted values may wander off and end up far from the observed values. If the model wanders off, it does not necessarily mean the model itself is a poor estimate, but there might be a shock present in the observed value, that the model is not able to capture. Leading it to a different level, but with much of the same movements. Since we know the variables included are more sensitive to negative shocks, from the descriptive data, the models will likely be upwards biased.

8.4.1. Predictions January 1988 – December 2019

For the full period, we will present several models. Both the one estimated in analysis one, as well as combining subpart one and subpart two to combined explain the full period.

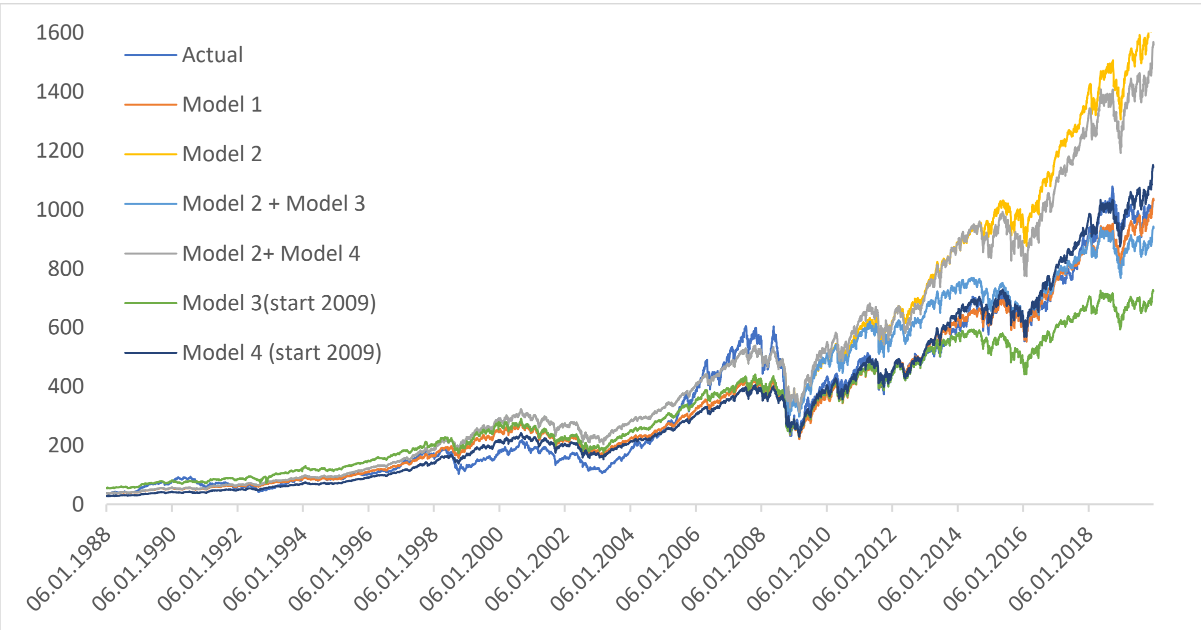


Figure 8.20: Predicted development over the full period.

In the graph above, we have used a total of four different models. The models are defined as follows

- Model 1: Results from analysis one, regression based on the full period
- Model 2: Results from analysis two, regression based on subpart one
- Model 3: Results from analysis three, regression based on subpart two
- Model 4: Results from analysis three, regression based on subpart two with RENIXX

As a general note, it is evident that all models are prone to negative shocks, as anticipated. All models fit actual data well, until the first negative shock in 1999. Model 2 gives better estimates relative to model 1 up to the financial crisis, just as model 3 for subpart 2 makes better estimates relative to model 1 after the financial crisis. Indicating that the models created for their respective subparts serve as a better estimation compared to model 1 estimated on the whole period.

8.4.2. Predictions January 1988 – August 2008

Figure 8.21 shows the predicted values of OSEAX from models based on analysis one and two.

Figure 8.21: Estimated development of OSEAX, results from analysis one and two.

As was mentioned, the estimated models tend to be upwards biased. Both models lay on a higher level after the shock around 1999. The models follow much of the same development from there on, with the same swings, but at a higher level. From 2003 and the rest of the analysis, model two captures the rise in the stock market and follow the actual data well. While model 1 falls short, not being able to capture the increase as efficient.

8.4.3. Predictions January 2009 – December 2019

Figure 8.22 shows the predicted values of OSEAX from models based on analysis one and two.

