Economics and Business Administration

MSc in Applied Economics and Finance

Master's Thesis

Managing Electricity Price Risk with Futures

An Empirical Analysis of Hedging Strategies in the Nordic Power Market

---

Authors:

Per Christian Enge (115475) and Benjamin Kjevik (115615)

Supervisor:

Peter Belling

Hand-in date: 05.15.2019

Characters (with spacing): 238.580

Pages: 116

Abstract

Electricity markets around the world have been subject to deregulation for the last few decades, and the highly volatile characteristics of the spot prices have increased the need for practical risk management tools for actors in the market. The objective of this thesis is to examine the performance of static and dynamic hedging models with monthly and quarterly futures in the Nordic power market from 2005 to 2018. This is done by constructing hedged portfolios with a time-invariant hedge ratio from the naïve and ordinary least squares model, and time-varying hedge ratios from the constant conditional correlation GARCH model and the dynamic conditional correlation GARCH model.

It is found that the best-performing GARCH model significantly outperforms the best-performing static model, both in-sample and out-of-sample. Furthermore, the results show that hedging performance varies considerably across periods. Specifically, an indication is found that the relative advantage of the GARCH models compared to the static models is greater when the volatility in the spot market is high. Measuring performance by mean-variance utility shows smaller differences in performance between the static and dynamic hedging models, but the GARCH models still rank highest overall. This suggests that dynamic hedging is beneficial also for hedgers that are utility-maximizing. Last, it is found that the hedging models obtain a significantly lower variance and value at risk when hedging with monthly contracts compared to quarterly contracts. In conclusion, the results show that futures hedging reduces the portfolio risk significantly compared to a no-hedge strategy, and this result is especially evident for the dynamic hedging models.

From a hedger's perspective, this thesis emphasizes the benefits of dynamic hedging with futures in the Nordic power market. The recommended strategy for actors hedging monthly and quarterly deliveries is to dynamically adjust the hedge ratio in their portfolios according to a constant conditional correlation GARCH model.

Acknowledgements

This thesis is written as a final part of our Master of Science in Applied Economics and Finance at Copenhagen Business School.

We wish to express our gratitude to our supervisor, Peter Belling, for the constructive comments and helpful advice throughout the process of writing the thesis. Moreover, we would like to thank Glenn Qvam Håkonsen and Sveinung Skjevrak in Konsesjonskraft IKS for giving access to the necessary data and insights. We are also thankful to Nord Pool for providing access to their FTP server and Professor Jim Hanly for advices on significance testing. Finally, we would like to acknowledge all the people who have been supportive along the way.

Table of Contents

List of Figures

List of Tables

1. Introduction

The first section of the thesis will cover the background and context for the topic along with the motivation behind the choice. The research question with corresponding sub-questions and the objectives will also be presented. The methodology used, including the delimitations of the thesis and a discussion of the data material used, will be discussed. The outline of the thesis will be presented at the end of the section.

1.1. Background and Motivation

In recent years, power markets around the world have been deregulated, and the structure has changed from regulated monopolies to become competitive open markets (Fiorenzani, 2006, p. 6). The scope of open electricity markets is to obtain economic efficiency, physical sustainability and grid balancing (Fiorenzani, 2006, p. 6). The deregulation has brought increased transparency into the market for power commodities and has led producers, wholesalers, and traders to be more active on the financial side of electricity markets (Nord Pool, n.d.-a).

Electricity markets differ from most other commodity markets due to several characteristics. The primary difference between electricity and other commodities such as gold, crude oil, and grain is that electricity is a non-fungible commodity with no possibility of being stored or transported in significant quantities (Zanotti, Gabbi, & Geranio, 2009), and the usage typically varies every minute (Bessembinder & Lemmon, 2002). Because demand and power generation need to match instantaneously, and since supply and demand shocks are impossible to control using inventories (Zanotti et al., 2009; Bessembinder & Lemmon, 2002), electricity markets become highly volatile with strong seasonal patterns, unlike most other commodity markets (Vehviläinen & Keppo, 2003). These characteristics lead electricity to be the most volatile commodity traded (EIA, 2002). Because of the high volatility in the market and the inelastic features of the demand and supply curves, hedging adverse spot price movements becomes essential for the participants operating in the market (Hanly, Morales, & Cassells, 2017).

The most common way of managing electricity price risk is to trade derivatives such as futures and forward contracts, thereby making future income and costs more predictable for sellers and buyers, respectively (Hanly et al., 2017). For instance, a producer of electricity worried about a fall in the electricity price could sell futures contracts and partially offset this risk. Fiorenzani (2006, p. 41) finds that traditional hedging instruments often lack attention to the potential dangers arising from the relatively unique market characteristics. Availability of practical risk management tools in power markets is therefore of high importance and can benefit actors operating in the industry.

This thesis focuses on hedging with the use of futures contracts in the Nordic power market, and there are several reasons why this topic has been chosen. Several exciting and challenging courses in the Applied Economics and Finance program partially motivate the choice of hedging and risk management. Furthermore, as the volatility of electricity prices is very high compared to most other assets and commodities, hedging becomes a key area of focus for actors in this industry. Sanda, Olsen, & Fleten (2013) found that more than 90% of the total electricity produced in Norway is hedged, which substantiates the argument. The choice of examining the electricity market is also motivated by its unique and complex market characteristics. Besides, as the liberalization of the electricity markets is still relatively new compared to other commodity markets, existing research on hedging in the Nordic power market is still somewhat limited (Hanly et al., 2017). The limited research on the field gives further motivation to investigate risk management in the Nordic power market as it provides an opportunity to extend on research in a market that is currently relatively narrow compared to other markets.

Although limited, there exists some prior research on hedging in the Nordic power market, with the most prominent studies being those of Byström (2003), Zanotti et al. (2009) and Hanly et al. (2017). These papers examine whether hedging strategies based on time-varying hedge ratios can outperform strategies with static hedge ratios and how effective the futures contracts are in reducing the risk for market participants. However, the papers differ in their setup and present somewhat contradictory results[1], increasing the motivation to further extend the research regarding hedging in the Nordic power market.

[1] The papers' findings will be more thoroughly discussed in the literature review in section 4.

1.2. Involved Parties

This sub-section will provide a brief presentation of parties of relevance for the thesis.

KIKS

The Norwegian company Konsesjonskraft IKS (KIKS) is a municipal collaboration established in 1984 (Konsesjonskraft IKS, n.d.). KIKS manages 871 GWh of concessionary[2] electricity on behalf of 20 Norwegian municipalities and the two counties of Agder (Konsesjonskraft IKS, 2017). KIKS' objective is to manage the concessional power according to the strategies developed in collaboration with the municipalities to ensure stable and predictable income for the municipalities, which is beneficial for the municipalities when developing their budgeting plans. As such, KIKS is a market participant with a strong need for adequate risk management policies and is thus representing a typical actor in the Nordic power market.

KIKS has contributed to the thesis by providing important insights on the Nordic power market and by providing historical futures data for the analysis. The thesis therefore aims to analyze hedging strategies that KIKS can benefit from in the future.

Nord Pool

Nord Pool is the physical exchange for power in the Nordic market. They have contributed to the thesis by sharing access to their FTP server for data collection. A more detailed description of Nord Pool and the functioning of the physical market for electricity will be given in sub-section 2.1.

Nasdaq OMX Commodities

Nasdaq OMX Commodities is the exchange for financial derivatives in the Nordic power market. Nasdaq has been involved in the thesis by providing fundamental insights into the financial market for Nordic power, as well as providing a thorough description of the settlement price and other contract specifications for the futures. Nasdaq OMX Commodities will be further introduced in sub-section 2.2.

[2] The Norwegian law of concessional energy states that municipalities that are affected by hydropower generation have a claim to receive up to 10% of the yearly production of the power plant (Lovdata, 1917).

1.3. Research Question and Objectives

The thesis takes the viewpoint of actors operating in the Nordic power market, such as producers or distributors of electricity, that have committed to selling electricity on the spot market in the future. An example of such an actor is the Norwegian company KIKS, as briefly introduced in the previous sub-section. Due to the highly volatile spot prices of electricity, the actors engage in hedging activities to mitigate the risk of adverse price movements in the spot price. As such, the actors take a short position in futures contracts to hedge the long position in the spot market. The size of the short position taken in futures contracts relative to the spot position is defined as the *hedge ratio*, which could either take a time-varying or time-invariant structure (Hull, 2012, p. 823). The thesis makes use of data on historical spot and futures prices to build econometric hedging models that are frequently found in the literature. These are in turn tested over a sample period that has, as far as knowledge goes, not been previously researched. The main research question of the thesis is therefore formulated as:

How can actors in the Nordic power market most effectively hedge electricity price risk with futures?

The literature uses different ways of measuring hedging effectiveness, and several performance measures will, therefore, be applied when evaluating the effectiveness of the models. Furthermore, to adequately address the main research question, the research question is divided into four sub-questions. The sub-questions, along with the rationale behind each sub-question, will be presented in the following.

The mentioned studies, covering hedging in the Nordic power market, report inconsistent findings when examining if hedging models with time-varying hedge ratios can outperform models with a static hedge ratio. Byström (2003) finds no benefits of hedging with dynamic hedge ratios compared to a static hedge ratio. Zanotti et al. (2009) and Hanly et al. (2017), however, report relatively higher hedging effectiveness of dynamic hedging models as compared to the static ones. Consequently, the first sub-question aims to address this.

1. *Are hedging models with dynamic hedge ratios more effective in reducing risk compared to models with static hedge ratios?*

The Nordic power market has a complex nature and is affected by various outside factors. Investigating if the hedging models perform differently depending on how the market behaves can help market participants adjust their hedging strategy accordingly. As an example, Zanotti et al. (2009) find that dynamic hedging models perform relatively better than static hedging models in the European Power Exchange (EEX) market when the volatility in the spot market is high, that is, in times when hedging is most important. The second sub-question aims to answer whether this is the case for the Nordic power market.

2. *Does the performance of the hedging models depend on the volatility in the spot market?*

The implementation of dynamic hedging models in an applied setting requires the participants to account for transaction costs associated with a frequent rebalancing of the hedged portfolio. The next sub-question will address this issue and aims to provide insights on the performance of the dynamic hedging models when transaction costs are considered.

3. *How do dynamic and static hedging models compare when accounting for transaction costs in a mean-variance utility framework?*

The electricity futures traded in the market can differ both in maturity and contract lengths. Consequently, analyzing different contracts can reach a broader audience and relate to different types of market participants. Monthly futures contracts are the most pronounced contracts in existing research, while the analysis of quarterly contracts adds to the existing research. Comparing how the different contracts perform relative to each other can lead to valuable insights for participants operating with different contract lengths. The last sub-question aims to compare the hedging performance of these two contract types.

4. *How does the hedging performance between the use of monthly and quarterly futures compare?*

In order to answer the main research question and the mentioned sub-questions, the thesis takes on the following research objectives:

1) Provide the reader with a detailed presentation of the Nordic power market and the available risk management tools in the market.
2) Review existing literature on both hedge ratio modeling in general and hedging in electricity markets.
3) Build econometric hedging models according to the Box-Jenkins methodology and carry out necessary diagnostic tests.

4) For each hedging model, estimate optimal hedge ratios through an in-sample analysis and perform forecasts through an out-of-sample analysis.
5) Assess the performance of each hedging model according to performance measures commonly used in the hedging literature.
6) Provide recommendations on hedging strategies for actors operating in the Nordic power market.

1.4. Methodology

The quantitative method is applied in this thesis, as all main conclusions are backed up by financial and statistical analyses. This method is appropriate when analyzing hedging strategies because of high data availability of historical spot and futures prices in the Nordic power market. Some of the analyses are descriptive in the way that they describe how the market has evolved, but the main objective of the thesis will be to find an optimal hedging strategy for an actor in the Nordic power market, and is therefore normative.

1.4.1. Research Scope and Delimitations

The focus of the thesis is limited to the Nordic power market, and other electricity markets will not be included in the analysis. The Nordic power market includes both the physical market, Nord Pool, and the financial market, Nasdaq OMX Commodities. The main reason for the choice of the Nordic market is that Nord Pool is the largest and most liquid power exchange in Europe (Torró, 2009; Europex, n.d.), but also because this is the market where the Nordic power firms operate.

The analysis will cover monthly and quarterly futures contracts, and it will not examine other contract lengths such as daily, weekly or yearly contracts. These contract types are primarily chosen because they are the most liquid futures contracts traded on the exchange (Botterud, Kristiansen, & Ilic, 2009), signifying their importance as hedging vehicles for actors in the market. Moreover, the monthly contracts are also chosen to enable the thesis to make comparisons of the results with existing literature on the field (Byström, 2003; Hanly et al., 2017). Quarterly contracts are less pronounced in the literature but will provide insights for market participants operating with longer delivery periods. Nasdaq offers other types of derivatives in addition to futures. However, these derivatives will be outside the scope of this thesis and will therefore not be considered.

In reality, the Nordic countries are divided into smaller sections which can have different electricity prices depending on the capacity of the power lines in the region (Statnett, 2018). If all power lines were the same with no congestion restrictions, the prices would be identical (Energifakta Norge, n.d.). However, in order to generalize the results and not complicate the analysis with the different area prices, the system price set by Nord Pool is treated as the spot price throughout the thesis. This is beneficial to do as the system price is the spot reference price for the power futures. The system price will be described in detail in sub-section 2.3.3.

Some transactions are taking place “over the counter” (OTC) in the bilateral market using forward contracts. Types of traders can be producers, distributors or end consumers, and most of the deals in the OTC-market are negotiated and done by a broker. All trades in the OTC market are reported to Nord Pool Clearing, with few exemptions (Olje- og energidepartementet, 2008). It is, however, challenging to obtain data on OTC transactions, and only transactions taking place in regulated exchanges will therefore be considered.

As previously mentioned, the research question of the thesis is related to hedging strategies and their ability to reduce the risk for actors operating in the market. Consequently, other aspects of hedging such as the effect of hedging on firm value or financial performance will not be investigated.

1.4.2. Data

Finding reliable data on historical spot and futures prices in the Nordic power market is unproblematic as the markets are open and transparent. The spot prices were directly obtained from Nord Pool’s FTP server. For the futures prices, extensive research was necessary because of a platform change in 2008. Bloomberg only shows historical futures prices dating back to 2014, but by contacting Nord Pool and KIKS directly, data from the period 2005-2018 was retrieved. Data received from Nord Pool and KIKS were compared with those retrieved from Bloomberg (2014-2018) and found indifferent. These sources are reliable, and the assumption is therefore that the datasets received are the actual prices with no errors. A further description of the data will be presented in section 5.

1.5. Outline

Section one has now given a presentation of the background and motivation for the topic of this thesis. The research question and objectives were also presented, together with the methodology, delimitations and a description of the data material used.

Section two will introduce the reader to the Nordic power market. This will include industry background of both the physical and financial market for electricity in the Nordic countries as well as a description of the electricity spot price formation.

Section three will first provide information about the relevance of hedging. Subsequently, a review of different pricing theories for futures and how they apply to the pricing of electricity futures will be presented. The concept of basis risk will also be discussed.

Section four will provide a literature review of optimal hedge ratio modeling and hedging in power markets, with an emphasis on the Nordic market.

Section five will describe the data used for the analysis in the thesis. Choices made regarding sampling interval, different sample periods and rollover procedure will be discussed. The section will also include descriptive statistics of the data and a test for stationarity in the time series.

Section six will start with an introduction of the different models that are used to estimate the optimal hedge ratios for the analysis. The chosen performance measures will also be presented in addition to the bootstrapping technique used to facilitate t-tests for statistical inference.

Section seven will present the empirical results of the hedging performance of the models, which includes both an in-sample analysis and an out-of-sample analysis.

Section eight will extend the empirical results by giving a further discussion of the results along with practical interpretations and implications of the findings. The results will be discussed in relation to the research question and previous studies on the field.

Section nine will conclude the thesis. The main research question with its corresponding sub-questions will be answered based on the obtained results.

Section ten will view the results from different perspectives and provide suggestions for further research.

2. The Nordic Power Market

This section will describe how the Nordic electricity market work, how the demand and supply curves evolve and how electricity prices are determined. This is to familiarize the reader with the basic characteristics of electricity markets, which is essential before entering the analysis.

2.1. The Physical Market - Nord Pool

The physical market for electricity in the Nordic countries is called Nord Pool and was established in 1993 as the first spot exchange for electricity in the world (Fiorenzani, 2006, p. 5). It is the largest and most liquid of all European electricity markets (Torró, 2009; Europex, n.d.), and in 2018 a total of 524 TWh[3] of power was traded through the exchange (Nord Pool, n.d.-b). Nord Pool offers trading, clearing, settlements, and associated services across nine Nordic countries[4] (Nord Pool, n.d.-b).

In October 2008, Nord Pool was acquired by Nasdaq (Nasdaq, 2008). After the acquisition, the market was divided into a physical market for electricity with the original name Nord Pool, and a financial market, Nasdaq OMX Commodities, where power derivatives are traded. These markets are non-mandatory, meaning that all producers and consumers choose to either interact in the market or to enter bilateral OTC contracts for the delivery of electricity in the short and long term (Zanotti et al., 2009). Most of the trades take place in the wholesale market, where power producers sell electricity to power suppliers, who further resell the electricity to their end consumers (Nord Pool, n.d.-b).

The Nordic electricity market uses a point tariff system (Nord Pool, n.d.-a). The idea of this system is that the producers pay a fee to the owner of the grid for the electricity they pour into the grid, and the end users pay a fee for the electricity they draw from the grid (Nord Pool, n.d.-a). The point is that somewhere on the grid, a producer must pour an amount of electricity that corresponds to the electricity a consumer has tapped from the grid (Energifakta Norge, n.d.).

Nord Pool offers two types of trading: day-ahead (Elspot) where electricity is bought for the next day, and intraday (Elbas) where electricity is bought one or more hours before delivery on the same day (Nord

[3] 1 Terrawatt hour = 1,000,000 MWh
[4] Norway, Denmark, Sweden, Finland, Estonia, Latvia, Lithuania, Germany, and the United Kingdom.

Pool, n.d.-c). The deadline for purchasing day-ahead power is 12 AM, and the deadline for purchasing intraday power is one hour before delivery (Nord Pool, n.d.-c). The system price, which will be further explained in sub-section 2.3.3, is calculated from an auction process in the day-ahead market. The system price is also used as the reference price for the power derivatives traded in the financial market (Nasdaq, n.d.-a). For that reason, the focus of this thesis will be on Elspot, commonly referred to as the day-ahead spot market (Huisman & Kilic, 2012; Botterud et al., 2009). Figure 1 shows an illustration of how the market works and the positioning of different participants in the market.