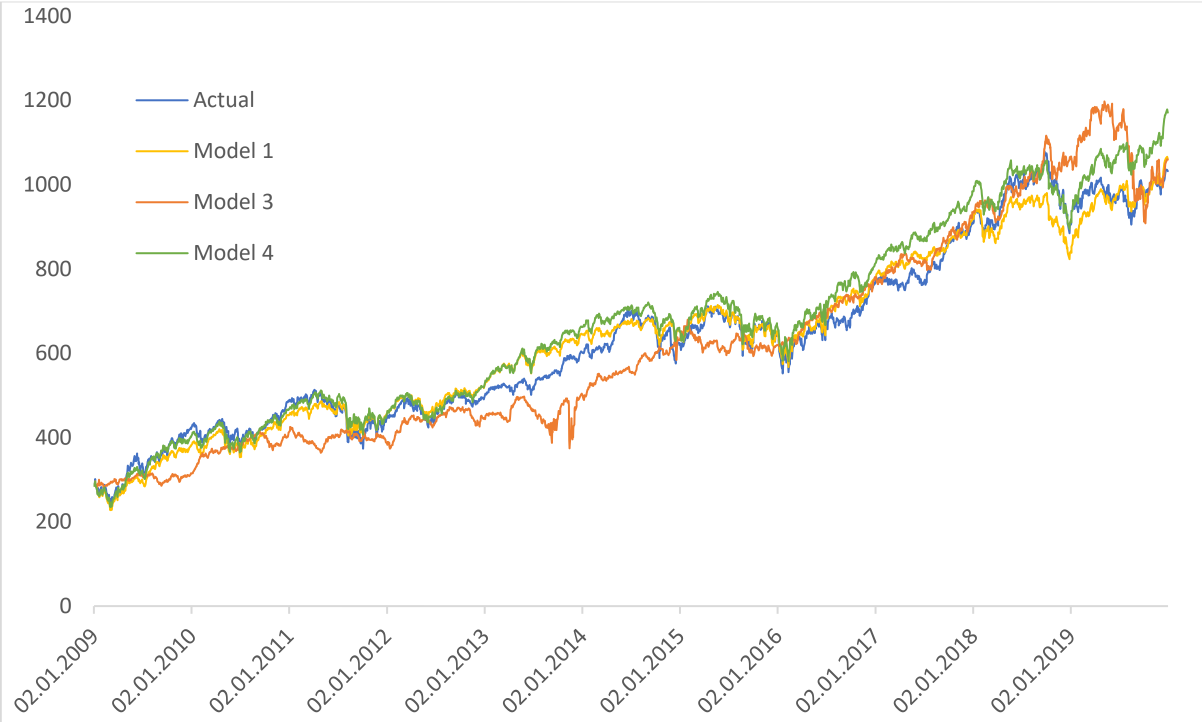


Figure 8.22: Estimated development of OSEAX, results from analysis two and three.

During a time of steadily increase of OSEAX, our models can estimate this increase well. All models follow the actual development of OSEAX well, with model four being more efficient compared to model three. Model three is inconsistent in its estimates, with often underestimating, then overestimating for other periods within the subpart. Towards the end of the period, we can see that model four starts to wander off, after following the actual data closely for a more extended period. This likely due to the green energy effect that has been observed for the renewable sector and outperforming the rest of the market.

9. Findings and Discussions

In the following section, we will take an in-depth look at the results from the different regression models. Separately, the various variables will be emphasized and discussed. Ultimately, the interaction between the variables will be interpreted.

9.1. Full results

Table 9.1 presents the beta coefficients for each variable, in all the respective regression models. The adjusted r squared value for the particular models are also provided. Prevailing for the various models is that both the international financial markets and the oil price compose a significant effect on OSEAX. However, these are all subject to variations over different periods.

Coefficients				
	Analysis 1	**Analysis 2**	**Analysis 3**	**Analysis 3, RENIXX**
Intercept	0.0002	0.0003	0.0001	0.0003
SP500	0.1081	0.0492	0.1092	0.0860
LagSP500	0.1442	0.2401		
Brent	0.0917	0.0590	0.1218	0.1193
LagBrent	0.0204	0.0342		
FTSEAX	0.6959	0.5811	0.7608	0.6819
NIBOR			-0.0161	-0.0164
USDNOK		0.0986	-0.1115	-0.1024
LagUSDNOK	0.0416			
RENIXX				0.1140
LagRENIXX				-0.0322
Adj R-squared	0.4711	0.3440	0.6557	0.6691

Table 9.1: Summary of the respective regression models

On a general note, we see that models for subpart two, include less lagged variables. Furthermore, the models in the later subparts are able to capture almost double the variance of the dependent variable relative to subpart one.

9.1.1. S&P 500

S&P 500 was included in the analysis as a measurement for the global economy. From the analysis, it appears evident that the S&P 500 and FTSEAX are competing with regard to the explanatory power of OSEAX. This was initially expected, as they both constitute major indexes, both serving as measures for the international financial markets. However, the S&P 500 represents the variable with the greatest difference in trading hours compared to OSEAX. This feature is reasonable concerning the significant lagged coefficient for model 1 and 2. Furthermore, the S&P 500 variable is significant for all estimated regression models, but its coefficient has changed over the periods. In subpart two, the variable receives a beta coefficient more than twice the value seen in subpart one.

Interestingly, the lagged variable of the S&P 500 is significant in subpart one but is removed from subpart two. A reasonable interpretation of this is that information travels faster during the subpart two, compared to subpart one. Factors like improved technology, together with globalization, may have led changes in the S&P 500 to be captured faster by OSEAX. With reference to EMH, this discovery is consistent in the sense that new information is more rapidly reflected in prices.

Interpreting the results, we find that a 1 percent increase in the S&P 500 results in a 0.11 percent increase for OSEAX, for the whole period. For the lagged variable, namely yesterday's movement, a similar increase will lead to a 0.145 percent increase in today's OSEAX. Even though these figures are considerable, they remain rather low compared to the other benchmark for the international financial markets, which is FTSEAX. If we were to drop FTSEAX from the model, most of its associated beta coefficient would then be divided and be assigned to both S&P 500 and its lagged variable. This would have been accomplished without affecting the other variables substantially. However, this would result in a weakened model based on the AIC score and adjusted R-squared. It was therefore chosen to include both variables in the model, without delving too much into the interpretation of the difference in their respective explanatory powers. Both should be interpreted as a measure of how OSEAX is dependent on the international financial market, and not so much on each individual stock index. It is then clear that from subpart one to subpart two, OSEAX has become increasingly dependent on the international stock market.