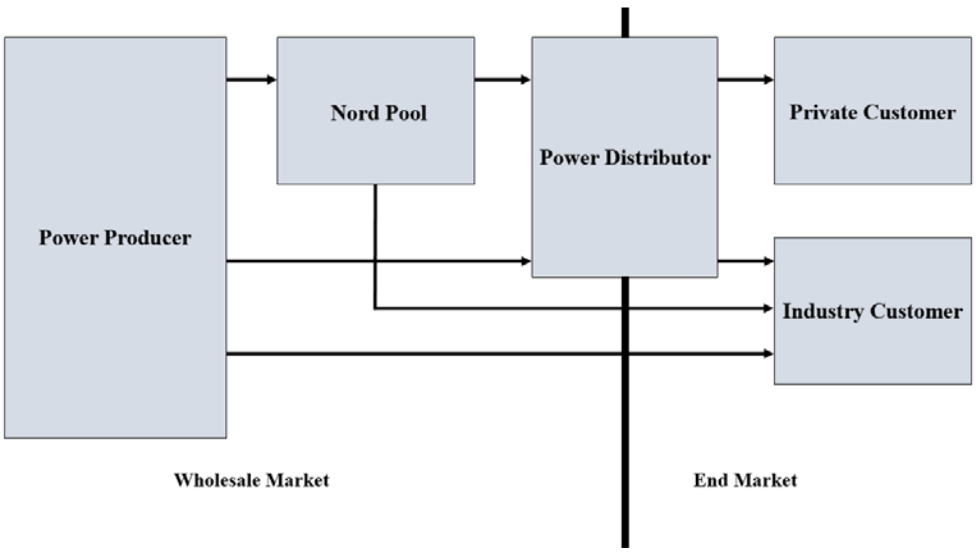


Figure 1 *- The organization of the Nordic power market (Energifakta Norge, n.d.)*

2.2. The Financial Market – Nasdaq OMX Commodities

As previously mentioned, the financial market for Nordic power is operated by Nasdaq OMX Commodities. The financial market is primarily used for price hedging and risk management, but there are also individual traders speculating in the derivatives because of cash settlements (Nasdaq, n.d.-b). Power derivatives traded at Nasdaq represent claims of future delivery of electricity, but the contracts are only financially settled with the system price from Nord Pool as the reference price (Nord Pool, n.d.-d). Nasdaq Clearing is acting as a counterparty for every transaction taking place in the market (Fiorenzani, 2006, p. 6). This eliminates all counterparty risk and is beneficial for the liquidity in the market (Nasdaq, n.d.-b).

High volatility in the market leads the market participants to engage in long-term financial derivatives contracts. By doing so, they can reduce the risk of price variations (Olje- og energidepartementet, 2015). The power derivatives offered on the exchange are futures, DS futures[5], options and EPADs[6] (Nasdaq, 2016). The contract lengths of the derivatives include day, week, month, quarter and year. As mentioned, only the Nordic power futures will be considered for the analysis in this thesis.

2.3. Price Determination

The Nordic electricity market is based on a demand-side bidding system where the electricity price is based on the marginal price rule (EA, n.d.). The following will describe how the electricity demand and supply curves are formed, and how electricity prices are determined.

2.3.1. Demand

The demand side of electricity markets is called the *load*. The load represents the total demand for electricity in the market at a given time.

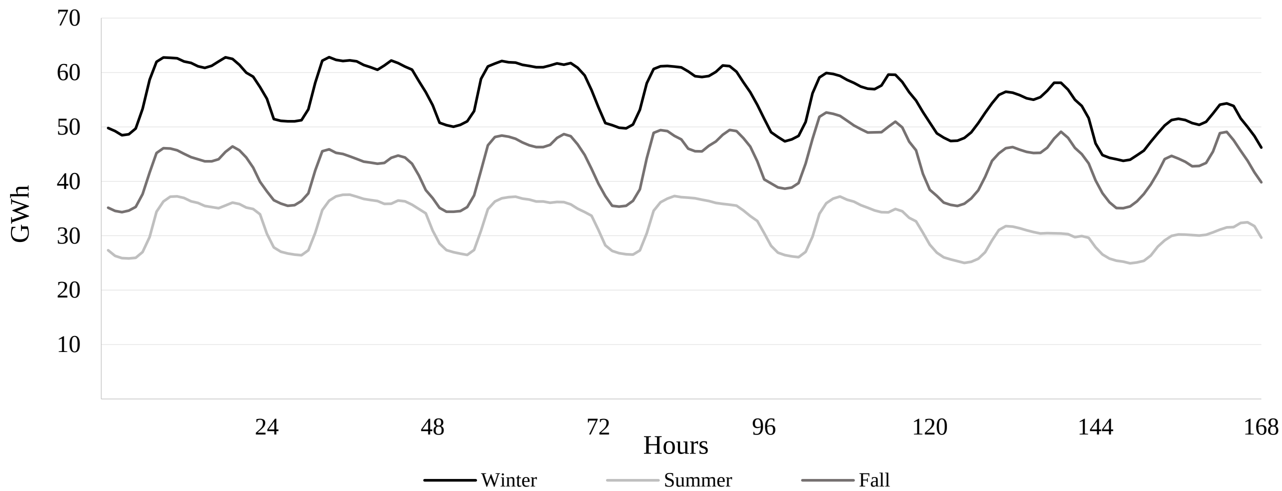


Figure 2 *- Total electricity turnover at Nord Pool. Winter (02/05/2018-02/11/2018), summer (07/23/2018-07/29/2018) and fall (10/22/2018-10/28/2018) (Source: Nord Pool FTP server and own calculations)*

[5] Previously known as forward contracts and was renamed to distinguish them from forward contracts taking place in the OTC market (Mäntysaari, 2015, p. 579). DS stands for 'deferred settlement', meaning that the derivatives are not marked-to-market and only settled at the maturity date (Nasdaq OMX, 2018).

[6] EPAD = Electricity Price Area Differential contract. These reflect the price difference between, e.g. an Area Price and the Nordic System Price (Nasdaq, 2016).

The pattern of the load over time, shown in Figure 2, is called a *load curve* and is affected by many factors (Posner, 2018). Colder and warmer temperatures lead to increased demand for heating and cooling, respectively (Posner, 2018). As can be seen from Figure 2, the demand is also dependent on the time of the day as the load is higher during the day and lower during the night for all seasons. Furthermore, it shows a higher load during breakfast and dinner times, and lower at the weekends when offices and schools are closed.

2.3.2. Supply

In electricity markets, each producer enters bids for how much electricity they want to sell at a given price (Posner, 2018). This means that each producer provides the market with an individual supply curve, which often leads the individual supply curves to be upward sloping. The system price depends on various supply factors in the market. Generation of water power is dependent on the hydrological situation, wind power depends on winds, and solar power depends on sunshine. Coal, gas and nuclear power are not dependent on weather conditions, but other factors such as fuel prices and emission permits, and the production of these energy sources are often used to balance the system price (Olje- og energidepartementet, n.d.). This means that in years with normal production of water, solar and wind power, the system price will highly depend on the cost of coal and gas power production (Olje- og energidepartementet, n.d.).

The *transmission system operator* (TSO) is responsible for regulating the power market to ensure a stable power in the transmission grid (Nord Pool, n.d.-a). There are seven non-profit TSOs[7] present at Nord Pool, and they are responsible for the collection of bids in their respective countries, arranging the bids in ascending price order based on marginal cost (Nord Pool, n.d.-a). This means that in periods when demand is high and the producers are approaching the capacity limit, producers with the highest marginal cost are used to generate the requested electricity. As found by Bessembinder and Lemmon (2002), this makes the supply curve convex as electricity prices increase in demand. The individual supply curves collected are aggregated into a ‘generation stack’ (Posner, 2018). In Figure 3, a typical generation stack

[7] Statnett SF, Svenska kraftnät, Fingrid Oy, Energinet.dk, Elering, Litgrid and Augstsprieguma tikls (AST) (Nord Pool, n.d.-a)

for the Nordic electricity market is presented, showing an approximation of the marginal costs for producers trading at Nord Pool (Huisman, Michels, & Westgaard, 2014).

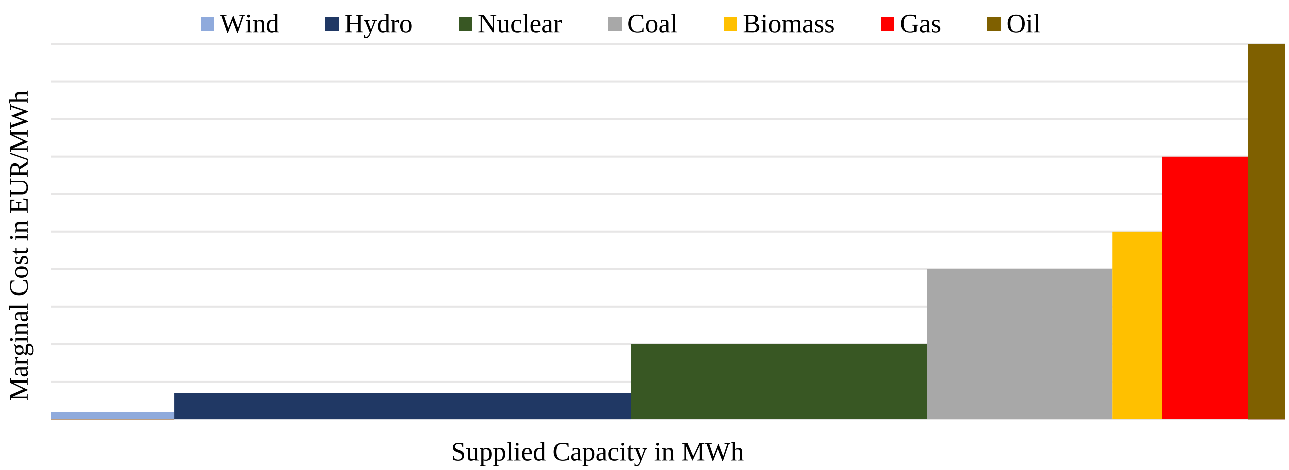


Figure 3 *- Nord Pool generation stack (Huisman et al., 2014)*

Figure 3 shows that hydropower, wind power, and other renewable energy sources have the lowest marginal costs. The marginal cost of production for these energy sources are in some cases close to zero (Olje- og energidepartementet, 2006). Coal, gas, and oil powered plants have the highest marginal costs, which can be explained by high fuel costs and emission permits (Huisman et al., 2014).

2.3.3. System Price

The Nordic system price is an unconstrained market clearing reference price and is calculated based on infinite capacity (Nord Pool, n.d.-e). The TSOs combine supply and demand curves that are specified by the market participants to obtain the production rule that minimizes the costs of aggregated demand (Fiorenzani, 2006, p. 5). This is done through a double auction process where the equilibrium price is found where the aggregated supply and demand curves intersect (Nord Pool, n.d.-f). Specifically, the system prices are calculated as the equally weighted average of the intersection of the aggregated supply and demand curves in each hour for all bidding areas after the day-ahead deadline (Nord Pool, n.d.-a). This eliminates differences in area prices and constitutes one common bidding area with one price, as the capacities are set to infinity (Nord Pool, n.d.-e). Thus, for the remainder of the thesis, ‘spot prices’ refers to the Nordic system prices set by Nord Pool in the day-ahead market.

The large variations in demand and the function of the generation stack lead to highly volatile prices (Huisman et al., 2014). This is clear from Figure 4, which shows the daily system prices in the Nordic power market from 2005 to 2018.

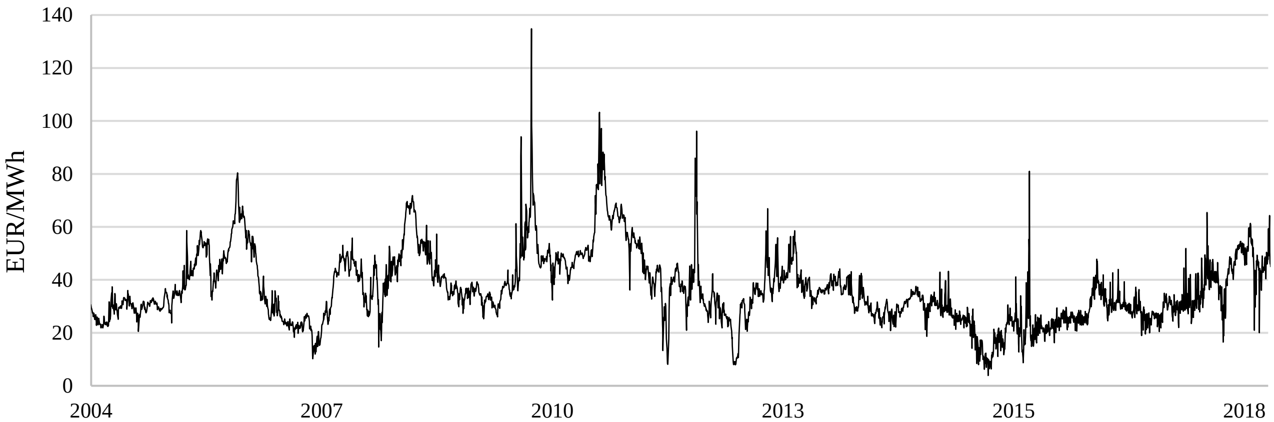


Figure 4 *- Historical system prices (Source: Nord Pool FTP server)*

This section has now presented an overview of the Nordic power market with a description of the physical and financial market. Furthermore, the price formation of the Nordic system price based on the evolution of the supply and demand curves was also described. The next section will present the relevance of hedging along with risk management tools offered in the financial market for Nordic power.

3. Risk Management

As mentioned, Nasdaq offers various types of derivatives for hedging and trading purposes. In this thesis, only hedging with futures will be analyzed, but a short discussion on how futures differ from forward contracts is found necessary. Two popular pricing theories for commodity futures will also be described as well as how they relate to electricity futures. Finally, the concept of basis risk and what implication this has for hedgers will be presented and discussed.

3.1. Relevance of Hedging

The literature uses various definitions of hedging, and one standard definition states that “the objective of hedging is to minimize the risk of the portfolio for a given level of return” (Ghosh, 1993). Hull (2012, p. 10) highlights that actors in derivatives markets can be categorized as hedgers, speculators, or arbitrageurs. Speculators seek to profit from derivatives trading by betting on futures price movements (Hull, 2012, p. 10) and arbitrageurs aim to make a riskless profit by trading in two or more markets simultaneously (Hull, 2012, p. 15). Hedgers engage in the market for risk reduction and predictability purposes. It is both feasible and reasonable for several actors to engage in hedging activities as they typically have commitments in the future. Because this thesis examines hedging as a way of managing risk, hedgers will be the actors of focus in this thesis.

Besides the risk reduction perspective of hedging, Brown, Crabb and Haushalter (2006) propose additional explanations for the extended use of hedging in electricity markets. They state that firms use selective hedging to identify potential value creation, and that past success can encourage managers to participate in hedging. Furthermore, the lack of a perfect theory on optimal hedge ratios allows for a variation of hedging decisions to be justified (Brown, Crabb, & Haushalter, 2006).

Sanda et al. (2013) surveyed Norwegian hydro-based electricity companies related to how they use power derivatives to hedge electricity price risk. Although the stated hedging policies differed across the companies, all the companies in the study engaged in power derivatives trading for hedging purposes (Sanda et al., 2013). This explains how the power derivatives at Nasdaq Commodities serve as valuable risk management tools for major actors in the industry, and that hedging is of high relevance in the Nordic power market.

3.2. Forward Contracts

A forward contract is a non-standardized agreement between two counterparties to buy or sell a security or commodity in the future at a price specified at the time of the agreement (Tuckman, 2002). A trader can take a long position by committing to buy a commodity, or a short position by committing to sell the commodity (Jovanovic, 2014, p. 6). The advantages of trading with forward contracts are that they reduce uncertainty, they are negotiated deals that offer great flexibility, and they are settled at maturity (Jovanovic, 2014, p. 6). The disadvantages are that it can be hard to find counterparties, it typically requires some guarantee, and the participants are subject to default risk (Jovanovic, 2014, p. 6). Forward contracts are traded OTC (over-the-counter) and are not marked-to-market. Consequently, the profit or loss is realized at maturity.

3.3. Futures Contracts

A futures contract is a standardized agreement between two counterparties of an exchange that will take place at a future date (Choudhry, 2007). When a party establishes a position in a futures contract, it can either run this position to maturity as with a forward contract, or close out the position before the maturity date (Choudhry, 2007). This allows the trader to sell a commodity that he does not have (Jovanovic, 2014, p. 10). If the position is held to maturity, physical settlement will take place. This involves delivery of the underlying asset specified in the contract. However, most futures contracts are closed out before delivery (Jovanovic, 2014, p. 10).

The clearinghouse, in this case Nasdaq Commodities, requires all market participants to deposit a margin with the exchange, and the size of the margin will depend on the size of the position held (Choudhry, 2007). The clearinghouse also acts as the ultimate buyer and seller to prevent default risk (Choudhry, 2007). Another characteristic of futures contracts distinguishing them from forward contracts are the daily settlements where the futures position is marked-to-market. This makes it possible to trade on the cash value on the futures contract (Bøhren, Michalsen, & Norli, 2012), which is not possible when trading forward contracts. If the price of the underlying asset falls such that the current margin does not cover the loss, additional funds must be deposited to the margin account (Jovanovic, 2014, p. 10). Futures contracts are widely used instruments for hedging, and Moschini and Myers (2002) highlights that hedging with futures reduce risk since spot and futures prices tend to move together, implying that spot price changes can be partially offset by price changes in the opposite futures position.

3.4. Pricing of Electricity Futures

This sub-section will review the standard theories on the pricing of commodity futures and how they relate to electricity futures.

Pricing of futures and forward contracts can be different, and factors affecting these differences are taxes, volatility in interest rates, transaction costs and the treatment of margins (Hull, 2012, p. 112). Nevertheless, according to Hull (2012, p. 112), the price of a futures contract is, in most cases, identical to the one of a corresponding forward contract when the risk-free rate is constant. He also highlights that the differences that may arise in short-term contracts when the risk-free rate is constant are in most situations sufficiently small to be ignored (Hull, 2012, p. 112). Wimschulte (2010) examined the forward and futures price differential in the Nordic electricity market and found no significant differences between the two. This also fits with the dataset retrieved from Bloomberg, which contains both forward and futures prices showing no differences.

There are in general two widely used pricing theories for commodity futures, which are the *theory of storage* and the *risk premium theory* (Fama & French, 1987). The theory of storage links the spot and futures prices through a no-arbitrage condition using interest rates, costs associated with storing the commodity, and the convenience yield of holding the commodity. This no-arbitrage argument is also called the *cash-and-carry trade*, which involves taking a long position in a commodity and a short position in a futures contract on the same commodity, thereby exploiting potential mispricing in the market (Fiorenzani, 2006, p. 85). This is possible when considering the principles of portfolio theory and replication, implying that the value of a portfolio replicating the payoff structure of another asset needs to have the same value as this particular asset (Bodie, Kane, & Marcus, 2011, p. 718). The spot-futures relationship is thus given by:

$$F_{t,T} = S_t e^{(r+u-y)T} \tag{1}$$

where S_t and $F_{t,T}$ are the spot and futures price at time t, respectively, r is the risk-free rate per annum with continuous compounding, u is the storage cost, y is the convenience yield[8], and T is the time to maturity of the futures contract in years (Hull, 2012, p. 120).

The risk premium theory has a different approach when linking the spot and futures prices. The risk premium theory assumes that the futures price contains the power to predict future spot prices and that the futures price, therefore, equals the expected future spot price plus a risk premium (Fama & French, 1987). According to this theory, the relationship between spot and futures prices can be formulated as:

$$F_{t,T} = E_t(S_T) + P_{t,T} \tag{2}$$

[8] The convenience yield reflects “the market’s expectations concerning the future availability of the commodity” (Hull, 2012 p. 120). For a full elaboration of the storage theory, see Hull (2012, p. 117-120).

where $F_{t,T}$ is the futures price at time t for delivery at time T, $E_t(S_T)$ is the time-t expected spot price for time T, and $P_{t,T}$ is the risk premium component (Fama & French, 1987).

Regarding pricing of electricity futures, both mentioned theories are applied in existing research. Most studies on the field are critical to the storage theory when it comes to electricity futures (Bessembinder & Lemmon, 2002; Weron & Zator, 2014). This is due to the non-storable characteristics of electricity, which is a necessary condition for the storage theory to hold. That being said, the Nordic power market is dominated by hydro-based electricity companies, and approximately 50% of the electricity on Nord Pool is generated from hydropower plants (Botterud et al., 2009). For that reason, Botterud et al. (2009) argue that the storage theory can be applied for the Nordic market, given that large hydro reservoirs make storage of electricity possible to some extent, noting further that significant changes in the reservoir levels have a substantial impact on the spot prices.