9.1.2. FTSEAX

Similarly to the S&P 500, FTSEAX was included to capture the effect from international markets, and more specifically, the European market. For all regression models conducted, FTSEAX is the most impactful independent variable with the greatest effect on OSEAX. Lagged variables for FTSEAX are throughout the models found insignificant, and therefore not included in any of them. This might be related to the compliance of trading hours between the two stock exchanges. With only a one hour difference in closing prices, it, therefore, seems like OSEAX manages to capture the information from FTSEAX during the same trading day.

For the whole period, a 1 percent rise in FTSEAX leads to a 0.70 percent increase in OSEAX. From subpart one to two, the effect experience an ample increase, where the coefficient moves from 0.58 to 0.76. Combined with a similar pattern for the S&P 500, it can be claimed that OSEAX has become increasingly more dependent on the international financial markets over the period. Compared to the S&P 500, the coefficients possessed by FTSEAX is substantial. As discussed under the interpretation of the S&P 500, these dynamics remain complex. However, there could be several possible explanations. Firstly, FTSEAX represents a superior indicator of the European financial market, narrowed down from the global market. In this regard, OSEAX is evidently more influenced by the European market, rather than the U.S market.

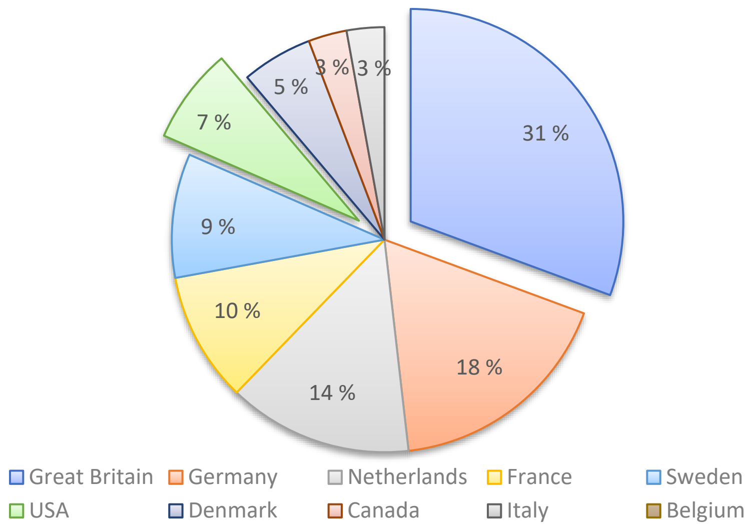


Figure 9.1: Top ten exporting destinations for Norway (Source: SSB, 2020).

Moreover, Norway represents a country highly reliant on export. Besides oil-related products, the country has also been increasingly dependent on the export of seafood (OEC, 2017). As can be seen from figure 9.1, Great Britain constitutes by far the largest export destination of Norway. Of the country's total export, 31 percent is exported to Great Britain, while only 7 percent is exported to the U.S. As the remaining top 10 countries are also mainly European countries, this demonstrates Norway's dependency on the European economy.

9.1.3. NIBOR

Besides the profits generated by a company and their expected future dividends, the discount rate is just as important for an investor valuing a stock. With respect to the formula for pricing stock in the theory section, a lower discount rate yields a higher value of equity. Accordingly, any stock market should be negatively correlated with an increasing interest rate. If the discount rates decline, the stock market should rise, all else being equal. NIBOR was correspondingly included in our model to investigate these dynamics.

Over the full period, NIBOR was omitted from the model, as it turned out to be insignificant. This is presumably due to the influence of subpart one, as it is also insignificant during this period. For subpart two, an increase of one percentage point would decrease OSEAX by 0.0161 percent. Therefore, the magnitude remains rather low, but the negative sign of the coefficient is consistent with financial theory.

Subpart one runs over a period where NIBOR both saw high levels and huge fluctuations, as can be seen in Appendix A. The main reason behind these immense levels is that Norges Bank used the interest rate to defend the currency. This implied that during difficult economic periods, the interest rates where risen (Kapital, 2018). With such an economic environment in force, it is reasonable to deduce that the considerable variations in NIBOR during subpart one causes the variable to become insignificant. However, in 2001 Norges Bank was given the mandate to target inflation (Ibid.). During subpart two, NIBOR has appeared more steadily and low, especially following the financial crisis. Consequently, in the later period, NIBOR is found significant.

9.1.4. USD/NOK

The USD/NOK variable is not included in the model for the full period. However, its lagged variable appeared significant and is therefore included. For both of the subparts, USD/NOK turned significant, but with opposite signs. In subperiod one, a one percent increase in the currency pair leads to an increase in OSEAX of 0.099 percent. This implies that a weakened NOK due to an increase in USD/NOK will accordingly lead to a rise in the Norwegian stock market.

In contrast, for subpart two, the coefficient has altered to a negative sign. In this case, a one percent increase in the currency pair will cause a decrease in OSEAX of 0.11 percent. Indicating that a weakened NOK leads to a decrease in the domestic exchange rate.

As oil is traded in USD, countries importing Norwegian oil will remain unaffected by changes in the level of the NOK. On the other hand, for Norwegian exporting companies, the level of the domestic currency is of great importance. As these companies will receive payments in USD, and subsequently translate this into NOK, they will benefit from weakened NOK. A weak NOK compared to USD will, therefore, increase the profit in the domestic currency, and vice versa. In this sense, a lower oil price can be compromised by a lower NOK. If we run the regression without the USD/NOK variable, the oil price coefficient remains the same. This suggests that changes in the oil price is not captured and rather assigned to the currency pair.

In subpart one, we observe a positive coefficient for USD/NOK, meaning that a weakened NOK will have a positive effect on the stock market. A possible explanation for this is that during this period, Norway steadily increased its export year by year (SSB, 2020x). A weakened NOK makes Norwegian products more competitive in the global market, leading to higher activity for Norwegian exporting companies. Export companies comprise a major part of the stock exchange, with Norway both exporting large values of petroleum and salmon. These dynamics, therefore, seems consistent with the discovery in this subperiod.

For subpart two, the sign has changed to be negative, which means that an increase in USD/NOK (depreciated NOK), has a negative effect on OSEAX. This can be translated to an increase of the NOK, which will lead to an increase in OSEAX. Increased stock markets signal confidence in a country's growth, creating positive investment opportunities. This finally leads to a higher demand for the Norwegian currency by foreign investors, strengthening the currency relative to others.

9.1.5. Oil Price

The primary independent variable for this thesis, Brent crude oil, is significant with a positive coefficient in all of the regression models conducted. Moreover, its lagged variable appear significant in model 1 and 2. For the entire period, a 1 percent increase in the oil price leads to an increase in OSEAX of 0.092 percent, where the associated lagged variable is given a coefficient of 0.02. Comparing subpart one and two, the coefficient moves from 0.059 to 0.122, respectively. However, together with the significant lagged variable, the total influence by the oil price in subpart one constitutes 0.093 percent. For the period from 2009 until 2019, the lagged oil price is no longer significant.

In subpart one, the coefficient for the lagged crude is almost half of the coetaneous. This is a strong indication that it takes longer for the oil prices changed to be captured and absorbed in the Norwegian stock market in this period.

During most of subperiod one, the OSEAX was trailing below 200, and the oil price was trading below 25 USD per barrel. Moreover, considering the residuals from regression model two and three, the former model exhibit lower fluctuations, indicating less volatility compared to the subsequent period. This could potentially be one of the reasons why the oil price receives a lower coefficient for this period.

In 2001, Equinor (then Statoil) was listed on OSEAX. During the same time, oil experienced increased attraction and correspondingly higher prices. A growing energy sector at OSEAX, together with high momentum in the oil price was prevalent at the end of subpart one. However, on average, the energy sector comprised 28.2 percent in the first period, in opposition to 39 percent in the second period. Therefore, the lower energy sector in subpart one, combined with more stable and relatively low levels of both oil price and OSEAX, serves as appropriate explanations as to why the coefficient is lower in this period.

For subperiod one, there was also found asymmetry in the oil price effect on OSEAX, with a higher coefficient associated with positive oil price changes, in relation to a smaller coefficient with oil price reductions. This complies with a lower correlation, which might cause the oil price to obtain a lower coefficient in the model. Results from a dummy variable inclusion to indicate positive/negative change can be found in Appendix I.

To test the asymmetric effect suggested by the difference in the oil coefficient being higher for positive changes than negative, we conduct a Z-test between the two variables. The null hypothesis and alternative hypothesis can be stated as:

$$H_0: \hat{\beta}_{up} = \hat{\beta}_{down}$$

$$H_A: \hat{\beta}_{up} \neq \hat{\beta}_{down}$$

By conducting the test using the formula for Z-test, we get a value of 1.42, with a critical value on the 95 percent confidence interval of 1.96. Meaning that we can, in fact, not reject the null of them being different. It still an indication that they differ just not at the required significance level. Note that this is for period one only.

In subpart two, the timeline is shortened with fewer observations. Nevertheless, the period is characterized by a higher weight of the energy sector, together with greater fluctuations in both OSEAX and the oil price. For this period, we found no signs of asymmetry between positive and negative changes in oil prices. Indicating that OSEAX will have the same reaction towards positive and negative changes in oil price, making the two more correlated. This, in turn, makes it easier to model the oil price effect on OSEAX and supports the higher coefficient during this period. Coefficients for positive and negative changes in oil prices can be found in appendix I. For period two, no lagged variables for the oil price is included. Indicating that the market absorbs the information faster.

9.1.6. RENIXX

RENIXX in only included in subpart two, due to its shorter existence. It represents a highly more untraditional measure when trying to explain OSEAX, and thus equivalent interesting. The variable cannot be compared to other periods but can be investigated through subperiod two, with and without the inclusion of RENIXX. Firstly, the variable is significant and does slightly increases the adjusted R-squared of the model. It can be derived that a 1 percent increase in RENIXX leads to a 0.114 percent rise in OSEAX. Compared to the oil price, this variable is approximately equal in terms of the beta coefficient. The combined effect of RENIXX and Brent crude is thus exceeding 0.2. The lagged RENIXX is also significant, but with a negative sign. Indicating that there might be a slight overreaction towards the renewables market.

It appears evident that the increasing interest with regards to renewable energy is applicable at OSE. A trend observed both globally and among companies listed at OSE, is that petroleum companies have started to rebrand themselves. At the Norwegian stock exchange, Equinor represents a prime example. Changing their name from Statoil, they have been reclassified from being a state-owned oil company into an energy company. This is reflected in their enhanced investments in renewable energy projects in recent years (Equinor, 2020). Several other OSE listed companies have followed this approach, including Aker, Kværner, and Bonheur. The combination of a small quantity of renewable energy companies listed at OSE, together with an increased amount of green investing petroleum companies, therefore, appears to be a possible explanation of the significant positive influence on OSEAX.