Most of the research focuses on the risk premium theory when analyzing the spot-futures relationship for electricity prices, and this theory has therefore acquired a greater acceptance in the electricity market. One of the first to present a risk premium model for electricity futures were Bessembinder and Lemmon (2002). They presented an equilibrium model with the assumption that the prices are determined by industry participants and not outside speculators. They found that the futures price is generally a biased forecast of the future spot price and that the futures prices exceed expected spot prices when the expected demand or demand volatility is high (Bessembinder & Lemmon, 2002). Gjolberg and Brattestad (2011) argue that the risk premium depends on the hedging demand in the market, and if the demand is balanced, the futures price will equal the expected future spot price. Despite an increasing amount of literature on the field, it cannot be said to be one entirely accepted theory explaining the prices of electricity futures, and prices will typically depend on a wide array of factors.

3.5. Basis Risk

Basis and basis risk are important concepts when it comes to hedging. The basis can be defined as “the difference between the spot price and the futures price of a commodity” (Hull, 2012, p. 792):

$$\text{Basis} = F_{t,T} - S_t = E_t(S_T) - S_t + P_{t,T} \tag{3}$$

Equation (3) shows how the basis, $F_{t,T} - S_t$, depends on the change in spot price and the expected risk premium (Huisman & Kilic, 2012). The sign of the basis can be used to characterize a market to be either in a *contango* or a *normal backwardation* situation (Hull, 2012, p. 123). A market in which the futures price is below the current spot price is said to be in a normal backwardation condition, and the opposite is called a contango market (Hull, 2012, p. 123). Botterud et al. (2009) analyzed the Nordic market with data spanning from 2002 to 2006, and found that seasonal patterns largely drove the market situation. Specifically, the market tended to be in normal backwardation in the first half of the year and contango in the second half (Botterud et al., 2009). In general, hedgers prefer to be net long in the case of contango and net short in the case of normal backwardation (Lee & Zhang, 2009)[9]. That is because futures prices are falling in contango markets and rising in normal backwardation markets.

If the underlying asset of the futures has the same characteristics as the asset being hedged, the basis will converge to zero when the futures contract approaches maturity (Hull, 2012, p. 26). However, this thesis examines the spot price of electricity and the futures prices of electricity with monthly and quarterly delivery periods. This implies that the asset underlying the futures contracts are not identical to the asset being hedged since the spot price applies for delivery of electricity for the following day. Consequently, this creates a basis risk for a hedger holding a contract until maturity. As emphasized by Byström (2003), the basis risk is especially notable in the electricity market, as there exist large temporary differences between spot and futures prices because of the non-storability of the commodity, thus causing a non-straightforward pricing relationship between the two. This suggests that obtaining satisfying hedging results could turn out to be difficult when compared to other energy markets.

This section has given an overview of forward and futures contracts, which are the most traded derivatives in the Nordic electricity market. The section has also described the pricing of futures contracts, as this will be the hedging instrument applied in the empirical analysis of the thesis. A short description of basis risk has also been presented to show how this risk can affect hedging decisions.

[9] For more information on normal backwardation and contango markets, see Lee & Zhang (2009).

4. Literature Review

The following will present a literature review on hedging in power markets, including different models of estimating the optimal hedge ratio with their respective empirical findings. This will create the basis for the selected models used for the analysis in this thesis, which will be presented in detail and analyzed in section 6 and 7, respectively.

As previously mentioned, the liberalization of power markets has led hedging with derivatives to become a common practice among most participants in the industry (Sanda et al., 2013). This has given rise to research on the area in order to obtain insights and improve understanding of how to better manage electricity price risk. The majority of research on the field focus on how to decide on the optimal hedge ratio, which is defined as "the ratio of the size of a position in a hedging instrument to the size of the position being hedged" (Hull, 2012, p. 823).

Research regarding optimal hedge ratio modeling in general is broadly covered. The optimal hedge ratio is often defined as the hedge ratio that minimizes the portfolio variance and is therefore commonly referred to as the *minimum-variance hedge ratio* (Ederington, 1979). Johnson (1960) introduced an approach for determining the optimal hedge ratio in a spot-futures[10] portfolio considering the existence of basis risk. Ederington (1979) further developed this concept by proposing a theoretical framework for hedging effectiveness in which hedging strategies are measured in terms of their percentage reduction in the return variance of the hedged portfolios compared to an unhedged portfolio, which will be described in detail in sub-section 6.4.1.

In other commodity markets, dynamic hedging models reduces in-sample portfolio variance significantly better than static hedging models in commodity markets (Kroner & Sultan, 1993; Brooks, Henry, & Persand, 2002), but show contradictory results when it comes to out-of-sample hedging effectiveness (Myers, 1991; Lin & Granger, 1994; Yang & Awoke, 2003). Yang and Awoke (2003) examined the hedging effectiveness of storable and non-storable agricultural commodity futures markets over the period 1997-2001. Using multivariate GARCH models, they found strong hedging effectiveness for all

[10] A portfolio containing a position both in the spot market and in the futures market.

storable commodities, but weaker for non-storable commodities (Yang & Awoke, 2003). These contradictory results suggest a critical view of an active risk management strategy for non-storable commodities, such as electricity.

Literature regarding hedging in power markets is limited compared to the broader hedging literature. The first to investigate hedging effectiveness in the Nord Pool market was Byström (2003). Byström (2003) examined whether hedging electricity prices with futures could reduce the variability of portfolio returns. This was done through an analysis with data spanning from 1996 to 1999 in which the sample was split into an in-sample estimation period and an out-of-sample test period of equal length. The test period was both analyzed in its entirety as well as split into three sub-periods. From this, he computed both dynamic and static hedge ratios constructed from five different models to analyze which were the most effective (Byström, 2003). The static models analyzed included the *naïve hedge* and the *OLS hedge ratio*, while the dynamic models included a moving average model and two versions of the *multivariate GARCH model*. Weekly futures contracts were used, and the dynamic hedging models were rebalanced daily (Byström, 2003). He found that all hedging models in the study reduced the portfolio variance, suggesting that power futures are adequate hedging tools for an actor in the industry. However, all the dynamic hedging models performed worse than the static models. As for the results, he reported a variance reduction of 17.79% for the naïve hedge, which unexpectedly showed to be the best performing model in the out-of-sample period. Furthermore, it is also worth noting the substantial differences across the three sub-periods. A variance reduction of 66.91% for the best-performing model in sub-period 3 (naïve) were reported, while the best performing model in sub-period 2 (OLS) only achieved a variance reduction of 6.04% (Byström, 2003).

Just as Byström (2003), Zanotti et al. (2009) also carried out a study on futures hedging with static and dynamic hedge ratios in power markets. In addition to Nord Pool, they included the European Energy Exchange (EEX/Phelix) market and the Powernext market[11], and the study was performed based on data from the period 2004-2006. Unlike Byström, they found that the dynamic hedges outperformed the static hedges over the examined period. Overall, they found the highest hedging effectiveness for Nord Pool

[11] In 2008, EEX (Phelix) and Powernext merged and is now called EPEX SPOT SE (the European Power Exchange) (EEX, 2017).

and the lowest for Powernext. Another key finding was the large variations in the hedging effectiveness both across time and across the different models (Zanotti et al., 2009). In particular, they reported that the GARCH models are most effective when volatility in the market is relatively high (Zanotti et al., 2009).

The non-storability property of electricity makes the cash-and-carry trade inappropriate (Skantze & Ilic, 2001, p. 54) and is explained by Torró (2009). Torró (2009) finds that the power markets' characteristics of high kurtosis, high volatility, jumps, positive skewness, mean-reversion, seasonality, and heteroscedasticity combined lead to an unusual low correlation between spot and futures prices. He points out that this could lead to poor performance of hedging strategies unless more sophisticated models are applied. He further shows that a better hedging performance can be obtained by using the model of Ederington and Salas (2008). This model minimizes the portfolio variance with the use of spot price forecasts under the assumption that changes in spot prices are partially predictable (Torró, 2009), which is a common feature of energy prices, as spot prices are partially predictable due to weather and demand seasonality. Torró (2009) also mentions that the poor effectiveness of the hedging strategies reported in previous studies is found because standard hedging approaches often underestimate the effectiveness of hedging.

Hanly et al. (2017) add to the studies by Byström (2003) and Zanotti et al. (2009) on futures hedging in power markets. In addition to the markets studied by Byström (2003) and Zanotti et al. (2009), they also covered futures hedging in the British market (APXUK), with data from 2004 to 2014. They examined both weekly and monthly hedging horizons by applying static and dynamic models. Adding to the studies of Byström (2003) and Zanotti et al. (2009), they used the percentage reduction in *Value at Risk* (VaR) as a measure of downside risk in addition to variance reduction when ranking the hedging performance of the models. They reported the best results for the OLS model for all weekly hedges and in 50% of the cases for the monthly hedges (Hanly et al., 2017). In cases where the GARCH model outperformed the OLS model, it was only for the Nord Pool market that a significant difference was found. Regarding the in-sample analysis, they reported a variance and VaR reduction from the GARCH model of 27.37% and 15.39%, respectively. Both of these risk reductions are significantly higher than those obtained with the OLS model. The results from the out-of-sample were somewhat lower with variance and VaR reductions from the GARCH model of 17.10% and 10.82%, respectively. However, the GARCH model was only significantly outperforming the OLS model when measured by VaR reduction in the out-of-sample analysis.

By analyzing historical data with different contract lengths and holding periods from 1996 to 2014, Byström (2003), Zanotti et al. (2009) and Hanly et al. (2017) show contradictory results when examining hedging effectiveness and performance, and present different results and relationships between static and dynamic hedging models. Since the Nordic electricity market is still young and evolving, the contradictory results give incentive for further research on the area by including more recent price observations in the analysis.

5. Data

This section will present the data material applied in the thesis. First, the spot and futures data used in the analysis will be presented. A description of how the returns are calculated and how the frequency for the data is chosen will follow. Thereafter, the rollover procedure for the futures will be described along with a description of the different sample periods examined in the analysis. Last, the concept of stationarity for time series data will be discussed, and a statistical test for this purpose will be conducted.

5.1. Spot and Futures Data

As previously mentioned, the spot price data for the analysis are daily data on the Nordic system price, which is set by Nord Pool. The prices are retrieved from Nord Pool's FTP server, and the daily system prices are computed as the daily arithmetic average of the hourly prices within each day in the sample. The spot price data spans from November 24, 2005, to November 21, 2018.

The futures price data has the same start and end dates as the spot data described above. Furthermore, the futures in the analysis have two different delivery periods: monthly and quarterly (three months). This implies that the first monthly and quarterly contract in the analysis applies for delivery of electricity throughout January 2006 and the first quarter (Q1) of 2006, respectively, both with December 30, 2006, as the last trading day. The last monthly and quarterly contracts in the analysis are the contracts for delivery in December 2018 and the first quarter (Q1) of 2019, respectively. This ensures enough data to form equally long return series for the different contract lengths. The monthly contracts can be traded six months in advance of maturity, while quarterly contracts can be traded up to two years before maturity (Nasdaq OMX, 2018). In total, 209 monthly futures contracts and 53 quarterly contracts are included in the dataset. The holding period of each futures, as well as the rollover dates, will be described in sub-section 5.3.

All futures prices in the analysis are daily base load settlement prices. *Base load* means that the futures cover delivery for all hours of all days in the delivery period, as opposed to *peak load* which only covers the hours 08:00 – 19:59 CET of the relevant day (Nasdaq OMX, 2018). Base load contracts are the most pronounced contracts in the literature (Byström, 2003; Zanotti et al., 2009; Hanly et al., 2017), and is for that reason chosen for this thesis. The prices are denominated in Euro and applies for a base size of 1 MWh for each contract. This means that one monthly futures contract or one quarterly futures contract implies delivery of 1 MW of electricity every hour for an entire month or quarter, respectively (Nasdaq OMX, 2018).[12] The settlement price of a futures contract is the price used for calculating daily gains and losses and potential margin requirements (Hull, 2012, p. 35). For the electricity futures, the settlement price is a theoretical price where a transaction could have taken place based on bids and asks in the market and pricing of other energy sources, such as oil and gas (Skjevrak, 2019). Settlement prices are chosen for the analysis to be consistent with existing research (Zanotti et al., 2009) and KIKS' analyses (Skjevrak, 2019). All futures contracts at Nasdaq OMX Commodities have cash settlement only, meaning that they are settled in cash instead of physical delivery (Nasdaq OMX, 2018).

High	**Low**	**Settlement**	**Close**	**Volume**	**Date**	**Contract Name**
73	73	73	73	16	09-07-2008 00:00	Jan-09
70	69	70	70	31	10-07-2008 00:00	Jan-09
71	70,95	72,5	70,95	2	11-07-2008 00:00	Jan-09
		72		0	14-07-2008 00:00	Jan-09

Figure 5 *- Small extraction of the dataset - monthly futures contracts.*[13]

[12] Base load months are normally in the range of 672-745 hours and base load quarters are normally in the range of 2159-2209 hours (Nasdaq OMX, 2018).

[13] The column in figure 5 shows prices from the first possible trading days for Jan-09 contracts, hence low traded volume.

The first two rows show how the data looks when the volume traded is relatively large. When volume is high, the settlement price will lie within the range of the lowest and highest price during the trading day. The third row shows another characteristic with the settlement price, which is that it is not bound by this range when liquidity is low. This is possible because the settlement price accounts for all possible trades in the market and not only the ones that have taken place. Prices can vary a lot during a day, and the settlement price can therefore be calculated below, inside, or above the high/low range (Skjevrak, 2019). The fourth row shows the advantage of using settlement price instead of the closing price. Days with no trades show no closing price, while the settlement price is reported based on Nasdaq's calculations of a theoretical market price. That being said, the settlement price is in most cases equal to the closing price.

The main reason why monthly and quarterly futures are chosen for the analysis is because they are the most liquid contracts in the market (Botterud et al., 2009). Monthly contracts are most liquid with most price data available (Redl & Bunn, 2013), and this contract length is also used in other studies on the field (Hanly et al., 2017; Zanotti et al., 2009). Quarterly contracts are less pronounced in the literature, but serve as valuable risk management tools for numerous actors in the industry (Håkonsen, 2019).

5.2. Return Computation and Sampling Interval

As Stock & Watson (2015, p. 572) point out, economic time series often exhibit approximately exponential growth, and the standard deviations of the series are usually proportional to their levels. For that reason, it is expedient to analyze economic time series as changes in their logarithms (Stock & Watson, 2015, p. 572). Considering this, the returns in this study will be computed as the logarithmic difference between the price at time t and time $t-1$:

$$r_t = \ln(p_t) - \ln(p_{t-1}) \tag{4}$$

For the spot prices, p_t refers to the system price at any day t. As for the futures prices, p_t denotes the settlement price of the particular futures contract for a given day t in the sample.

What also needs to be decided is the sampling interval to use when computing the returns. This is discussed by Stoll & Whaley (1993, p. 59), who argue that the choice of sampling interval of returns is essentially a compromise between capturing more information in the data with smaller intervals versus eliminating more noise from the data by choosing a longer interval, thus minimizing the uncertainty in the estimated returns. By examining daily, weekly and biweekly returns for the S&P 500 index futures, Stoll & Whaley (1993, p. 59) found the optimal sampling interval to be weekly returns. The main reason for this finding was that weekly data successfully removed bid-ask effects present in the daily returns. For that reason, weekly non-overlapping returns will be used in this analysis. Since the futures market is closed in the weekends and holidays, five trading days are used when computing weekly returns. Weeks that go from December to January are merged in order to obtain a consistent dataset such that all weeks include five days. By excluding all non-trading days, the dataset consists of 652 weekly observations for each of the return series; spot returns, monthly futures returns, and quarterly futures returns. Considering hedging literature in electricity markets, this is consistent with the procedure of Hanly et al. (2017) but contradicts the procedure of Byström (2003) and Zanotti et al. (2009) who employs daily returns.

Regarding how often the hedger rebalances the hedged portfolio for the dynamic strategies, Benet (1992) found that hedging is most effective when a weekly rebalancing is used. Malliaris and Urratia (1991) obtained the best results when the rebalancing matched the frequency of the data, also arguing in favor of using weekly data. For that reason, weekly rebalancing has been chosen for the dynamic models, while for the static models, the rebalancing naturally matches the contract's delivery period of one month or one quarter.

5.3. Rolling Over the Futures

What also needs to be decided is how the futures contracts will be rolled over. One alternative is that the hedger holds the contract until maturity before taking a new position in the next nearest contract. However, previous studies on hedging in power markets argue that a better way to represent a typical actor in the market is to roll over to the next futures contract one week before the expiration of the current contract (Byström, 2003; Zanotti et al., 2009). The arguments used for closing out the current position before expiration is to avoid thin markets (low liquidity) and expiration effects, meaning increased volatility in the futures prices close to the maturity date (Byström, 2003). This is consistent with the *Samuelson hypothesis*, which states that the volatility of changes in futures prices increases when the contract is close to maturity (Samuelson, 1965). Samuelson (1965) argues that this is because of higher information asymmetries with longer time to maturity, and vice versa. The idea is that new information about expected market conditions, which in the Nordic electricity market could typically be weather forecasts or information about the balance of supply and demand (Botterud et al., 2009), will lead to a greater reaction in the futures market when the time to maturity is short. By the same logic, information about expected market conditions longer time before the maturity date will have a smaller impact on the futures prices, as the information is less certain. The Samuelson hypothesis has been proven to hold for electricity markets by Koekebakker and Ollmar (2001), as they show that in the short end of the term structure, the volatility increases sharply as the time to maturity decreases. Rolling over to the next futures contract one week (five trading days) before expiration thus prevents thin markets and increases the probability of enough liquidity for trading, and this rollover procedure will therefore be applied in this thesis.

This rollover method implies that the holding period of each futures equals the contract length, meaning that the monthly contracts have a 1-month holding period and the quarterly contracts have a 3-month holding period. Figure 7 shows an example of how the chosen rollover procedure works when applied to the dataset.

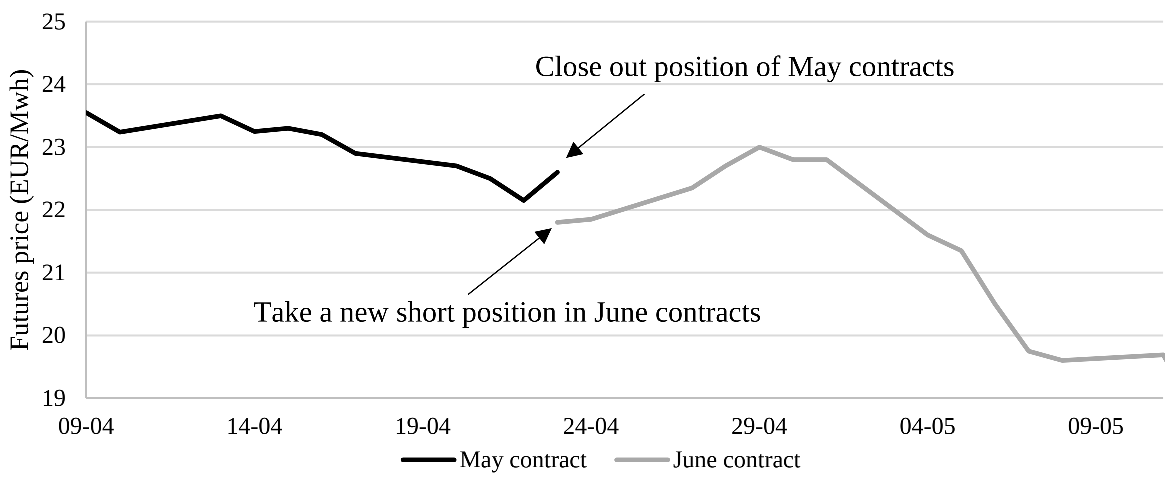


Figure 6 *- An example of how futures contracts are rolled over with the case of monthly futures in 2015. Five days before the expiration of the May contract, that is, on April 23, the hedger closes out the position he holds in May contracts and enters a new short position in June contracts. The same principle applies to the quarterly futures.*

To illustrate how the daily returns are computed, the daily log return on April 23 is computed as $\ln(€22.60/€22.15) = 2.01\%$, where both prices refer to the May contract. The following day, April 24, the return is computed as $\ln(€21.85/€21.80) = 0.23\%$, with both prices now referring to the June contract. Consequently, the logarithmic difference is always calculated over the same contract to ensure the futures returns are correctly computed. As described in sub-section 5.2, five daily logarithmic returns are combined to obtain observations of weekly returns.