9.2. Discussion of findings

Prevailing for the analysis and results is that there is no constant relationship between any of the independent variables and OSEAX. Substantial variations in the beta coefficients across periods are discovered, as well as alterations of signs for some coefficients. Moreover, lagged variables found significant in one regression model, have turned insignificant in another.

In light of the significant lagged variables for the oil price, this is consistent with the finding by Driesprong et al. (2008), as for especially subperiod one, there seems to prevail a delayed reaction response to changes in the oil price. Our findings are moreover coherent with the majority of previous research on the field that similarly finds a positive correlation between OSE and the oil price. However, compared to Bjørnland (2009), our beta coefficients appeared moderately lower. Concerning the asymmetric effect investigated, we found similar patterns as Sørensen (2009) but could neither show that positive and negative fluctuations in the oil price are significantly different.

To further investigate the relationship between OSEAX and the examined macro variables, it seems favorable to look deeper into the interplay between OSEAX, USD/NOK, NIBOR, and RENIXX, with emphasis on the period surrounding the oil price shock in 2014. Since oil is traded in USD, the USD/NOK cross-currency plays a vital role, as it is likely to suppress the effect of a weakened oil price. Moreover, in turn, the interest rate will be able to affect the currency and vice versa. From mid-2014, the oil price experienced a substantial decrease, while the OSEAX remained rather stable and later increased. During the same period, the USD/NOK also increased, implying a lowered NOK towards the dollar. From figure x, it can be implied that the Norwegian stock market was partly preserved by a depreciated NOK, in the time of the descending oil price. Initially, one would expect the OSEAX to follow the downfall of the oil price.

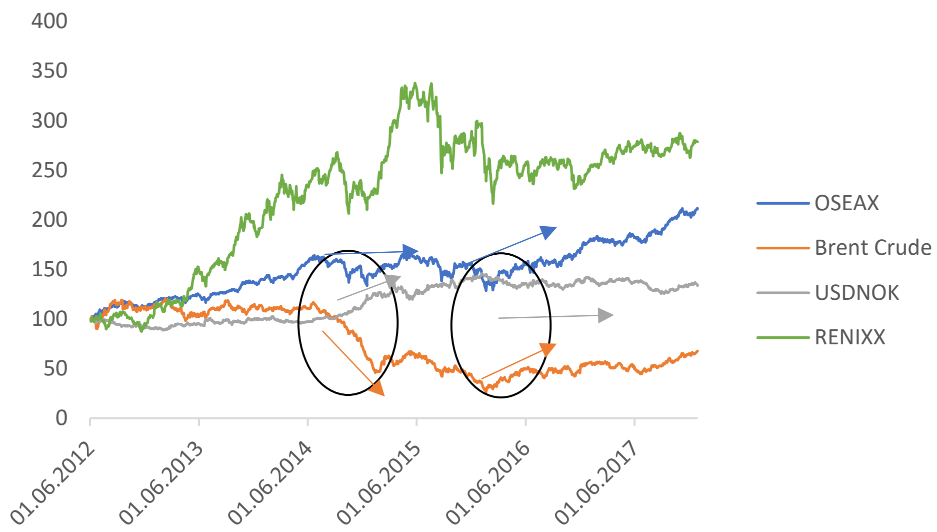


Figure 9.2: Development of selected variables after financial crisis recovery.

In the first circle outlined, it suggests that there is a relation between the currency and the oil price, allowing the stock market to be unaffected. In the second circle in, however, it is indicated that the currency is unaffected, and the stock market captures the effect of increased oil prices. There seems to be no general rule by looking at each variable independently, but one must look at and understand the dynamics between them.

9.3. Weaknesses in analysis

A general weakness of this type of analysis is that the data material will be susceptible to small changes and data manipulation. Thus, it has been vital for us to be transparent in the changes we make and why. Likewise, it has been important to argue well for the changes we make and why it is a necessary change. Many choices and specifications towards what is the best solution, is highly subjective and may lead to a different outcome in our models.

The analysis will also be highly sensitive to the period selected and what is chosen not to be included. It is likely that if another time horizon, or subparts where chosen, the results would be entirely different. We have illustrated periods that go against the model we get from our analysis, both to give a better view of the data we use and explain the dynamics between variables selected.

Due to a large number of observations over such a long horizon, it could also be argued that more subparts should be examined. It is evident that the variables have changed a lot over the period, and no coefficients can be regarded as constant over the period. By splitting into several subparts, one could learn more about how certain events can affect the relationship, how fast it can change, and what drives the change.

A final possible weakness is not to include more variables as a benchmark for industries. RENIXX serves as a benchmark for the renewable energy industry. When included, this variable was able to explain as much of the movements in OSE as the oil price. This is an indication that the oil sector and the renewable energy sector affect OSE similarly. Questioning if all industries would have a similar effect as oil and renewables, or if we, by coincidence, included two industries with the same effect on OSE.

10. Conclusion

Since Norway first discovered oil, the commodity has played an essential part in the country's economy. The petroleum industry constitutes the country's largest industry and has served as the dominant driving force for today's welfare society in Norway. However, during recent decades, both environmental and economic concerns related to the petroleum sector have gained increased traction.

Based on these underlying conditions, this thesis examines the historical influence of the commodity on the country's stock exchange, OSE. Moreover, the conceivable change over time has been investigated. In order to discover the true relationship between the two factors, several other factors have been embedded in an OLS regression. This is done to avoid problems with omitted variable bias. Further, the employed variables have been thoroughly certified by various statistical tests to prevent problems with spurious regression.