5.4. Sample Periods and Descriptive Statistics

This sub-section aims to explain how the data sample is split into sub-periods to prepare it for the analysis. As previously mentioned, the dataset used in the thesis consists of spot and futures prices of Nordic power from 2005 to 2018. The reason for the length of the dataset is that a new market structure was introduced in 2003 (Nasdaq OMX, 2003). Blocks contracts were replaced with monthly contracts and seasons contracts were replaced with quarterly contracts (Nasdaq OMX, 2003). In 2005, Nasdaq started calculating the settlement price for each trading day and quoting each contract in Euro (Nasdaq OMX, n.d.). As a result of this, excluding the years before 2005 will make the dataset consistent in terms of contract lengths of the futures, and ensure the availability of a Euro-denominated settlement price for each trading day.

Table 1 *- An overview of the sub-periods in the full sample*

Period	Date	Events during sub-periods
Sub-period 1	12/01/05 – 02/24/09	Futures contracts quoted in Euro
		Introduction of the Settlement price
		EU ETS regulation is implemented (European Union, n.d.)
Sub-period 2	03/03/09 – 05/25/12	Norway joins the Elbas intraday market
		Implementation of a negative price floor in the Elspot market
Sub-period 3	06/04/12 – 08/28/15	Bidding area and Elbas introduced in Lithuania and Latvia
		Introduction of a new web-based marketplace
Sub-period 4	09/04/15 – 11/21/18	Nord Pool appointed NEMO (Nominated Electricity Market Operator)
		The transition of liquidity from forward to futures contracts

The dataset is split into four sub-periods to improve the robustness of the analysis and to check if the results vary depending on different market situations, which is related to the second sub-question of the thesis. To be able to observe tendencies and thoroughly compare the obtained results, each sub-period contains equally many weekly observations, which is in accordance with previous studies on the field (Bystrom, 2003; Zanotti et al., 2009; Hanly et al., 2017). The dates listed in Table 1 are the ending dates of the weeks used for the weekly returns in the time series, and the dataset ends at 11/21/2018 for both contract lengths to be able to compare the futures prices with the spot price on the same day.

Sub-periods 1-3 function as the in-sample estimation period, while sub-period 4 function as the out-of-sample evaluation period in the thesis. The analysis for sub-period 4 will be among the first studies examining this period. An illustration of how the data sample is split between an in-sample estimation period and an out-of-sample forecast evaluation period can be seen in Figure 7.

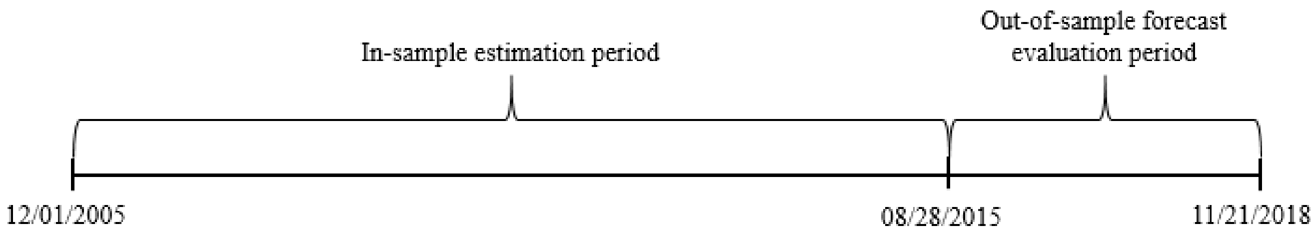


***Figure 7** - Presentation of how the data is split into an in-sample estimation period and an out-of-sample forecast evaluation period*

Sub-period 1 starts right after the implementation of the EU ETS regulations and the introduction of Euro quoting and settlement price calculation of the contracts (Nord Pool, n.d.-g). Sub-period 2 starts 163 weeks after the beginning of sub-period 1, and an important event in this period was that Norway joined the Elbas market, which could make the intraday market less volatile due to the characteristics of the Norwegian electricity market where most of the electricity stems from hydropower production. Another event in this period was the implementation of a negative price floor in the Elspot market (Nord Pool, n.d.-g). Both events could affect the volume traded during this period, and hence also the prices. During sub-period 3, Elbas was introduced in Lithuania and Latvia, and a new web-based marketplace was set up (Nord Pool, n.d.-g). The new web structure could make it easier for new traders to enter the market and hence lead to more market participants. At the beginning of sub-period 4, Nord Pool was appointed NEMO, and the liquidity in the market was switched from forward to futures contracts (Skjevrak, 2019). The reason for this change was that Nasdaq wanted to attract more traders into the market (Skjevrak, 2019). General descriptive statistics of the return series are given in Table 2 and the price series are displayed graphically in Figure 8.

Table 2 *- Descriptive statistics of weekly returns for all return series*

Period	Series	Mean (%)	Std. dev. (%)	Skewness	Kurtosis	JB statistic	Corr.
Full period	Spot	0.06	19.28	-0.13	15.68	4367.34***	-
	Monthly	-0.32	7.20	-0.06	3.01	0.46	0.35
	Quarterly	-0.16	5.72	-0.34	2.18	30.63***	0.26
Sub-period 1	Spot	0.04	12.81	0.41	2.05	42.60***	-
	Monthly	-1.07	8.59	-0.07	0.97	112.62***	0.42
	Quarterly	-0.78	7.22	-0.41	0.76	154.83***	0.33
Sub-period 2	Spot	-0.23	22.48	1.11	16.65	5198.19***	-
	Monthly	-0.26	6.96	0.44	1.31	98.69***	0.48
	Quarterly	0.02	5.50	0.53	0.53	196.34***	0.40
Sub-period 3	Spot	-0.25	18.51	0.01	4.90	98.26***	-
	Monthly	-0.90	7.20	-0.20	8.38	789.58***	0.34
	Quarterly	-0.66	4.55	-1.39	10.85	1886.91***	0.22
Sub-period 4	Spot	0.70	21.92	-1.65	16.94	5577.25***	-
	Monthly	0.94	5.66	-0.02	0.04	16.94***	0.20
	Quarterly	0.79	5.16	-0.16	0.27	205.36***	0.15

*The Jarque-Bera (JB) statistic measures normality of the return series and asymptotically follows a chi-squared distribution with two degrees of freedom (Brooks, 2008 p. 163). The critical values for the test are 4.61 (10%), 5.99 (%), and 9.21 (1%). *, ** and *** indicate rejection of the null hypothesis at the 10%, 5%, and 1% significance level, respectively. The correlation coefficients are between both futures return series and the spot return series in the same period.*

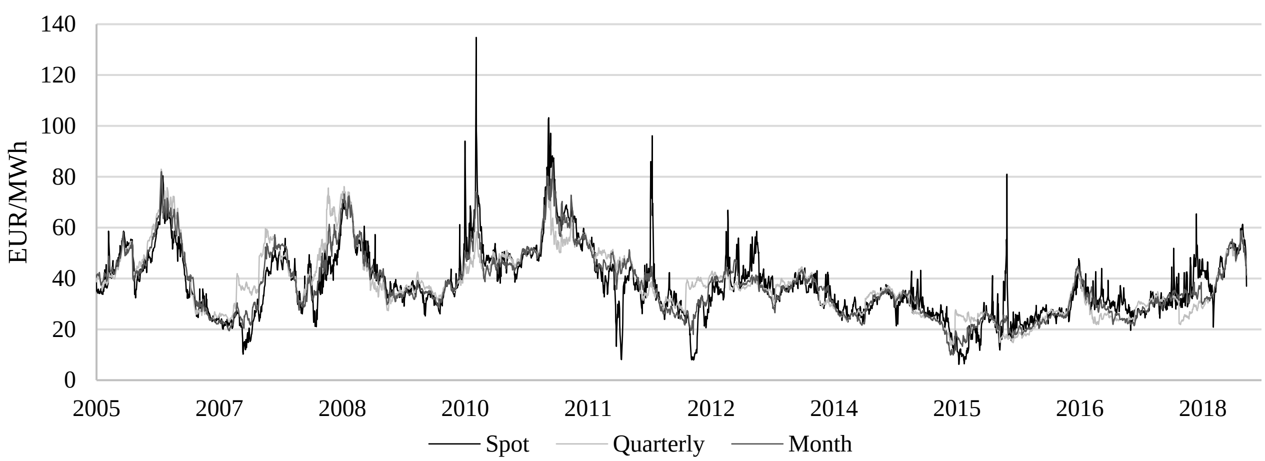


Figure 8 *– Historical spot, monthly and quarterly futures prices*

From Table 2, weekly mean returns show a history of being very low or negative in all sub-periods except sub-period 4. Since the historical low electricity price of 6.23 EUR/MWh in the middle of 2015, the general trend has been steadily increasing electricity prices up to the end of the sample, which results in a high mean return in sub-period 4 for all series.

The standard deviation of a traded asset's logarithmic returns is a common way in finance to describe the volatility of an asset as it measures the deviations from the expected return (Brooks, 2008, p. 383; Bodie et al., 2011, p. 132). Examining the standard deviations of the spot returns in the different sub-periods can be helpful to infer if the performance of the hedging models depends on the spot market volatility. From Table 2, it can be inferred that the spot market is more volatile than the futures market. This is intuitive as the delivery period for the spot prices are for the following day, whereas the futures contracts have delivery periods of one or three months and hence react less to shocks in the market, which is visible in Figure 8. Additionally, the monthly futures returns display slightly higher volatility than the quarterly futures returns, probably due to the same reasons. The results also show that the electricity spot and futures series are much more volatile than other assets, such as crude oil or equities that typically have standard deviations of about 5% and 2.5%, respectively (Hanly et al., 2017). Moreover, it can also be seen that the volatility is changing over time. As an example, the standard deviation for the spot series is only 12.81% in sub-period 1, while it is as high as 22.48% in the period after. This phenomenon will be further discussed in sub-section 6.2.3, which motivates the use of time-varying hedging models.

The distributions' skewness measures "how much a distribution deviates from symmetry" (Stock & Watson, 2015, p. 69). About half of the series have a skewness between -0.5 and 0.5 and are said to be approximately symmetrical (Jani, 2014, p. 114). Distributions with skewness less than -1 or higher than 1 are said to be highly skewed (Jani, 2014, p. 114), which applies for approximately one fourth of the series in the dataset. The last fourth is in the range [-1, -0.5] and [0.5, 1], and is considered to be moderately skewed (Jani, 2014, p. 114). The skewness of the distributions are in other words varying a lot across the different series.

The kurtosis of a distribution measures "how much mass is in its tails" (Stock & Watson, 2015, p. 71), but it is also a measure of how peaked the distribution is around its mean (Brooks, 2008, p. 162). For the return series examined, the kurtoses are mostly above 3. This indicates a leptokurtic distribution, meaning that it has fatter tails and is more peaked around the mean than a normal distribution, which is commonly found in financial time series (Brooks, 2008, p. 162). Five of the series display a kurtosis of less than 3, which indicates a platykurtic distribution with thinner tails and lesser peak around the mean compared to a normal distribution (Brooks, 2008, p. 162). The high and varying skewness and kurtosis of the return data examined are common characteristics for energy time series (Hanly et al., 2017).

The *Jarque-Bera* test statistic measures whether the series have a normal distribution and is given by:

$$JB = n\left[\frac{\text{skewness}^2}{6} + \frac{(\text{kurtosis} - 3)^2}{24}\right] \tag{5}$$

where n denotes the number of observations (Jarque & Bera, 1980). The null hypothesis of the test is that the series has a normal distribution, and it is rejected at the 1% level of significance in all periods and for all return series, except for the monthly contract in the full period. As a result, the series cannot be confirmed to be normally distributed. This is common for general financial asset returns and also said to be a stylized fact for electricity markets (Chevallier, 2010).

The correlation between the spot returns and the monthly futures returns are higher than for the quarterly futures returns for all series analyzed. This makes intuitive sense as it is plausible that contracts with shorter delivery periods capture more of the current fundamental market information than contracts with longer delivery periods. These results give reason to expect higher hedging effectiveness from hedging with monthly futures compared to quarterly futures. The reason for this is that the idea of hedging is to offset the risk of adverse price changes by a similar price change in the hedging instrument, in which an opposite position is taken. To what extent a price reaction in the spot market is causing a price change in the hedging instrument is measured by the correlation coefficient and is hence a determinant of hedging performance (Charnes & Koch, 2003).

5.5. Stationarity

Stationarity is an important concept when it comes to time series modeling. According to Stock and Watson (2015, p. 587), a time series is stationary "if its probability distribution does not change over time". The general idea of stationarity is that historical relationships can be generalized to the future (Stock & Watson, 2015, p. 587). If a time series does not fulfill the requirements of stationarity, it is said to be *non-stationary*. A consequence of non-stationarity is that shocks to the system will not die out, which is problematic when modeling time series (Brooks, 2008, p. 230). An alternative way of saying that a time series is non-stationary is to say that it contains one or more *unit roots* (Stock & Watson, 2015, p. 600).

A way of testing for stationarity in time series is to use the *Augmented Dickey-Fuller* (ADF) test (Stock & Watson, 2015, p. 603). Although there exist other tests for this purpose, the ADF test is the one that is most used in practice and shown to be one of the most reliable (Stock & Watson, 2015, p. 603). The ADF test relates to a regression in which the first difference of the time series in question, ΔY_t, is regressed on the first lagged observation of the series and first-differenced lagged values, ΔY_{t-i} (Stock & Watson, 2015, p. 603). The regression can be expressed as:

$$\Delta Y_t = \beta_0 + \delta Y_{t-1} + \sum_{i=1}^{p} \gamma_i \Delta Y_{t-i} + \varepsilon_t \tag{6}$$

where β_0 denotes the intercept, δ is the coefficient of Y_{t-1}, γ_i denotes the coefficients of the first-differenced lagged values and ε_t is the error term. The null hypothesis of the test is that $\delta = 0$, implying that the series contains a unit root. The alternative hypothesis is that $\delta < 0$, or in other words, the series is stationary (Stock & Watson, 2015, p. 605).[14]

An alternative way of specifying the ADF test is to include a deterministic trend in the regression, which would transform the alternative hypothesis to be that the series is stationary around a deterministic time trend (Stock & Watson, 2015, p. 605). However, by looking at the series graphically in Figure 9 below, none of them seems to display any clear trends as they are all fluctuating around a zero mean. Hence

[14] For more details on unit roots and the Augmented Dickey-Fuller test, see Stock & Watson (2015) p. 596-605 and Dickey & Fuller (1979).

none of the conducted tests is specified with a deterministic trend. Furthermore, a lag length for the regression in equation (6) needs to be chosen, and studies have shown that it is better to have too many lags than too few when conducting the ADF test, and an application of the Akaike Information Criterion (AIC) is therefore recommended (Stock, 1994; Haldrup & Jansson, 2005).[15] Consequently, the lag length of the test for each series is estimated by minimizing the AIC.

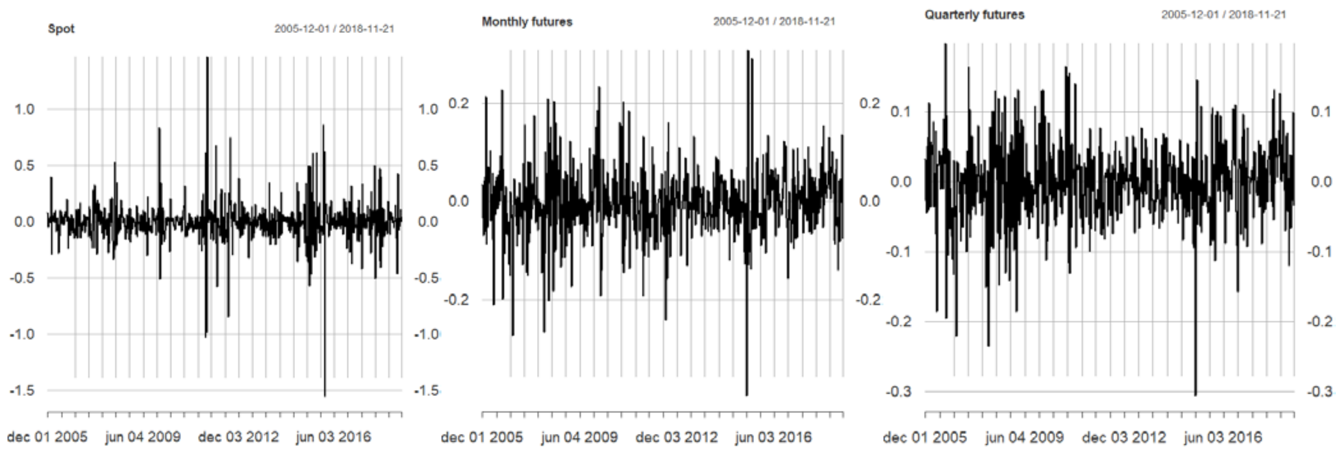


***Figure 9** – Return series displayed graphically: spot, monthly futures, and quarterly futures. Note the different scaling of the vertical axes.*

Although the return series are the essential inputs for hedge ratio modeling, it is also interesting to know whether the price series are stationary or not due to the concept of *co-integration*. Co-integration refers to the case when two series are integrated of the same order d, that is, when they both are $I(d)$ (Enders, 2015, p. 347). This means that they both need to be differenced d times to become stationary (Stock & Watson, 2015, p. 696). If two series are co-integrated, it intuitively means that they share a long-run equilibrium relationship (Enders, 2015, p. 353). In such a situation, their relationship can be more accurately described by adding an *error correction term* (ECT) to the estimated mean models[16] (Enders, 2015, p. 353). This is also adopted in hedging literature by Lien (1996), who stresses that a smaller than optimal futures position will be taken if a co-integration relationship is omitted in the hedging model.

[15] The use of information criterions for lag length selection is described in more detail in sub-section 6.2.2.

[16] The mean models will be more thoroughly described in sub-section 6.2.2

For that reason, ADF tests have been performed for the price series[17] to examine whether error correction terms can improve the accuracy of the models, and ultimately the hedging performance. All the price series have a weekly frequency and apply for the entire sample period. The test results for the ADF test for both the return series and the price series are given in Table 3.

Table 3 *- Augmented Dickey-Fuller (ADF) test for stationarity for return series and price series*

Time series		**ADF statistic**	**Lags used**
Return series	Spot	-11.84***	1
	Monthly futures	-15.29***	1
	Quarterly futures	-8.71***	5
Price series	Spot	-3.13**	6
	Monthly futures	-2.94**	1
	Quarterly futures	-2.94**	3

*The critical values for the ADF test are -3.43 (1% level), -2.86 (5%), and -2.57 (10%). *, **, and *** indicate rejection of the null hypothesis at the 10%, 5%, and 1% significance level, respectively. The lags used in each test are selected according to the AIC.*

As can be seen from Table 3, the null hypothesis of no stationarity is rejected at the 1% level for all return series as the test statistic is more negative than all the critical values. Furthermore, the table also reveals that the null hypothesis of no stationarity for the price series is rejected at the 5% significance level. This means that all the return series and price series are stationary without differencing, or said differently, they are $I(0)$. As a result, they cannot be cointegrated since each is stationary on its own, and the estimation models in section 6 will therefore not include error correction terms.

The electricity demand in the Nordic electricity market shows high evidence of hourly, weekly and yearly seasonality, and the price levels are, therefore, also varying in a seasonal pattern. As previously mentioned, weekly logarithmic returns are used and not the price levels. This leads to weaker evidence of seasonality in general (Byström, 2003). The hourly and weekly seasonality are not affecting the time

[17] The price series are constructed by splicing the futures data into continuous time series according to the *Nth Nearest contract* method one week before maturity. There will therefore be some irrational noise in the price series due to the fact that there cannot be more than one price observation per day. However, this will not affect the return series as each daily return has been computed based on the same contract, as described in sub-section 5.2. For a full elaboration on different splicing methods with their corresponding pros and cons, see Masteika, Rutkauskas & Alexander (2012).

series because of the use of weekly logarithmic returns and the short-term hedging analysis in this thesis will not be affected by the long-term yearly seasonality (Byström, 2003).

This section has presented the data material applied in the thesis, how the returns are computed, how the futures contracts are rolled over, the division of the dataset into sub-periods, a discussion of seasonality in the data, and a test for stationarity in the time series data. To sum up the most essential information from this section, a brief overview will follow. The futures data consist of monthly and quarterly contracts, while the returns are calculated as the weekly logarithmic returns. The holding period for each monthly and quarterly contract is one month and three months, respectively. Rolling over to the next futures contract is done one week prior to the maturity of the current contract, and the dynamic models are rebalanced weekly.

6. Model Selection and Performance Measures

This section will present the hedging models that will be analyzed along with the chosen performance measures. This is to familiarize the reader with the models and the rationale behind the application of them. As previously mentioned, the analysis takes the viewpoint of an actor in the Nordic power market that has committed to selling electricity on the spot market in the future. The actor takes a short position in futures contracts to hedge the price risk of the long position in the spot market. The dynamic models incorporate a time-varying structure of the hedge ratios while the static models assume a constant hedge ratio. The static hedging models that will be presented include the naïve hedge and the OLS model. The dynamic models include two versions of the multivariate GARCH model, which are the constant conditional correlation (CCC) GARCH and the dynamic conditional correlation (DCC) GARCH. Both hedging models will be modeled according to the Box-Jenkins procedure. The measures used for evaluating the models' performance will be presented along with the bootstrapping technique used to assess the statistical significance of the results.

6.1. Static Hedging Models

A static hedging model is a model that does not require a rebalancing of the portfolio when the price or the volatility of the underlying asset changes (Miffre, 2004). The hedge ratio in such models are therefore time-invariant and decided for the entire hedging horizon. Static hedging methods depend less on active risk management and are often preferred by market participants (Sanda et al., 2013). Another potential benefit from static hedging methods is reduced transaction costs, as the hedging position will remain the same during the hedging period. The two most common static hedging models from the literature review, that is, the naïve model and the OLS model, will be described in the following.

6.1.1. The Naïve Hedge

According to traditional hedging theory, hedgers should protect a long position of x units in the spot market by simultaneously selling x units of futures contracts (Anderson & Danthine, 1980; Kroner & Sultan, 1993). This implies a hedge ratio of 1 and is often called the naïve hedging strategy or a 1-to-1 hedge (Kroner & Sultan, 1993). However, this hedging method ignores the basis risk facing the hedger and will, therefore, be a suboptimal strategy in most cases (de Jong, De Roon, & Veld, 1997). As it does not account for the imperfect correlation between spot and futures prices, the naïve hedge transforms spot price risk into basis risk (de Jong, De Roon, & Veld, 1997). Despite this drawback, the naïve hedge has the advantages of being both simple to implement as well as not requiring rebalancing over time.

The primary motivation for including the naïve hedging strategy in the analysis is for comparison purposes, as it is a widely used benchmark model in most studies on hedge ratio modeling in general (Kroner & Sultan, 1993; Garcia, Roh & Villaplana., 1995) and in electricity markets (Byström, 2003; Zanotti et al., 2009; Hanly et al., 2017).

6.1.2. Ordinary Least Squares (OLS)

The *ordinary least squares* (OLS) estimated hedge ratio builds on the principles of portfolio theory and has been popularized by Ederington (1979). This hedge ratio is found simply by minimizing the return variance of a spot-futures portfolio, and it is therefore also commonly referred to as the minimum-variance hedge ratio (Hull, 2012, p. 57). By denoting $r_{s,t}$ and $r_{f,t}$ as the returns in the spot and futures market, respectively, and β as the static hedge ratio, the return of the spot-futures portfolio $r_{\pi,t}$ is given by:

$$r_{\pi,t} = r_{s,t} - \beta r_{f,t} \tag{7}$$

From the properties of the variance of an expected value, the return variance of the portfolio is defined as (Rohatgi & Ehsanes Saleh, 2015, p. 6):

$$var(r_{\pi,t}) = var(r_{s,t}) + \beta^2 var(r_{f,t}) - 2\beta \cdot cov(r_{s,t}, r_{f,t}) \tag{8}$$

where a variance-minimizing hedger solves:

$$\min_{\beta} var(r_{s,t}) + \beta^2 var(r_{f,t}) - 2\beta \cdot cov(r_{s,t}, r_{f,t}) \tag{9}$$

By minimizing the return variance of the portfolio with respect to the hedge ratio, β, the optimal hedge ratio, β^*, can be expressed as[18]:

$$\beta^* = \frac{cov(r_{s,t}, r_{f,t})}{var(r_{f,t})} = corr(r_{s,t}, r_{f,t}) \frac{\sigma_{r_{s,t}}}{\sigma_{r_{f,t}}} \tag{10}$$

where $\sigma_{r_{s,t}}$ and $\sigma_{r_{f,t}}$ denote the standard deviations of spot and futures returns, respectively. The expression in equation (10) is equivalent to that of the estimated slope coefficient in an OLS regression (Stock & Watson, 2015 p. 163). Thus, the optimal hedge ratio can also be found by running the following linear regression:

$$r_{s,t} = \alpha + \beta r_{f,t} + \varepsilon_t, \qquad t = 1,2,\dots,n \tag{11}$$

where α is the intercept of the population regression line, β is the slope of the population regression line, and ε_t is the error term assumed to follow a normal distribution (Stock & Watson, 2015, p. 159). The estimated slope coefficient from the regression, $\hat{\beta}$, will correspond to the optimal hedge ratio in equation (10). The estimated parameters from the OLS regression line is found by minimizing the sum of squared residuals (Brooks, 2008, p. 33).

The static OLS hedging model has one clear advantage compared to the traditional naïve hedging model as it recognizes the imperfect correlation between spot and futures prices (Ederington, 1979). The static OLS hedge ratios in the analysis of this thesis will be computed by running the regression presented in equation (11). It should also be noted that no other static hedge ratio can outperform the OLS hedge ratio based on variance reduction in-sample, as it is the single hedge ratio that on average reduces the most

[18] For a full derivation of equation (10), see Ederington (1979)

variance over the sample (Hull, 2012, p. 57). This is however not the case when considering the out-of-sample analysis, and the comparison of the OLS estimated hedge ratio and the naïve hedge ratio is hence more interesting for this part of the analysis.

In an OLS regression, the R^2 represents the fraction of the variance in the dependent variable that can be explained by the independent variable(s) (Stock & Watson, 2015, p. 823). Therefore, the R^2 can be interpreted as the proportion of the sample variance in a spot portfolio that can be eliminated by hedging with futures (Hull, 2012, p. 58). A higher R^2 will thus imply a greater variance reduction.

6.1.2.1. Assumptions of the OLS model

The traditional OLS model has five underlying assumptions. The assumptions concern the disturbance terms and the interpretation of them, and will need to hold in order to make valid inferences about the actual coefficient values from the estimated parameters (Brooks, 2008, p. 44). If assumptions 1-4 hold, then the BLUE (best linear unbiased estimator) of the coefficients will be those estimated by the OLS model (Brooks, 2008, p. 44). In this context, 'best' refers to having the lowest variance among the class of linear unbiased estimators (Brooks, 2008, p. 45). This sub-section will provide a brief presentation of the underlying assumptions for the OLS model in equation (11), as well as the econometric tests conducted. The test results will be reported in Table 4.

Assumption 1: The errors have zero mean

$$E(\varepsilon_t) = 0 \tag{12}$$

The first assumption of the OLS model is that the expected value of the disturbances equals zero. If a regression model includes an intercept, this assumption will always be satisfied (Brooks, 2008, p. 131).

Assumption 2: The variance of the errors is constant and finite over all values of x_t

$$var(\varepsilon_t) = \sigma^2 < \infty \tag{13}$$

This assumption is commonly referred to as the assumption of homoscedasticity. The opposite is when the error terms have a non-constant variance, known as heteroscedasticity. The most common way of dealing with the presence of heteroscedasticity is to use heteroscedasticity-consistent standard error estimates. This is also known as White standard errors, and is easily employed in most statistical software programs (Brooks, 2008, p. 138). This will produce valid standard errors in the presence of heteroscedasticity, at least in large samples (Wooldridge, 2012). If the assumption of homoscedasticity is violated, the estimated coefficients will still be unbiased and consistent, but they will not be BLUE. Consequently, the standard errors could turn out to be wrong, and wrong inferences might be drawn because of this. Detection of heteroscedasticity is done by conducting White's (1980) test for heteroscedasticity for each in-sample period. The test statistic is chi-squared distributed with 2 degrees of freedom under the null hypothesis of homoscedasticity against the alternative of heteroscedasticity (Brooks, 2008, p. 135).

Assumption 3: The errors are linearly independent of one another (no autocorrelation)

$$cov(\varepsilon_i, \varepsilon_j) = 0, \qquad \forall\, i \neq j \tag{14}$$

The consequences of using OLS in the presence of autocorrelation are the same as those when heteroscedasticity is present. Application of a variance-covariance estimator that is consistent in the presence of both heteroscedasticity and autocorrelation (Newey & West, 1987) is the most common way of satisfying this assumption (Brooks, 2008, p. 152). Two statistical tests for detecting autocorrelation are the Durbin-Watson test and the Breusch-Godfrey test (Brooks, 2008, p. 148). The Durbin-Watson test is the simplest as it only tests for serial correlation between an error and its first lagged value, as opposed to the Breusch-Godfrey test which involves doing a joint test for autocorrelation between an error and several of its lagged values simultaneously (Brooks, 2008, p. 148). Only the Breusch-Godfrey test has been applied for the OLS model due to its strong advantages compared to the Durbin-Watson test, and it has been conducted for all sample periods. The error terms, ε_t, from equation (11) are modeled in the following way:

$$\varepsilon_t = \rho_1 \varepsilon_{t-1} + \rho_2 \varepsilon_{t-2} + \cdots + \rho_{52} \varepsilon_{t-52} + \upsilon_t, \qquad \upsilon \sim N(0, \sigma_\upsilon^2) \tag{15}$$

where ρ_i $(i = 1, \ldots, 52)$ is the autocorrelation coefficient between the error term and one if its lagged values, and υ_t is the disturbance term with a mean of zero and assumed to follow a normal distribution. The number of lags of the residuals for the test is set to 52. The rationale for this is to follow the rule of thumb by having the lag length corresponding to the frequency of the data used (Brooks, 2008, p. 149; Asteriou & Hall, 2011, p. 160). Hence, it is tested whether the errors at any point in time are related to any of the errors in the previous year. The null hypothesis is no autocorrelation in the errors, whereas the alternative hypothesis is that at least one of the lagged errors are related to the current error term (Brooks, 2008, p. 148). The test statistic follows a chi-squared distribution with degrees of freedom corresponding to the number of lags specified in the test (Brooks, 2008, p. 149).

Assumption 4: There is no relationship between the error and the corresponding x variate

$$cov(\varepsilon_t, x_t) = 0 \tag{16}$$

This assumption implies that x_t is non-stochastic in repeated samples, meaning there is no sampling variation in x_t and its value is solely determined outside the model (Brooks, 2008, p. 160). It can be showed that if the first assumption holds, the OLS estimator will still be unbiased, even if the regressors are stochastic[19].

Assumption 5: The disturbances are normally distributed

$$\varepsilon_t \sim N(0, \sigma^2) \tag{17}$$

The coefficient estimators will still be BLUE even if this assumption is violated, but it is required to hold in order to make valid inferences regarding the population parameters from the sample parameters (Brooks, 2008, p. 43). A common way of testing for normality in the residuals is the Jarque-Bera (JB) test (Brooks, 2008, p. 161). The JB test statistic, which is determined by the skewness and kurtosis of the residuals' distributions and the sample size, is given by equation (5), and is previously described in sub-section 5.4.

[19] For a formal proof, see (Brooks, 2008, p. 160)

6.1.2.2. Testing the OLS assumptions

This sub-section will present the results of the econometric tests for the OLS assumptions described in the previous sub-section.

***Table 4** – Test results for the OLS assumptions*

Period	Futures contract	Homoscedasticity: White test	No autocorrelation: Breusch-Godfrey test	Normality in errors: Jarque-Bera test
Sub-period 1	Month	6.43**	59.02	27.20***
	Quarter	7.23**	49.65	21.59***
Sub-period 2	Month	4.55	73.72***	1888.46***
	Quarter	1.07	79.22***	184.75***
Sub-period 3	Month	2.48	52.59	2093.51***
	Quarter	0.05	50.51	2116.93***
Full in-sample period	Month	2.44	140.86***	4426.85***
	Quarter	0.56	126.95***	4420.82***

*Critical values (significance levels in parentheses); White and Jarque-Bera (df=2): 4.61 (10 %), 5.99 (5%) and 9.21 (1%). Breusch-Godfrey (df=52): 65.42 (10%), 69.83(5%) and 78.62 (1%). *, **, ***, indicate rejection of the null hypothesis at significance levels 10%, 5% and 1%, respectively.*

Regarding the first assumption concerning zero mean in the error terms, it can be seen from equation (11) that the estimated OLS model is specified with an intercept. Consequently, this assumption is not violated.

The second assumption regarding homoscedasticity is tested by conducting a White test for each sample period. Table 4 reveals that the null hypothesis is rejected in only two of the cases, which are for both models in the first sub-period. Although evidence of heteroscedasticity is not found in most cases, White standard errors are applied for all sample periods. The reason is that in large samples, always reporting only the heteroscedasticity-consistent standard errors has become common practice over the years (Wooldridge, 2012). Therefore, in order to be consistent with the standard procedures from most research papers (Wooldridge, 2012), all standard errors reported are correcting for heteroscedasticity.

For assumption 3, the null hypothesis of no autocorrelation is rejected in two of the four periods examined. These include the full sample period and sub-period 2. These cases are corrected for by using

Newey-West standard errors in order to obtain the correct interpretation for the statistical significance of the regression estimates (Stock & Watson, 2015, p. 650).

As previously mentioned, it can be shown that if the first assumption holds, violation of assumption 4 will not make the OLS estimator biased.

Regarding assumption 5 of normally distributed residuals, the test results show that the null hypothesis is rejected in all cases. However, the *central limit theorem* states that "under general conditions, the distribution of the standardized sample average is well approximated by a normal distribution when *n* is large" (Stock & Watson, 2015, p. 98). According to Stock & Watson (2015, p. 98), this is "typically a very good approximation to the distribution" for sample sizes larger than 100. As all the estimated OLS models have sample sizes well above 100 observations, the central limit theorem ensures that valid statistical inferences about the estimated parameters can be drawn.

As highlighted by Brooks (2008, p. 164), it is often the case in financial modeling that a few 'extreme residuals', known as *outliers*, cause the assumption of normally distributed errors to be rejected.

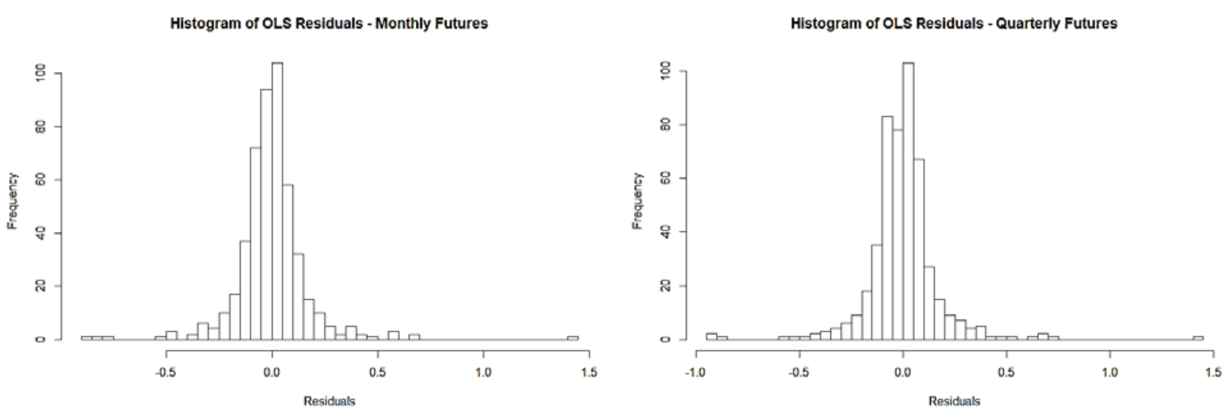


Figure 10 *- Histogram of residuals of estimated OLS model (full in-sample period)*

The histograms of the residuals in Figure 10 looks normally distributed, but there exist a few outliers which are affecting the results of the test. This is confirmed when looking at the normal probability plots in Figure 11, as the distributions approximate a normal distribution when compared to the straight line but outliers on each side cause an S-shape of both plots. This indicates a leptokurtic normal distribution of the residuals, which are commonly found when examining financial time series (Brooks, 2008, p. 162).

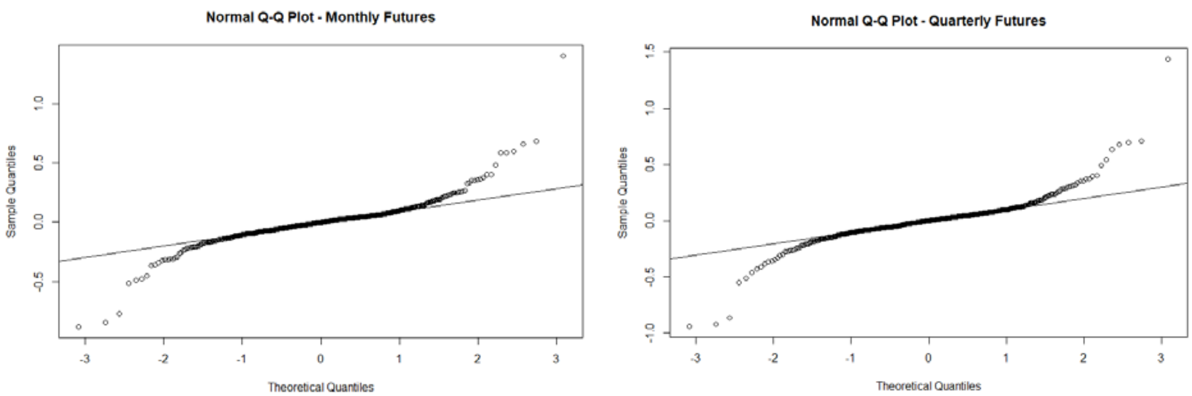


Figure 11 *- Normal probability plot of OLS residuals (full in-sample period)*

Brooks (2008, p. 166) notes that outliers could have a large impact on the coefficient estimates, and some practitioners therefore remove those outlying observations, although it is most common to keep them as all data represents useful information. This is especially important for hedging as the objective is to reduce the risk associated with large outliers, and all observations are therefore kept in the dataset.