For the whole period, 1988-2019, we find that OSEAX is most positively affected by international financial markets, and primarily the London Stock Exchange, represented by the FTSEAX index. Similarly, the Norwegian stock market evidently responds to fluctuations in the S&P 500 stock index, both to the coetaneous and the lagged version. In comparison to the latter, the effect of the oil price is found relatively high. Where a one percent increase in the commodity price leads to a 0.092 percent increase in OSEAX. Similar to the S&P 500, the lagged variable of the oil price is also found significant, with an obtained coefficient of 0.02. With respect to the other included variables, none of them appear to be significant, except for the lagged version of the cross-currency USD/NOK, which comprise a lower influence on OSEAX than the oil price.

Investigating the dynamics in two different subperiods, several changes materialize. For the international financial markets, the reaction by OSEAX is lower in subperiod one (1988-2008), while they both increases in subperiod two (2009-2019), compared to the entire period. Similarly, OSEAX becomes more affected by the oil price during the second subperiod compared to both the first and the entire period. Compared to the remaining variables, the oil price has a greater influence on the Norwegian stock market in the final period, even though the exchange rate have gained a significantly higher impact .

When including a variable for the renewable energy market in subperiod two, RENIXX, the oil price remains rather unaffected. Interestingly, the additional variable obtain a coefficient corresponding to the oil price's. Implying that the two captures different information.

With the inclusion of a variable representing the renewable energy index, RENIXX, the oil coefficient remains unaffected. This implies that the oil price and renewable index capture different information, both individually relevant when explaining OSE. Furthermore, RENIXX takes on a coefficient close to the oil price, indicating that the renewable index has approximately the same effect on OSEAX as the oil price. Indicating that the renewable industry is as influential as the oil industry for the Norwegian stock market for the second subperiod.

Further research

To assess the findings in this thesis, further research could examine the dynamics at a stock exchange in an oil importing country. As stated in the literature review, there has previously been found that oil has a significant negative effect on equity markets in these countries. With the vigorously growing interest by both politicians and investors for renewable energy in recent years, it would have been interesting to examine whether there has been a modification in the influence of the oil price. Moreover, the incentive for these countries to transition would be important to discover. In periods with falling oil prices, one would presume that the impetus for alternative energy sources will diminish, where a low commodity price would benefit the domestic stock market. However, during times with high oil prices or high volatility in the oil prices, the incentive to intensify investments into renewable energy would likely enhance.

Norway represents a unique country within these circumstances. On the one hand, domestic green hydropower is covering all electricity consumption in the country. Norway also have the largest proportion of electric cars per capita in the world. Simultaneously, the country represents the Western Europe's largest producer of oil. The outlook for Norway with this special combination, and hereby the developments in the equity market therefore contributes an appealing proposition for a further research at a later stage. To what degree the companies that today comprise the energy sector at OSE manages to build on existing offshore and oil technologies and move into areas of renewable energy will likely make a huge impact on the forthcoming reliance on the oil price. In the same way will the global oil demand in the coming decades determine the continuation of the oil companies at OSE. A rapid disruption in oil supply today would represent a harmful development for the world economy. Therefore, the oil price and the revenues related will in all probability serve as important for the foreseeable future for Norway.

Bibliography

Amadeo, K. (2020, February 15). *The OECD and Member Countries*. Retrieved from The Balance: https://www.thebalance.com/organization-economic-cooperation-development-3305871

Antwi, S. K., & Hamza, K. (2015). Qualitative and Quantitative Research Paradigms in Business Research: A Philosophical Reflection. *European Journal of Business and Management Vol.7, No.3*, 217-224.

Baltagi, B. H. (2011). *Econometrics.* New York: Springer.

Barstad, S. (2015, July 15). *De største oljerelaterte selskapene i Norge har fått verdiene redusert med en tredel*. Retrieved from Aftenposten: 15https://www.aftenposten.no/okonomi/i/PXOJ/de-stoerste-oljerelaterte-selskapene-i-norge-har-faatt-verdiene-redusert-med-en-tredel?

Berk, J. B., & DeMarzo, P. M. (2017). *Corporate Finance.* Harlow: Pearson Education.

Bernhardsen, T., & Røisland, Ø. (2000). *Hvilke faktorer påvirker kronekursen?* Retrieved from Norges Bank: https://www.norges-bank.no/globalassets/upload/publikasjoner/penger_og_kreditt/2000-03/bernh.pdf

Bernhardsen, T., Kloster, A., & Syrstad, O. (2012). *Risikopåslagene i Nibor og andre lands interbankrenter*. Retrieved from Norges Bank: https://www.norges-bank.no/aktuelt/nyheter-og-hendelser/Signerte-publikasjoner/Staff-Memo/2012/Staff-Memo-202012/

Bjørnland, H. (2009). Oil Price Shocks and Stock Market Booms in an Oil Exporting Country. *Scottish Journal of Political Economy, Vol. 56*.

BloombergNEF. (2019). *Gloabal Trends In Renewable Energy Investment 2019.* Frankfurt: UNEP.

Bodie, Z., Kane, A., & Marcus, A. (2014). *Investments.* Berkshire: McGraw-Hill Education.

BP. (2019, January 23). *BP Statistical Review o f World Energy .* BP. Retrieved from BP.com: https://www.bp.com/en/global/corporate/energy-economics/statistical-review-of-world-energy.html

Brander, A. S., Brekke, H., & Naug, B. E. (2013). *Economic commentaries.* Norges Bank.

Brooks, C. (2017). *Introductory Econometrics for Finance.* Cambridge: Cambridge University Press.

Bryman, A., & Bell, E. (2015). *Business Research Methods.* Oxford: Oxford University Press.

Depersio, G. (2020, March 5). *Why Did Oil and Natural Gas Prices Plummet in 2014?* Retrieved from Investorpedia: https://www.investopedia.com/ask/answers/030315/why-did-oil-prices-drop-so-much-2014.asp

Dickey, D., & Fuller, W. (1981). The Likelihood Ratio Statistics For Autoregressive Time Series With a Unit Root. *Econometrica*, 1058-1072.

Driesprong, G., Jacobsen, B., & Maat, B. (2008). Striking oil: Another puzzle? *Journal of Financial Economics, vol. 89, issue 2*, 307-327.

EIA. (2020, Januaray 21). *What drives Crude oil prices*. Retrieved from EIA.org: https://www.eia.gov/finance/markets/crudeoil/

EIA. (2020, January 21). *What drives crude oil prices: Balance*. Retrieved from EIA.gov: https://www.eia.gov/finance/markets/crudeoil/balance.php

EIA. (2020, January 21). *What drives crude oil prices: Demand Non-OECD*. Retrieved from EIA.org: https://www.eia.gov/finance/markets/crudeoil/demand-nonoecd.php

EIA. (2020, January 21). *What drives crude oil prices: Demand OECD*. Retrieved from EIA.gov: https://www.eia.gov/finance/markets/crudeoil/demand-oecd.php

EIA. (2020, January 21). *What drives crude oil prices: Financial Markets*. Retrieved from EIA.gov: https://www.eia.gov/finance/markets/crudeoil/financial_markets.php

EIA. (2020, Januar 21). *What drives crude oil prices: Spot Prices*. Retrieved from EIA.gov: https://www.eia.gov/finance/markets/crudeoil/spot_prices.php

EIA. (2020, January 21). *What drives crude oil prices: Supply Non-OPEC*. Retrieved from EIA.gov: https://www.eia.gov/finance/markets/crudeoil/supply-nonopec.php

EIA. (2020, January 21). *What drives crude oil prices: Supply OPEC*. Retrieved from EIA.gov: https://www.eia.gov/finance/markets/crudeoil/supply-opec.php

Engle, R. (2011). GARCH 101: The Use of ARCH/GARCH Models in Applied Econometrics. *Journal of Economic Perspectives—Volume 15, Number 4*, 157-168.

F.Fama, E., & French, K. R. (1996). Multifactor Explanations of Asset Pricing Anomalies. *The Journal of Finance, Volume 51*, 55-84.

Fama, E. f. (1970). Efficient Capital Markets: A review of theory and empirical work. *The yournal of Finance*, 383-417.

Fama, E. F. (1995). Random Walks in Stock Market Prices. *Financial Analysts Journal, Vol. 51*, 75-80.

Finans Norge. (2017). *Nibor – the Norwegian Interbank Offered Rate*. Retrieved from Finanse Norge: https://www.finansnorge.no/en/interest-rates/nibor---the-norwegian-interbank-offered-rate/

Fransvold, T. (2016, January 7). *Ingen av de norske feltutbyggingene lønnsomme*. Retrieved from Dagens Næringsliv: https://www.dn.no/uro-i-markedene/uro-i-oljemarkedet/oljepris/ingen-av-de-norske-feltutbyggingene-lonnsomme/1-1-5548349

Grossman, S. J., & Stiglitz, J. E. (1980). On the Impossibility of Informationally Efficient Markets. *The American Economic Review, Vol. 70*, 393-408.

Hammoudeh, S., & Li, H. (2005). Oil sensitivity and systematic risk in oil-sensitive stock indices. *Journal of Economics and Business, vol. 57, issue 1*, 1-21.

Hecht, A. (2020, May 1). *Understanding the Crude Oil Market*. Retrieved from The Balance: https://www.thebalance.com/crude-oil-brent-versus-wti-808872

Hovland, K. M., & Lorentzen, M. (2019, December 3). *Equinor er sin egen sjette største eier: – Aksjen ville nok vært enda mer ned uten tilbakekjøpene*. Retrieved from E24: https://e24.no/boers-og-finans/i/8mLWV1/equinor-er-sin-egen-sjette-stoerste-eier-aksjen-ville-nok-vaert-enda-mer-ned-uten-tilbakekjoepene

IWR. (2020). *RENIXX World - global stock index*. Retrieved from IWR Faszination Energie: https://www.iwr-institut.de/en/services/renewable-economic-research/renixx-world-en

Jegadeesh, N., & Titman, S. (1993). Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency. *The Journam of Finance, Vol. 48*, 65-91.

Jung, H., & Park, C. (2011). Stock market reaction to oil price shocks: A comparison between an oil-exporting economy and an oil-importing economy. *Journal of Economic Theory and Econometrics, Vol. 23*, 1-29.

Kendall, M. G., & Hill, A. B. (1953). The Analysis of Economic Time-Series-Part I: Prices. *Journal of the Royal Statistical Society. Series A (General), Vol. 116*, 11-34.

Khashman, A. (2019). *A Transition in Energy Investment*. Retrieved from Pexapark: https://pexapark.com/blog/green-energy-investment/

Lintner, J. (1965). The Valuation of Risky Assets and the Selection of Risky Investment. *The Review of Economics and Statistics*, 13-37.