6.2. Dynamic Hedging Models

As can be seen from the expression of the OLS estimated hedge ratio in equation (10), the optimal hedge ratio is based on the sample variance of the futures returns and the sample covariance between the spot and futures returns. A critical assumption of Ederington's hedging framework is, therefore, that the volatility in the spot and futures markets is constant over time, implying a static hedge ratio regardless of when the position in futures contracts is entered. This is a strong assumption that contradicts the reality of most financial markets, as it is seldom the case that risk is constant over time (Stock & Watson, 2015, p. 710).

One case of time-varying volatility is denoted as *volatility clustering* and was first described by Mandelbrot (1963) as "large changes tend to be followed by large changes, of either sign, and small changes tend to be followed by small changes". As previously mentioned, electricity markets have characteristics that make them relatively volatile compared to other financial markets, and volatility clusters have been proven to exist also for the Nordic power market (Simonsen, 2005). This can be seen graphically in Figure 12 below, where the volatility of the spot market for Nordic power seems to be higher in some periods than others. For example, the periods from August 2011 to September 2012 and June 2015 to February 2016 can be characterized as volatile, whereas the period from July 2010 to June 2011 seems to be rather tranquil in comparison.

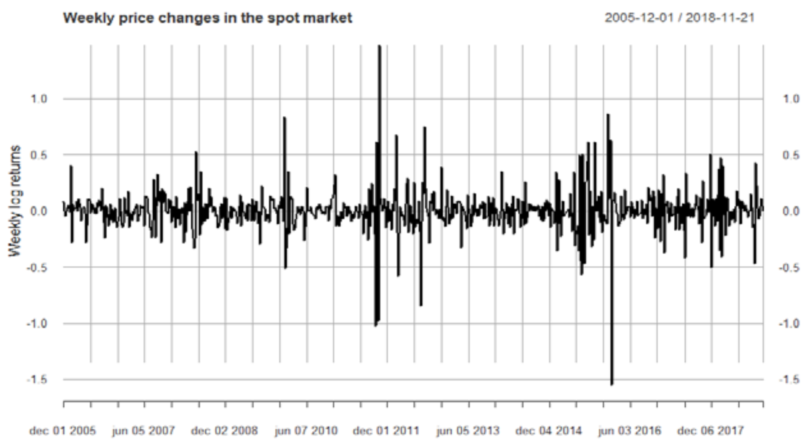


Figure 12 *- Weekly price changes in the spot market (Source: Nord Pool FTP server and own calculations)*

Because volatility clusters in time series data are not uncommon, the framework of Ederington (1979) has been subject for critique in later years. Among others, Kroner & Sultan (1993) argue that since asset prices are characterized by time-varying distributions, the optimal hedge ratio should also be time-varying. Models that can capture this information is therefore assumed to be superior to the static hedging models mentioned previously (Kroner & Sultan, 1993). The second class of hedging models included in the analysis is, therefore, dynamic models with time-varying hedge ratios. For these models, following the notation of Baillie & Myers (1991), the optimal hedge ratio in equation (10) can be slightly modified and expressed as:

$$\beta_{t-1}^{*} = \frac{cov\left(r_{s,t}, r_{f,t} | \Omega_{t-1}\right)}{var\left(r_{f,t} | \Omega_{t-1}\right)} \tag{18}$$

where Ω_{t-1} denotes the information set at time $t-1$, and the rest of the notation is the same as in equation (10). Thus, the distinction between equation (18) and equation (10) is that the variables on the right-hand side of equation (18) are conditional on information from period $t-1$. Consequently, the optimal hedge ratio now contains a time subscript as it is time-varying and dynamically set for each hedging period.

There exist several models that account for time-varying dynamics in covariances and variances. The most basic model is known as the *simple moving average* (SMA) model (Chiulli, 1999, p. 234). One major limitation of this model is that it has a fixed window length with equally weighted observations (Chiulli, 1999, p. 234). An extension of the SMA model is the *exponentially weighted moving average* (EWMA) model, which puts more weight on more recent observations when estimating variance and covariance (Brooks, 2008, p. 384). However, most hedging research use even more sophisticated models such as *autoregressive conditionally heteroskedastic* (ARCH) and *generalized autoregressive conditionally heteroscedastic* (GARCH) models. Examples of prior hedging research employing GARCH models are Kroner & Sultan (1993) for foreign currency hedging, Chang, McAleer, and Tansuchat (2011) who examine crude oil hedging strategies and Hanly et al. (2017) who cover electricity price hedging, among many others. Due to the widespread use of these models in the hedging literature, GARCH models will be modeled for the dynamic hedging strategies in this thesis. The following sub-sections will describe the theory and the specifications of the dynamic hedging models that will be applied in the analysis.

6.2.1. ARCH and GARCH Models

The ARCH model was introduced by Engle (1982). ARCH models are used for modeling volatility clustering by allowing the conditional variance to depend on the squared errors in the preceding periods (Engle, 1982). The model can in its general form be written as:

$$h_t = \omega + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 \tag{19}$$

where h_t is the conditional variance at time t, ε_{t-i}^2 $(i = 1, \dots, q)$ are the squared errors from the preceding periods and α_i $(i = 1, \dots, q)$ and ω are coefficients to be estimated. In ARCH models, the variance estimates are denoted as conditional variance, as they are estimated conditional on past information (Brooks, 2008, p. 387). To distinguish it from the sample variance, which is normally denoted by σ_t^2, the conditional variance is typically denoted by h_t in the literature (Brooks, 2008, p. 388).

To ensure that the conditional variance, h_t, is positive, the coefficients take on the following requirements: $\omega > 0$, $\alpha_i \geq 0 \; \forall \; i = 1, \dots, q-1$ and $\alpha_q > 0$ (Andersen, Davis, Kreiss, & Mikosch, 2009, p. 19). The residuals referred to are from an estimated conditional mean model, making equation (19) only a partial model as it stands (Brooks, 2008, p. 388).

Bollerslev (1986) used Engle's model to develop the more commonly used model called GARCH. The GARCH model estimates the conditional variance as a function of both the lagged squared errors and lagged estimates of conditional variances and can generally be expressed as:

$$h_t = \omega + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j h_{t-j} \tag{20}$$

The notation is the same as in equation (19) with β_j $(i = 1, \dots, p)$ being coefficients on the lagged conditional variances. These are required to be non-negative, $\beta_j \geq 0 \; \forall \; j = 0,1, \dots, p$, to ensure positive variance estimates (Andersen et al., 2009, p. 20). The requirements for the coefficients on the lagged errors and the constant are the same as in the ARCH model.

6.2.2. Mean Model Order Selection

The purpose of multivariate GARCH models is to make estimates and forecasts of variance and covariance (Brooks, 2008, p. 432). However, and as previously mentioned, it is necessary to specify a conditional mean equation when building a GARCH model. The reason is that the lagged squared errors in the conditional variance equation is extracted from a conditional mean equation (Brooks, 2008, p. 388). As Brooks (2008, p. 390) points out, given that the variance is specified around the mean, any misspecification of the mean is likely to lead to a misspecification of the variance. Therefore, it is crucial to devote a fair amount of consideration into the specification of the mean model in order to get as precise as possible estimates of the conditional variances and covariances, even though the mean itself is not of direct interest for hedge ratio computation (Brooks, 2008, p. 390).

The conditional mean equation in GARCH models typically takes the form of an *autoregressive* (AR) model or an *autoregressive-moving-average* (ARMA) model (Enders, 2015, p. 125), as will also the mean equations for the GARCH models in this thesis. An AR model is a model where the dependent variable, which in this case is either the weekly spot or futures returns, is modeled as a function of its own lagged values (Brooks, 2008, p. 215). The AR model can be expressed as:

$$r_t = \mu + \sum_{i=1}^{p} \phi_i r_{t-i} + \varepsilon_t \tag{21}$$

where μ is the constant term, r_{t-i} denotes the lagged values of the dependent variable with corresponding coefficients ϕ_i, and ε_t is the error term. Equation (21) represents an AR(p) model as it includes p lagged values. An ARMA model extends the AR model by modeling the dependent variable as a combination of both an autoregressive (AR) process and a moving-average (MA) process (Brooks, 2008, p. 223):

$$r_t = \mu + \varepsilon_t + \sum_{i=1}^{p} \phi_i r_{t-i} + \sum_{i=1}^{q} \gamma_i \varepsilon_{t-i} \tag{22}$$

The ε_{t-i}'s are representing lagged white noise error terms with corresponding coefficients denoted by γ_i, while ε_t denotes the disturbance term in period t (Enders, 2015, p. 50). Equation (22) includes p autoregressive terms and q moving-average terms and is in this general form referred to as an ARMA(p, q) model (Enders, 2015, p. 51).

The conditional mean and variance models in this thesis are specified according to the *Box-Jenkins methodology*, which is a widely used method for appropriate model selection for time series data (Enders, 2015, p. 76). The method has three steps (Enders, 2015, p. 76):

1) Identification
2) Estimation
3) Diagnostic checking

The identification step involves "determining the order of the model required to capture the dynamic features of the data" (Brooks, 2008, p. 230). In an ARMA context, this means deciding the lag length of the AR(p) and MA(q) terms. According to Stock & Watson (2015, p. 593), choosing the lag length order requires a trade-off between additional information from the lagged values and additional estimation error due to more estimated coefficients. One common way of dealing with this trade-off when choosing the lag order is to use *information criteria* (Brooks, 2008, p.232). An information criterion is a computed statistic that incorporates two factors: one that is a function of the model's residual sum of squares (RSS), and one that penalizes the loss of degrees of freedom when including additional lags (Brooks, 2008, p.232). As a result, adding additional variables to the ARMA model will, on the one hand, decrease the RSS, but the negative impact from the 'penalty term' will also increase (Brooks, 2008, p. 232). When using information criteria for ARMA model selection the objective is to minimize the criterion. Hence, additional lags should only be added to the ARMA model if the decrease in RSS is more than enough to offset the increase of the penalty term (Brooks, 2008, p. 232). Two of the most common information criteria used in practice for ARMA model selection are the *Akaike information criterion* (AIC) and the *Bayes information criterion* (BIC) (Brooks, 2008, p. 232), which are given by:

$$AIC = \ln(\hat{\sigma}^2) + \frac{2k}{T} \tag{23}$$

$$BIC = \ln(\hat{\sigma}^2) + \frac{k}{T}\ln(T) \tag{24}$$

where $\hat{\sigma}^2$ is the RSS divided by the sample size, T, and k is the number of parameters estimated, that is, the number of lags plus a possible constant term. Another common way of determining the model order is to examine the *autocorrelation function* (ACF) of the time series used for ARMA modeling (Enders, 2015, p. 66). This involves interpreting the model order from the patterns displayed in the ACF plot and is, therefore, a more subjective way of determining the order of the model. A combination of the two methods will be used when identifying the optimal mean models in this thesis. Accordingly, values of AIC and BIC for various ARMA(p, q) models as well as the ACF for all three univariate series are presented in Table 5 and Figure 13.

***Table 5** – AIC and BIC values for different ARMA(p, q) models*

	Spot		**Monthly futures**		**Quarterly futures**	
ARMA(p, q)	**AIC**	**BIC**	**AIC**	**BIC**	**AIC**	**BIC**
(0,0)	-293.51	-293.49	-1576.81	**-1567.85**	-1878.59	**-1869.63**
(1,0)	-332.18	-318.74	-1575.33	-1561.89	-1878.55	-1865.11
(0,1)	-342.59	-329.15	-1575.29	-1561.85	-1878.35	-1864.91
(1,1)	-353.21	**-335.29**	-1573.96	-1556.04	-1880.28	-1862.36
(2,0)	-339.98	-322.05	-1574.52	-1556.60	-1878.55	-1860.63
(2,1)	-353.38	-330.98	-1575.26	-1552.86	-1881.55	-1859.15
(1,2)	-357.29	-334.89	-1575.23	-1552.83	**-1881.67**	-1859.27
(2,2)	**-358.93**	-332.05	**-1581.80**	-1554.92	-1880.14	-1853.26

Figures in bold are denoting the minimum of the reported values.

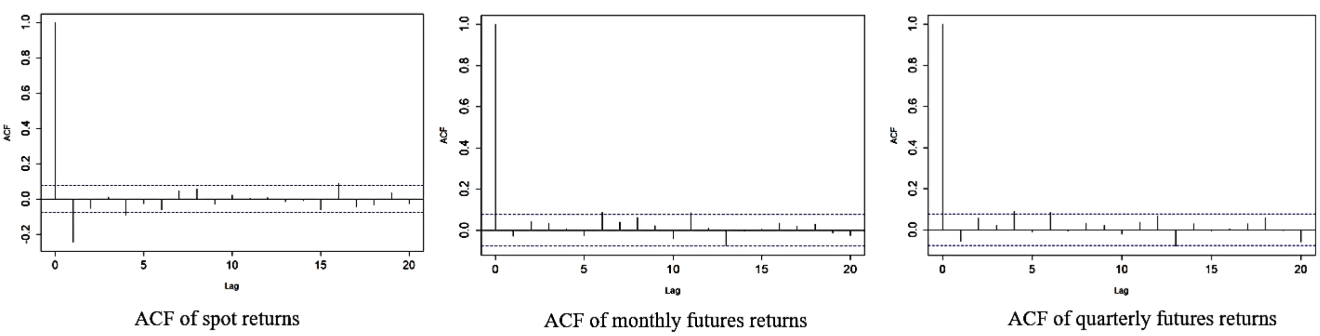


***Figure 13** – ACF plots of spot, monthly futures, and quarterly futures return series.*

As Enders (2015, p. 76) points out, a central idea in the Box-Jenkins methodology is the principle of *parsimony*, as Box and Jenkins argue that parsimonious models produce better estimates than more complex models with too many parameters. Consequently, models of a higher order than an ARMA(2,2) were not considered.

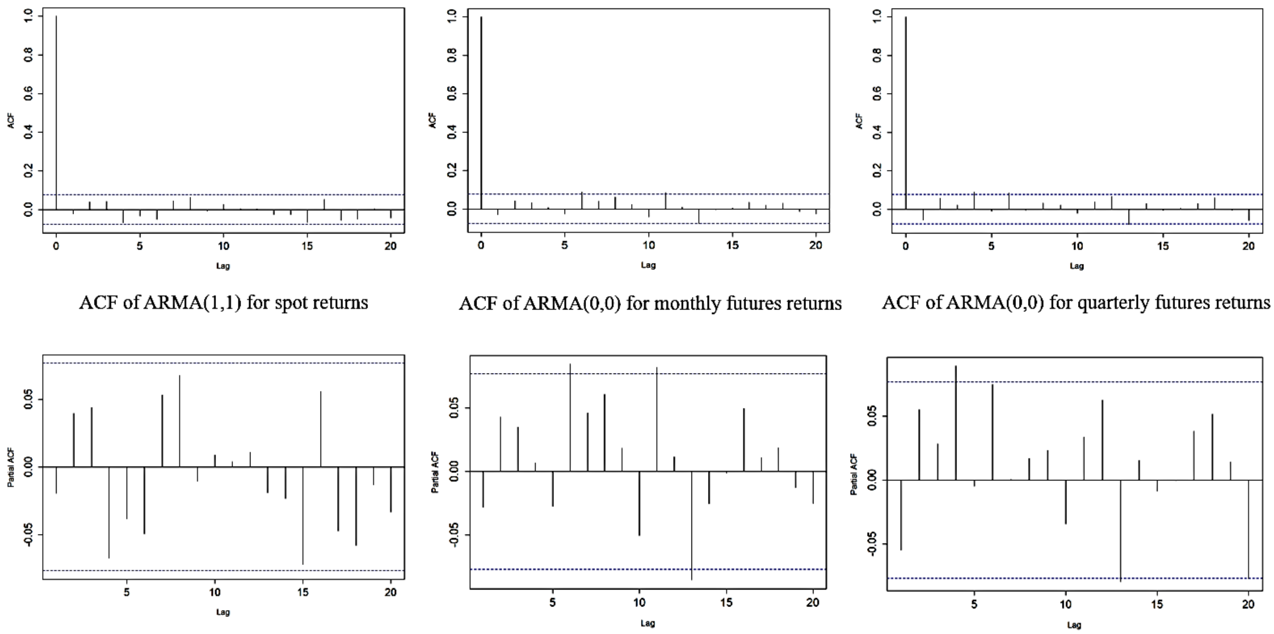
For the spot series, the suggested models according to AIC and BIC are ARMA(2,2) and ARMA(1,1), respectively. Regarding the ACF plots, the blue lines are 95% confidence bands drawn under the null hypothesis of no autocorrelation. From the ACF for the spot return series, it can be inferred that the autocorrelation coefficient on the first lag is highly significant with a negative sign, while the following lags are all insignificant and displaying an oscillating decay. Such a property of the ACF suggests an

ARMA(1,1) process (Enders, 2015, p. 66), which confirms the model suggested by the BIC. Besides, Medel & Salgado (2012) found that BIC produces better estimates and forecasts than AIC in large samples (>50 observations). For these two reasons, an ARMA(1,1) model is chosen for the spot series.

As for the monthly and quarterly futures return series, both AIC and BIC propose an ARMA(0,0) model. This is consistent with the ACFs of the series, as there are no significant autocorrelations displayed, indicating a white noise process (Enders, 2015, p. 66). A white noise process can be defined as a sequence in which "each value in the sequence has a mean of zero, a constant variance, and is uncorrelated with all other realizations" (Enders, 2015, p. 49). Accordingly, an ARMA(0,0) model for both futures return series is selected.

The second step in the Box-Jenkins method is estimation, in which each of the three chosen models is fit under the goal of selecting a stationary and parsimonious model with a good fit (Enders, 2015, p. 76). The estimated coefficients will be reported in sub-section 7.1.2.

The third step in the Box-Jenkins procedure is diagnostic checking, which is to ensure that the estimated models are adequate for the data (Enders, 2015, p. 76). The most commonly used method of diagnostic checking is to perform autocorrelation tests (Brooks, 2008, p. 231). Autocorrelation tests imply examining the autocorrelation function (ACF) and the partial autocorrelation function (PACF) of the residuals of the estimated model to ensure that they mimic a white noise process (Enders, 2015, p. 76). If the ACF and PACF of the residuals display white noise, it intuitively means that the estimated model has extracted all meaningful information from the data and that the model has a good fit (Fabozzi, Focardi, & Kolm, 2010, p. 4).



ACF of ARMA(1,1) for spot returns | ACF of ARMA(0,0) for monthly futures returns | ACF of ARMA(0,0) for quarterly futures returns

PACF of ARMA(1,1) for spot returns | PACF of ARMA(0,0) for monthly futures returns | PACF of ARMA(0,0) for quarterly futures returns

Figure 14 *– ACFs and PACFs of the selected models*

As can be seen from Figure 14, almost all the correlations are within the 95% confidence bands, implying insignificant autocorrelations and partial autocorrelations (Brooks, 2008, p. 234). There are a few exceptions where marginal spikes are visible outside of the 95% confidence bands. However, the confidence bands are built with a 95% significance level, meaning that they will reject the null hypothesis of no autocorrelation when it is true with a 5% probability (Enders, 2015, p. 72). Consequently, the marginal spikes can be regarded as random errors. It can thus be concluded that the residuals are mimicking white noise processes and that the selected models are suitable for the data. An overview of the selected mean models is displayed in Table 6.