Morningstar. (2019). *norwegian-interbank-offer-rate-(nibor)*. Retrieved from Morningstar: https://www.morningstar.no/no/glossary/102696/norwegian-interbank-offer-rate-(nibor).aspx

Mossin, J. (1966). Equilibrium in a Capital Asset Market. *Econometrica*, 768-783.

Munk, C. (2018). *Financial Markets and Investments.* Copenhagen.

Myrseth, S. (2020, February 3). *Det Store Hamskiftet*. Retrieved from Dovre Forvaltning: https://www.dovreforvaltning.com/no/content/det-store-hamskiftet

Neuman, W. (2014). *Social Research Methods: Qualitative and Quantitative Approaches.* Essex: Pearson.

Nilsen, A. A. (2019, June 17). *Forvaltere ser grønn bølge i aksjemarkedet: – Nesten litt frenetisk*. Retrieved from E24: https://e24.no/boers-og-finans/i/0ngEnB/forvaltere-ser-groenn-boelge-i-aksjemarkedet-nesten-litt-frenetisk

Norges Bank. (2020, February 13). *Økonomiske perspektiver*. Retrieved from Norges Bank: https://www.norges-bank.no/aktuelt/nyheter-og-hendelser/Foredrag-og-taler/2020/2020-02-13-arstalen/

Norwegian Petroleum. (2020, April 27). *PRODUCTION FORECASTS*. Retrieved from Norwegian Petroleum: https://www.norskpetroleum.no/en/production-and-exports/production-forecasts/

Norwegian Petroleum. (2020, April 27). *THE GOVERNMENT'S REVENUES*. Retrieved from Norsk Petroleum: https://www.norskpetroleum.no/en/economy/governments-revenues/

Næs, R., Skjeltorp, J., & Ødegaard, B. (2009). What factors affect the Oslo Stock Exchange? *Norges Bank*, 1-61.

Opland, H. (2017, December 22). *Derfor er rentenivået viktig for prisingen av aksjer*. Retrieved from Odin Bloggen: https://blogg.odinfond.no/renteniva-prisingen-av-aksjer

Oslo Børs. (2018, July 11). *Energi på nye høyder*. Retrieved from Oslo Børs: https://www.oslobors.no/Oslo-Boers/Om-Oslo-Boers/Nyheter-fra-Oslo-Boers/Energi-paa-nye-hoeyder

Oslo Børs. (2019, December 17). *Utenlandsk eierskap på høyeste nivå siden finanskrisen*. Retrieved from Oslo Børs: https://www.oslobors.no/Oslo-Boers/Om-Oslo-Boers/Nyheter-fra-Oslo-Boers/Utenlandsk-eierskap-paa-hoeyeste-nivaa-siden-finanskrisen

Oslo Børs. (2020, February 7). *Oslo Børs All-share Index*. Retrieved from Oslo Børs: https://www.oslobors.no/markedsaktivitet/#/details/OSEAX.OSE/overview

Park, J., & Ratti, R. (2008). Oil price shocks and stock markets in the U.S. and 13 European countries. *Energy Economics, vol. 30, issue 5*, 2587-2608.

Perersen, C. V., & Plenborg, T. (2012). *Financial Statement Abalysis.* Harlow, England: Pearson Education.

Pickl, M. (2019). The renewable energy strategies of oil majors – From oil to energy? *Energy Strategy Reviews, Vol. 26*.

Saunders, M. N., Lewis, P., & Thornhill, A. (2019). *Research Methods for Business Students, 8th Edition.* New York: Pearson.

Sharpe, W. F. (1964). Capital Asset Prices: A Theory of Market Equilibrium under Conditions. *The Journal of Finance Vol. 19, No. 3*, 425-442.

Singh, N., Nyuur, R. B.-b.-i., & Richmond, B. (2019). Renewable Energy Development as a Driver of Economic Growth: Evidence from Multivariate Panel Data Analysis. *Sustainability, Vol 18*.

Siripurapu, A., & Chatzky, A. (2020, January 21). *OPEC in a Changing World*. Retrieved from Council on Foreign Relations: https://www.cfr.org/backgrounder/opec-changing-world

SSB. (2020). *Utenrikshandel med varer*. Retrieved from SSB: https://www.ssb.no/statbank/table/08804/tableViewLayout1/

Statista. (2019). *Largest stock exchanges in Europe as of December 2019, by domestic market capitalization*. Retrieved from Statista: https://www.statista.com/statistics/693587/stock-exchanges-market-capitalization-europe/

Sørensen, L. Q. (2009). *Oil Price Shocks and Stock Return Predictability.* Bergen: Norwegian School of Economics.

Treynor, J. L. (1961). Market Value, Time, and Risk. *Originally unpublished*.

Trocmé, J., & Kruse, K. (2016). *Oil in the Nordics.* Nordea Markets.

Tuli, F. (2010). The Basis of Distinction Between Qualitative and Quantitative Research in Social Science: Reflection on Ontological, Epistemological and Methodological Perspectives. *Ethiopian Journal of Education and Sciences*, 97-108.

UNEP. (2019). *A decade of renewable energy investment, led by solar, tops USD 2.5 trillion*. Retrieved from UN Environment Programe: https://www.unenvironment.org/news-and-stories/press-release/decade-renewable-energy-investment-led-solar-tops-usd-25-trillion

Wooldridge, J. M. (2012). *Introductory Econometrics A Modern Approach, 5. edition.* Michigan: South-Western Cengage Learning.