Table 6 *- Selected mean models according to the Box-Jenkins methodology*

Series	**Model**	**Equation**
Spot	ARMA(1,1)	$r_t = \mu + \phi r_{t-1} + \gamma \varepsilon_{t-1} + \varepsilon_t$
Monthly futures	ARMA(0,0)	$r_t = \mu + \varepsilon_t$
Quarterly futures	ARMA(0,0)	$r_t = \mu + \varepsilon_t$

6.2.3. Testing for ARCH Effects

According to Brooks (2008), before estimating a GARCH model, it is expedient to first conduct Engle's (1982) test for ARCH effects to ensure that such a model type is appropriate for the data. Since ARCH-type models imply ARMA models for the squared residuals, ε_t^2, Engle (1982) showed that a Lagrange Multiplier (LM) test for ARCH effects in the data could be conducted with the following auxiliary regression (Brooks, 2008, p. 390):

$$\varepsilon_{i,t}^2 = \gamma_0 + \gamma_1 \varepsilon_{i,t-1}^2 + \cdots + \gamma_p \varepsilon_{i,t-p}^2 + \nu_t \tag{25}$$

where the ε_t^2's are the squared residuals with corresponding coefficients γ_i, and ν_t is an error term. The squared residuals are extracted from the conditional mean models, which are the selected ARMA models described in the previous sub-section.

The null hypothesis of the test is that there are no ARCH effects, and the null is rejected if the computed test statistic exceeds its corresponding critical value (Brooks, 2008, p. 389). The LM test statistic is the product of the R^2 of the auxiliary regression and the sample size, and it is chi-squared distributed with p degrees of freedom:

$$LM = nR^2 \sim \chi^2(p) \tag{26}$$

Engle's LM test has been conducted for all three univariate time series to make sure GARCH modeling for the hedge ratio computations is appropriate. Regarding the choice of lags for the test, there is no one right answer, and this is essentially a trade-off between keeping degrees of freedom moderate while not losing sensitivity to realistic alternative hypotheses (Brooks, 2008, p. 58). For that reason, the test is conducted for three different lag lengths. Table 7 reports the computed test statistics for all the mentioned lag lengths in the auxiliary regression from equation (25).

Table 7 *- Engle's Lagrange Multiplier (LM) test for ARCH effects*

Lags (p)	**Spot**	**Monthly futures**	**Quarterly futures**
3	100.41***	47.07***	23.26***
5	102.98***	48.65***	27.05***
10	105.25***	58.81***	28.08***

*The chi-squared critical values the 1% significance level are 11.34, 15.09 and 23.21 for 3, 5, and 10 lags, respectively. *** indicate rejection of the null hypothesis at the 1% significance level.*

Table 7 shows evidence of ARCH effects in all three series given that the null hypothesis is rejected at the 1% level of significance for all chosen lag lengths. The test results therefore indicate that the time series data serve as good candidates for GARCH modeling.

6.2.4. GARCH Model Order Selection

A feature that needs to be decided when specifying a GARCH(p, q) model is the number of lags for p and q. Brooks and Burke (2003) show that modified versions of the AIC and BIC can be used for appropriate model selection when it comes to GARCH models just as with ARMA modeling. AIC and BIC values for different combinations of GARCH(p, q) variance models with mean models corresponding to those selected in sub-section 6.2.2 are therefore computed and reported in Table 8.

Table 8 *- AIC and BIC values for different univariate GARCH(p, q) models*

	Spot		Monthly futures		Quarterly futures	
GARCH(p, q)	**AIC**	**BIC**	**AIC**	**BIC**	**AIC**	**BIC**
(1,1)	**-0.923**	**-0.882**	- 2.474	- 2.447	- 2.949	**- 2.923**
(2,1)	-0.919	-0.871	- 2.473	- 2.440	- 2.947	- 2.913
(1,2)	-0.919	-0.871	**- 2.490**	**- 2.456**	**- 2.952**	- 2.918
(2,2)	-0.917	-0.862	- 2.487	- 2.446	- 2.949	- 2.908

Figures in bold denote the minimum of the reported values.

As Table 8 shows, both BIC and AIC suggest a GARCH(1,1) order for the spot series, while AIC and BIC suggest a GARCH(1,2) and a GARCH(1,1) order for the quarterly futures series, respectively. Accordingly, a GARCH(1,1) model has been selected for both mentioned series. The GARCH(1,1) was selected over the GARCH(2,1) for the quarterly futures series in order to follow the principle of parsimony discussed in sub-section 6.2.2.

When it comes to the monthly futures series, both the AIC and BIC suggest a GARCH(1,2) model. However, when running this model, the first of the two GARCH coefficients (β_1) was estimated to be almost zero, while also being highly insignificant according to the p-value. As this is found not to be a desirable feature to have in the selected model (Brooks, 2008, p. 104), the second best model according to the BIC was used, which was the GARCH(1,1) model. Given that the GARCH(1,1) model produced notably lower standard errors for the coefficients, and since there is only a marginal difference in the BIC values of the models, the GARCH(1,1) was chosen over the GARCH(1,2). A GARCH(1,1) model is also in accordance with the principle of parsimony and existing hedging literature (see, e.g., Kroner & Sultan, 1993; Baillie & Myers, 1991; Hanly et al., 2017).

Diagnostic checking of the selected models confirms that the estimated models are adequately specified. Enders (2015, p. 150) notes that a common way of performing diagnostic checks on GARCH models is to compute the Ljung-Box Q-statistics of the models' squared standardized residuals, η_t^2 (Enders, 2015, p. 150). The standardized residuals, η_t, can be computed by dividing the residuals of the model by the conditional standard deviations, $\eta_t = \varepsilon_t/\sqrt{h_t}$ (Enders, 2015, p. 150). If the null hypothesis that the Q-statistics are equal to zero is rejected, there are no remaining GARCH effects, and the estimated models are properly specified (Enders, 2015, p. 150). The test is specified with 15 lags as this has become a common practice among researchers when employing the Ljung-Box test for GARCH models (Burns, 2002). The test statistic is given by:

$$Q = T(T+2)\sum_{i=1}^{n}\frac{\rho_i^2}{T-i} \tag{27}$$

where T is the sample size, the ρ_i's denote the correlation coefficients and n is the maximum lag length. Furthermore, the test statistic is chi-squared distributed with n degrees of freedom. (Brooks, 2008, p. 209). The test results are reported in Table 9.

Table 9 *– Ljung-Box test for the estimated GARCH models*

	GARCH(p, q) order	**Test statistic**	**p-value**
Spot	(1,1)	5.811	0.983
Monthly futures	(1,1)	18.792	0.223
Quarterly futures	(1,1)	9.666	0.840

*Critical values for 15 lags: 22.31 (10% level), 25.00 (5%) and 30.58 (1%). *, ** and *** indicate rejection of the null hypothesis at 10%, 5% and 1% significance level, respectively.*

As Table 9 shows, the null hypothesis cannot be rejected at any conventional significance level, hence no remaining autocorrelation in the squared standardized residuals are found. Therefore, it can be concluded that the non-linear dependence in the data has been properly modeled by the estimated GARCH(1,1) models. The univariate GARCH(1,1) processes for each return series is formulated as:

$$h_{i,t} = \omega_i + \alpha_i \varepsilon_{i,t-1}^2 + \beta_i h_{i,t-1} \tag{28}$$

where i denotes the type of asset returns, which in this case are the spot, monthly futures, and quarterly futures. The rest of the notation is the same as in equation (20).

6.2.5. Multivariate GARCH Models

The specified GARCH models discussed so far are all univariate models, meaning that they produce estimates of conditional volatility for only one time series (Brooks, 2008, p. 429). However, in order to estimate time-varying hedge ratios, it is also necessary to compute estimates of the conditional covariance between the spot and futures series, as this is an input in the hedge ratio formula (see equation (10)). Consequently, it is necessary to specify multivariate GARCH models, which in addition to estimating conditional variances also estimates conditional covariances, thereby extending the univariate GARCH model (Brooks, 2008, p. 429). The following sub-sections will describe the two chosen multivariate GARCH models for hedge ratio computation in this thesis, which are the CCC-GARCH and the DCC-GARCH model.

6.2.5.1. Constant Conditional Correlation (CCC) GARCH

This sub-section will present the CCC-GARCH model developed by Bollerslev (1990), which is the first of the two multivariate GARCH (MGARCH) models employed in the thesis. The CCC-GARCH model is the simplest type of multivariate GARCH models as it models the conditional correlations as constant, thereby reducing the number of coefficients to be estimated and simplifying the estimation process. It has also been extensively employed in the hedging literature (see, e.g., Byström, 2003; Zanotti et al., 2009; Hanly et al., 2017).

The first step in the model setup is the conditional mean models, which are represented by colon vectors of returns denoted by r_t. In the context of this thesis, the return vectors denote spot and futures returns. As argued in sub-section 6.2.3, r_t is modeled as an ARMA(1,1) process for the spot returns and an ARMA(0,0) for both of the futures returns series:

$$r_t = \mu + \phi r_{t-1} + \gamma \varepsilon_{t-1} + \varepsilon_t \tag{29}$$

$$r_t = \mu + \varepsilon_t \tag{30}$$

Equation (29) applies for the spot return series and equation (30) applies for both futures return series. Furthermore, the residuals for each series are modeled as (Bollerslev, 1990):

$$\varepsilon_t = D_t \eta_t \tag{31}$$

where $D_t = \text{ diag}\left(\sqrt{h_{i,t}}, \dots, \sqrt{h_{N,t}}\right)$ is a diagonal matrix containing the conditional standard deviations of r_t and η_t is an independently and identically distributed colon vector of random errors so that the standardized residuals can be written as:

$$\eta_t = \varepsilon_t D_t^{-1} \tag{32}$$

What is essential in the CCC model is that the conditional covariance matrix of the return vectors, H_t, is defined as the product of the square root of the corresponding conditional variances and the constant conditional correlation (Bollerslev, 1990). The conditional covariance matrix is given by:

$$H_t = \begin{pmatrix} h_{s,t} & h_{sf,t} \\ h_{sf,t} & h_{f,t} \end{pmatrix} = D_t R D_t \tag{33}$$

In equation (33), R denotes the constant correlation matrix and is positive definite with $\rho_{ii} = 1, \forall i$ (Bollerslev, 1990). Furthermore, $h_{s,t}$ and $h_{f,t}$ denote the conditional variance of spot and futures returns, respectively, and $h_{sf,t}$ denotes the conditional covariance of spot and futures returns. The conditional variances can be modeled as any univariate GARCH model (Chang et al., 2011), and by following the argumentation in sub-section 6.2.4, all of the univariate time series in this thesis are modeled as GARCH(1,1) processes (see equation (28)).

Although the CCC-GARCH model is an attractive model because of its simplicity and that it is superior compared to other simple moving average models, many studies have found the assumption of a constant correlation to be too restrictive when applied in empirical research (Bauwens, Laurent, & Rombouts, 2006). It has also been found that the model has difficulties in capturing the interactions among the assets in the model (Zanotti et al., 2009). A natural extension of the CCC-GARCH model that can cope with this limitation is the DCC-GARCH model developed by Engle (2002).

6.2.5.2. Testing for Non-Constant Correlation

As previously mentioned, the assumption of constant correlation is somewhat restrictive and improving the model by introducing dynamic correlations can be beneficial in many cases. Therefore, a test for constant correlation is conducted to see if the use of time-varying correlations can be expected to improve the hedging results in the analysis. A useful econometric test for this purpose is Engle & Sheppard's test for non-constant correlation (Engle & Sheppard, 2001). The null hypothesis is that the correlation between the series is constant:

$$H_0: R_t = \bar{R}, \quad \forall t \in T \tag{34}$$

versus the alternative hypothesis that there exists time dependency in the correlation between the series:

$$H_a: vech^u(R_t) = vech^u(\bar{R}) + \sum_{i=1}^{n} \beta_n vech^u(R_{t-n}) \tag{35}$$

In the above expression, $vech^u$ is a modified type of vectorization that only selects elements above the diagonal (Engle & Sheppard, 2001). The lags of 10 and 15 are chosen to be consistent with prior research (Isogai, 2015; Abubaker, 2016), and to keep degrees of freedom moderate while still being able to test for realistic alternative hypotheses. The test results are shown in Table 10.

Table 10 *- Engle & Sheppard´s test for non-constant correlation*

	10 lags		**15 lags**	
	Statistic	**p-value**	**Statistic**	**p-value**
Monthly	16.92	0.11	30.48***	0.02
Quarterly	14.65	0.20	20.77	0.19

The first row reports the results when testing for non-constant correlation between the spot and monthly futures returns, while the second row applies for the correlation between spot and quarterly futures returns.

As Table 10 shows, the p-values for the test of non-constant correlation between the spot and monthly futures returns are 0.11 and 0.02 for the chosen lag lengths. Consequently, evidence of non-constant correlation is found when running the test with 15 lags, but not with 10 lags. Hence, the hedging model for monthly contracts can show improvements when incorporating a dynamic structure of the conditional correlations when estimating the optimal hedge ratios.

The p-values for the test of non-constant correlation between the spot returns quarterly futures returns are 0.20 and 0.19 for the chosen lag lengths. Consequently, there is not sufficient evidence to reject the null hypothesis of a constant correlation between the series. This suggests that extending the CCC-GARCH model to a model that includes dynamic conditional correlations will lead to no or relatively small improvements in the hedging performance for quarterly contracts.

As the test provides evidence of non-constant correlation between spot and monthly futures returns, and because the assumption of a constant correlation is found to be restrictive, the DCC-GARCH model will be modeled for both futures contracts.

6.2.5.3. Dynamic Conditional Correlation (DCC) GARCH

As mentioned, the most popular extension of the CCC-GARCH model is the DCC-GARCH (dynamic conditional correlation) model developed by Engle (2002). The DCC model copes with the limitation of the CCC model by modeling a time-varying correlation between the return series. The DCC-GARCH model used in the analysis is presented in the following and is estimated in the same manner as it was originally proposed by Engle (2002). The conditional covariance matrix of the returns vectors is now defined as:

$$H_t = \begin{pmatrix} h_{s,t} & h_{sf,t} \\ h_{sf,t} & h_{f,t} \end{pmatrix} = D_t R_t D_t \tag{36}$$

where R_t is the conditional correlation matrix containing the time-varying conditional correlations and the rest of the notation is the same as in equation (33). The DCC-GARCH model is estimated in two stages as the matrices of the conditional standard deviations and the conditional correlation matrix are estimated separately. This two-stage approach of the model maintains much of the simplicity that lies in the CCC model as it has both "the flexibility of univariate GARCH but not the complexity of conventional multivariate GARCH" (Engle, 2002). The mean equations and the univariate GARCH specifications are the same as for the CCC-GARCH model.

Since H_t is a covariance matrix, it is required to be positive definite. This means that the main diagonal of the matrix will contain all positive numbers, and the matrix will be symmetrical about this diagonal (Brooks, 2008, p. 434). Due to the decomposition of H_t in equation (36), the conditional correlation matrix, R_t, also needs to be positive definite. Also, all elements in R_t are required to be less than or equal to one by the definition of the correlation coefficient (Stock & Watson, 2015 p. 78). To ensure that these requirements are satisfied, R_t is defined as (Engle, 2002):

$$R_t = \text{diag}\{Q_t\}^{-1} Q_t \text{diag}\{Q_t\}^{-1} \quad (37)$$

Engle (2002) further defines the positive definite matrix Q_t as:

$$Q_t = \bar{Q}(1 - \theta_1 - \theta_2) + \theta_1(\eta_{t-1}\eta'_{t-1}) + \theta_2 Q_{t-1} \quad (38)$$

where η denotes a vector of the standardized residuals ($\eta_{i,t} = \varepsilon_{i,t}/\sqrt{h_{i,t}}$), Q_t and $\bar{Q}$ are the conditional and unconditional covariance matrices of the standardized residuals, respectively, and θ_1 and θ_2 are non-negative scalars (Engle, 2002; Chang et al., 2011). It is also required that $\theta_1 + \theta_2 < 1$ to ensure that the dynamic correlation process is mean reverting (Engle, 2002). The DCC-GARCH model is non-linear and is estimated by maximizing a likelihood function using a two-stage approach (Engle, 2002). Intuitively, this is done by finding the most likely values of the parameters given the actual data (Brooks, 2008, p. 395). The maximum likelihood estimation[20] is performed using the statistical software R with the package ‘rmgarch’ (Ghalanos, 2019).

Following Zanotti et al. (2009), the time-varying variances and covariances from the estimated covariance matrix H_t are used in order to compute the dynamic hedge ratios as shown in equation (18) for both the CCC-GARCH model and the DCC-GARCH model. The estimated GARCH parameters along with the time-varying hedge ratios will be reported in section 7.

6.3. Out-of-Sample Forecasting

This sub-section will describe the method used for the out-of-sample analysis and how it differs from the in-sample analysis. Even though the in-sample analysis provides a good indication of the futures’ effectiveness as hedging instruments, an out-of-sample analysis produces even more accurate and realistic results (Kroner & Sultan, 1993). As Kroner and Sultan (1993) point out, it is more interesting for hedgers to know how well they can do in the future if they use a different hedging strategy than how

[20] For a full elaboration of the likelihood function and DCC-GARCH models in general, see Engle (2002) and Bauwens et al. (2006).

well they could have performed in the past by using another hedging model. This point is also emphasized by Zhou (2016), as he stresses that hedging decisions have a forward-looking nature in the sense that hedgers must decide for the following period based on currently available information. For these reasons, an out-of-sample analysis is a better way to assess the performance of the proposed hedging strategies than an in-sample analysis, and also more applicable for market participants. In general, empirical evidence from out-of-sample forecasting is considered more trustworthy in comparison to evidence from in-sample performance, and it also presents a more accurate representation of the forecaster in 'real time' (Eurostat, 2015). By withholding the last one-fourth of the total sample, an evaluation period[21] of 163 observations for out-of-sample forecasting is constructed. Consequently, sub-period 4 is used for the out-of-sample analysis of the hedging models.

The static OLS hedge ratio for the out-of-sample analysis is computed based on data from the in-sample period, which is a total of 489 weeks. Thereafter, the performance of the OLS hedge ratio is tested in the evaluation period. It is thus examined whether the optimal hedge ratio estimated in one period can perform well in the successive period. The naïve hedge ratio is also tested in the estimation period for comparison purposes. Since the OLS hedge ratio is now estimated out-of-sample, there is no guarantee that it will outperform the naïve hedge when it comes to variance reduction, which was definitely the case for the in-sample analysis.

One way to forecast out-of-sample is to forecast the hedge ratios for several weeks ahead in time, which is called *multi-step-ahead forecasts* (Brooks, 2008, p. 245). However, multi-step-ahead forecasts produce inaccurate forecasts if the forecast horizon is long (Brooks, 2008, p. 246), which is the case in this thesis as there are 163 weeks in the out-of-sample period. It is therefore recommended to use *one-step-ahead forecasts* with a new estimation period for the model parameters in each period (Brooks, 2008, p. 246). This forecasting scheme has another advantage as it allows the model coefficients to vary in each period, in contrast to the GARCH models in-sample which are linear in parameters (Andersen et al. 2009, p. 171).

[21] Also known as a *holdback period* (Enders, 2015, p. 83)

The out-of-sample analysis for the dynamic hedging models will be performed by conducting one-step-ahead forecasts for the optimal hedge ratio for each week in the evaluation period. As previously mentioned, the CCC-GARCH model assumes a constant correlation, so the forecast of the conditional correlation is equal to the in-sample estimated correlation matrix. However, the diagonal matrix of the conditional standard deviations is time-varying, so one-step-ahead forecasts for the matrix will be performed in each week (Engle & Sheppard, 2001). These forecasts can be expressed as:

$$\widehat{D}_t = E[D_{t+1}|\Omega_t] = \operatorname{diag}\left(\sqrt{\hat{h}_{i,t+1}}, \dots, \sqrt{\hat{h}_{N,t+1}}\right) \tag{39}$$

where Ω_t signifies information given at time t, '^' denotes that it is a forecasted value and the rest of the notation is the same as in sub-section 6.2.5.1 (Bollerslev, 1990). The forecasts for the time-varying conditional covariance matrix for the CCC-GARCH model are expressed as (Engle & Sheppard, 2001):

$$\widehat{H}_{t+1} = E[H_{t+1}|\Omega_t] = \widehat{D}_{t+1} R \widehat{D}_{t+1} \tag{40}$$

The DCC-GARCH model, on the other hand, models a dynamic conditional correlation. Hence, the forecasts for the conditional covariance matrix are expressed as (Engle & Sheppard, 2001):

$$\widehat{H}_{t+1} = E[H_{t+1}|\Omega_t] = \widehat{D}_{t+1} \hat{R}_{t+1} \widehat{D}_{t+1} \tag{41}$$

As mentioned in sub-section 6.2.5.3, the DCC-GARCH model estimates the covariance matrix in two steps, so the forecast in equation (41) implies separate forecasts of D and R. The forecasts of the conditional standard deviations are the same as in the CCC-GARCH model in equation (40), while the forecasts of the conditional correlation matrix can be expressed as:

$$\hat{R}_{t+1} = E[R_{t+1}|\Omega_t] = \operatorname{diag}\{\hat{Q}_{t+1}\}^{-1} \hat{Q}_{t+1} \operatorname{diag}\{\hat{Q}_{t+1}\}^{-1} \tag{42}$$

Since the DCC-GARCH model forecasts each product of the conditional covariance matrix, in contradiction to the CCC-GARCH model, it is expected that the DCC-GARCH model will outperform the CCC-GARCH model out-of-sample. The forecast function in the R package (Ghalanos, 2019) performs the forecasts for the DCC model using the approximation method by Engle and Sheppard (2001).

The estimation period for the one-step-ahead forecasts can follow either a *recursive* or a *rolling window* (Brooks, 2008, p. 246). A recursive window has a fixed start date of the estimation period, but additional observations are added for each forecast (Brooks, 2008, p. 246). Contrarily, a rolling window has a fixed length of the in-sample estimation period for each one-step-ahead forecast (Brooks, 2008, p. 246). This means that the start and end date of the rolling window successively increase by one observation for each forecast (Brooks, 2008, p. 246). The rolling window method is chosen such that the window length of the time-varying hedge ratios corresponds to the estimation period of the OLS estimated hedge ratio. In addition, the choice of a rolling window is in accordance with Hanly et al. (2017). As a result, the rolling window for the one-step-ahead forecasts will have a fixed length of 489 observations, which implies that the first forecast in the evaluation period makes use of an in-sample estimation period spanning from 12/01/2005 to 08/28/2015. The first observation from this period will then be removed and the first observation from the evaluation period will be added, leaving the window length unchanged. This means that for the forecast in the second week of sub-period 4, the in-sample estimation period will span from 12/08/2005 to 09/04/2015. This process will then be repeated, giving a total of 163 one-step-ahead forecasts for the dynamic hedge ratios in sub-period 4. The forecasted hedge ratios will be presented and discussed in section 7.

6.4. Performance Measures

This sub-section will present the performance measures used for assessing and comparing the performance of the hedging models presented in the previous sub-section.

Traditional literature describes hedging effectiveness as the success of reducing the variability of cash prices (Ederington, 1979; Lindahl, 1989), and the effectiveness is further said to be determined by the hedging relationship, that is, the correlation between price changes in the hedged asset and the hedging vehicle used, and the hedge ratio (Charnes & Koch, 2003). Decreased variability is most often the objective for a hedger, but not always (Working, 1953). Therefore, hedging effectiveness and performance can be measured in various ways. The focus of this thesis will be on the percentage reduction in the portfolio’s variance (Ederington, 1979) and value at risk (Hanly et al., 2017) compared to an unhedged portfolio and the maximization of a utility function for the hedger (Anderson & Danthine, 1980). The rationale behind these choices along with a more thorough description of the performance measures will be presented in the following sub-sections.

6.4.1. Portfolio Variance and Ederington's Hedging Effectiveness

The first measure that will be used to assess the performance of the hedging models is the sample variance of the hedged portfolios' returns:

$$var(r_\pi) = var\left(r_s - \beta r_f\right) \tag{43}$$

The notation in the above expression is the same as in equation (7). Both the standard deviation and the variance of a portfolio's return are widely used measures of risk, as they both measure the deviations from the expected return (Berk & DeMarzo, 2014, p. 317). By using the portfolio variance of a hedged portfolio, Ederington (1979) introduced a measure of hedging effectiveness, which is the percentage reduction obtained by a hedged portfolio relative to an unhedged spot portfolio, known as the Ederington's hedging effectiveness measure (EHE):

$$EHE = 1 - \frac{var(r_\pi)}{var(r_s)} \tag{44}$$

In equation (44), $var(r_\pi)$ is the sample variance of the hedged portfolio containing both spot and futures positions, and $var(r_s)$ is the sample variance of the unhedged spot portfolio. This way of measuring hedging effectiveness has been employed extensively in the hedging literature since Ederington's paper from 1979 (see, e.g., Baillie & Myers, 1991; Kroner & Sultan, 1993; Chang et al., 2011). For the OLS model, the EHE will equal the R^2 of the regression. Disadvantages of this measure are that it ignores the transaction costs and that it does not differentiate between downside and upside risk (Bodie et al., 2011, p. 132).

As emphasized by Byström (2003), a large portfolio variance in one period could be very sensitive to a small number of very large returns, even though most of the returns are smaller than the returns of the portfolio they are compared to. For that reason, he proposes an additional measure, which is to examine how often the weekly returns from hedging is smaller than the weekly spot returns in absolute terms, that is, to count the observations where:

$$\left|r_{\pi,t}\right| < \left|r_{s,t}\right| \tag{45}$$

As before, $r_{s,t}$ and $r_{\pi,t}$ denote a weekly return observation of the unhedged spot portfolio and a hedged portfolio, respectively. A portfolio containing a small number of relatively large absolute returns would benefit from this alternative approach. Hence, this approach will also be used when evaluating the performance of the hedging models.

6.4.2. Value at Risk

The various hedging strategies will also be measured by the reduction in Value at Risk (VaR) compared to an unhedged position. This is to obtain a broader view of the performance of the models. Additionally, VaR is the most common way of measuring a portfolio's downside risk and differs in that way from the portfolio variance which does not distinguish between downside and upside risk (Bodie et al., 2011, p. 132). This feature makes it especially useful when examining hedging models, given that one of the primary purposes of hedging is to reduce exposure to unexpected negative events (Working, 1953).

VaR can be defined as the maximum percentage loss of a financial position that can be expected during a given period for a specified level of confidence (Wang, Yeh, & Chuang, 2015). The confidence level for the estimation in the thesis is set to 95% given its widespread use in practice (Bodie et al., 2011, p. 138). As such, the VaR will be interpreted as the maximum percentage loss the hedger can expect to occur over the next week at a 95% confidence level. It is calculated as:

$$VaR_{95\%} = \mu_\pi - 1.645 * \sigma_\pi \tag{46}$$

where μ_π is the mean return of the portfolio, 1.645 is the one-sided critical value from the Student's t-distribution at the 95% level of confidence (Stock & Watson, 2015, p. 805), and σ_π is the standard deviation of the portfolio return.

In order to evaluate the hedged portfolios more intuitively, the percentage reduction in VaR compared to an unhedged spot position is calculated as:

$$VaR_{95\%} \text{ reduction} = 1 - \frac{VaR_{95\%}(r_\pi)}{VaR_{95\%}(r_s)} \tag{47}$$

In the above expression, $VaR_{95\%}(r_{\pi})$ denotes the VaR from a hedged portfolio, and $VaR_{95\%}(r_s)$ denotes the VaR of an unhedged spot portfolio, both at the 95% confidence level. This performance measure has also been adopted in papers investigating hedging strategies (see, e.g., Zhou, 2016; Hanly et al., 2017).

6.4.3. Mean-Variance Utility

Although risk reduction parameters such as the two mentioned are the most common ways of measuring hedging effectiveness in the literature, an obvious drawback of the mentioned methods is that they do not consider the transaction costs of rebalancing the portfolio. Specifically, the proposed hedging models imply a weekly rebalancing for the two GARCH models, whereas the naïve and OLS hedge only require a rebalancing when rolling over to the next futures contract five trading days before maturity. This difference in transaction costs is especially notable when considering the hedges with quarterly futures contracts, as the naïve and OLS hedge only require four interactions in the futures market per year when rolling over the quarterly futures contracts, while the hedged portfolios from the GARCH models requires rebalancing every week throughout the year. This implies that the time-varying hedge models are more costly to implement compared to the time-invariant models (Zhou, 2016).

For transactions costs in general, this does not only concern the monetary costs involved when interacting in the futures market, but also the time and effort put into it (Nasdaq, n.d.-c). In addition to the costs of physically moving the asset from seller to purchaser, and commission and clearing fees, the costs also include the time and labor costs associated with the trade, as well as costs associated with the bid-ask spread (Nasdaq, n.d.-c).

Due to transaction costs from interacting in the futures market, a widely used hedging performance measure is utility maximization (Anderson & Danthine, 1980). A way of comparing utility of different hedging models when accounting for transaction costs and the portfolio returns can be done by applying the framework of Kroner and Sultan (1993). They examined the hedging effectiveness of dynamic and static hedge ratios with the use of foreign currency futures and introduced a way of measuring the models based on a mean-variance utility function. This approach has been employed in several studies in the hedging literature since it was introduced, see Park & Switzer (1995), Moon, Yu, & Hong (2009) and

Zhou (2016) among others. Specifically, Kroner and Sultan (1993) assume that the typical hedger can be represented by the following utility function:

$$U(r_\pi) = E(r_\pi) - y - \lambda \mathrm{var}(r_\pi) \tag{48}$$

where U measures the hedger's utility in each period as a function of the returns from the hedged portfolio, r_π. Furthermore, $E(r_\pi)$ denotes the expected return on the hedged portfolio, y represents the reduced returns caused by the transaction costs of trading futures contracts, and λ is a measure of the hedger's risk aversion (Kroner & Sultan, 1993). A larger λ is associated with a higher level of risk aversion, meaning that the hedger's utility becomes more negative for a given level of variance of the portfolio return (Kroner & Sultan, 1993). A hedger represented by this utility function solves the following problem when deciding on the hedge ratio:

$$\max_{\beta} E[U(r_\pi)] \tag{49}$$

It can be shown that when solving this maximization problem, the hedger's optimal hedge ratio can be expressed as[22]:

$$\beta^* = \frac{2\lambda cov(r_s, r_f) - E(r_f)}{2\lambda var(r_f)} \tag{50}$$

The optimal hedge ratio for a hedger with this utility function corresponds to the minimum-variance hedge ratio shown in equation (10) if the *martingale assumption* for the futures returns hold, that is, if $E(r_f) = 0$ (Kroner & Sultan, 1993). The martingale assumption implies that the expected return on futures is zero, meaning the number of futures contracts held does not affect the expected return of the hedged portfolio (Kroner & Sultan, 1993). Several papers in the hedging literature continue their analyses as if this assumption holds (see, e.g., Kroner & Sultan, 1993; Park & Switzer, 1995). The assumption that futures prices are martingales helps to simplify the analysis without affecting the accuracy of the results substantially (Ankirchner, Pigorsch, & Schweizer, 2013). Recall that the selected conditional mean equations for the futures returns are modeled as white noise processes (see Table 6), which essentially means that the best estimate of the futures price tomorrow is today's price (Enders, 2015, p. 129). Therefore, it makes sense to assume that the futures prices behave as martingales in this case.

[22] For a full mathematical derivation of equation (50), see Kroner & Sultan (1993).

Consequently, the optimal hedge ratio for a hedger with the utility function in equation (48) corresponds to the minimum-variance hedge ratio in equation (10).

Zhou (2016) emphasizes the advantages of applying the mentioned utility function when examining the effectiveness of hedging strategies. The primary objective of hedging is to minimize the risk of a portfolio, but that does not mean that the hedger is indifferent with regard to the portfolio return or the impact of transaction costs (Zhou, 2016). Therefore, representing the hedger by the proposed utility function can bring important and valuable insights as to how much economic benefit the different strategies generate (Zhou, 2016).

By applying the utility function in equation (48), the average weekly utility of each of the different models can be compared. In the computation of the average weekly utility, the expected return equals the average weekly portfolio return (Zhou, 2016). Following Zhou (2016), the utility gain of the different hedging strategies relative to the unhedged strategy will also be reported, which is computed as:

$$\text{Utility gain} = U(r_\pi) - U(r_s) \tag{51}$$

As before, r_π denotes the return of any of the hedged portfolios and r_s denotes the return of an unhedged spot portfolio (Zhou, 2016). This represents an intuitive way of examining which hedging model that brings the most utility gain compared to a no-hedge strategy.

In addition to the mentioned approach, Kroner and Sultan (1993) introduce a variant of this model that has the advantage that the hedger only rebalances the hedged portfolio if the expected utility of rebalancing is greater than the expected utility of not rebalancing. The expected utility is based on the forecasts of the conditional variances and covariances as described in sub-section 6.3. It can be shown that the expected utility of the portfolio in the following period, if it is rebalanced, is[23]:

$$\begin{aligned} E_t\left[U\left(r_{\pi,t+1}^{\text{rebalance}}\right)\right] &= E_t\left(r_{\pi,t+1}\right) - y - \lambda \text{var}\left(r_{\pi,t+1}\right) \\ &= -y - \lambda\left(\hat{h}_{s,t+1} - 2\beta_t^* \hat{h}_{sf,t+1} + \beta_t^{*2} \hat{h}_{f,t+1}\right) \end{aligned} \tag{52}$$

[23] For a full derivation, see Kroner & Sultan (1993).

In equation (52), $\hat{h}_{s,t+1}$ and $\hat{h}_{f,t+1}$ denote forecasts of conditional variances of the spot and futures returns in the following week, respectively and $\hat{h}_{sf,t+1}$ denotes forecasts of conditional covariance between the spot and futures returns in the next week. In accordance with Kroner and Sultan (1993), the expected return in the following week is assumed to be zero for simplicity. Similarly, it can be shown that the expected utility by not rebalancing the hedged portfolio is:

$$\begin{aligned} E_t\left[U\left(r_{\pi,t+1}^{\text{no rebalance}}\right)\right] &= E_t\left(r_{\pi,t+1}\right) - \lambda \text{var}\left(r_{\pi,t+1}\right) \\ &= -\lambda\left(\hat{h}_{s,t+1} - 2\beta_{t'}^{*}\hat{h}_{sf,t+1} + {\beta_{t'}^{*}}^{2}\hat{h}_{f,t+1}\right) \end{aligned} \tag{53}$$

The only two differences between the expected utility of rebalancing in equation (52) and the expected utility of no rebalancing in equation (53) is that the expected utility of no rebalancing does not account for the transaction costs, y, and the hedge ratio for the portfolio is denoted by $\beta_{t'}^{*}$, indicating that it is the hedge ratio from the most recent rebalancing (Kroner & Sultan, 1993). In contrast, the hedge ratio set if the portfolio is rebalanced is denoted by β_t^{*}, indicating that it is the most optimal hedge ratio to implement at time t according to the newest information about the conditional variances and covariances (Kroner & Sultan, 1993). The hedger therefore decides every week whether rebalancing of the hedged portfolio is worthwhile, that is, the hedger rebalances if:

$$E_t\left[U\left(r_{\pi,t+1}^{\text{rebalance}}\right)\right] > E_t\left[U\left(r_{\pi,t+1}^{\text{no rebalance}}\right)\right] \tag{54}$$

For this method, the conditional forecasts from the DCC-GARCH model is used when measuring the expected utility, and consequently, only the DCC-GARCH model is considered. This variant of Kroner and Sultan's (1993) utility framework is applied in the out-of-sample period as the method requires forecasts of the conditional variances and covariances. To rank the models according to this method, the accumulated utility from each week in the out-of-sample period will be computed for each model. The results will be reported in sub-section 7.2.2 when presenting the empirical results from the out-of-sample analysis.

6.5. Bootstrapping

To compare the hedging performance of the models in a statistical sense, a bootstrapping technique will be used. This is essential as it helps to ensure that the outperformance of the models is indeed significant and not due to chance. Bootstrapping is a common way of resampling data, and although different versions exist, the original method proposed by Efron (1979) will be applied. This is among the simplest methods of resampling data, but it is frequently used in the hedging literature to obtain statistical inference (see, e.g., Kavussanos & Nomikos, 2000; Byström, 2003; Poominars, Cadle, & Theobald, 2003).

The bootstrap method by Efron (1979) works by systematically drawing one random observation (with replacement) from the obtained return series until a new return series with the same length as the original series is constructed. This process will be repeated 1000 times for each hedging model in each sub-period such that a total of 1000 return series for the hedged portfolios are obtained. Based on the resampled series, distributions of new variance and VaR estimates for the hedged portfolios will be obtained, thus facilitating t-tests for differences in hedging performance of the models.

The tests that will be conducted are two sample t-tests for differences in means (Stock & Watson, 2015, p. 128), and a total of four different tests will be carried out. All tests will be based on differences in the mean of the variance and VaR distributions for the models. The tests will be constructed for: the unhedged model compared to the hedging models, the best static model compared to the best dynamic model, the CCC-GARCH model compared to the DCC-GARCH model, and the monthly contracts compared to the quarterly contracts. The tests will be conducted for each sub-period (see Appendix 1), and the results will be discussed when comparing the hedging performance in section 7.

Section 6 has now presented the hedging models that will be used in the analysis, which are the naïve hedge, the OLS model, the CCC-GARCH model and the DCC-GARCH model. Econometric modeling of the dynamic hedging models was done according to the Box-Jenkins methodology. This includes the conditional mean and variance models for each model, and the setup for the conditional covariance matrix in each of the models. In addition to model selection, section 6 has presented reasons for performing an out-of-sample analysis as well as explaining the forecasting method that will be applied. After that, the performance measures that will be used when assessing the hedging results of the models were presented. These measures include portfolio variance, Value at Risk and the utility maximization corresponding to a mean-variance utility function. One alternative measure for each of these is also described and will be used to assess the performance of the hedging models when compared to an unhedged strategy. Lastly, the bootstrapping technique and the statistical tests conducted to obtain statistical inference between the performances of the models have been explained. Section 7 will present the empirical results of the thesis.

7. Empirical Results

Section 7 will present the results from the empirical analysis of hedging in the Nordic power market based on the discussions in the previous sections. The goal is to obtain empirical evidence for answering the research questions and corresponding sub-questions presented in section 1. The section is divided into two sub-sections where the first covers the in-sample analysis and the second covers the out-of-sample analysis. The in-sample analysis refers to the analysis of the estimated hedge ratios that are generated from the same sample that was used to estimate the model's parameters (Brooks, 2008, p. 245). In contradiction, the out-of-sample analysis makes use of forecasted hedge ratios (Enders, 2015, p. 82).

7.1. In-Sample Analysis

The following sub-sections will present the estimated hedge ratios from the in-sample analysis along with the hedging performance of the models for the same period.

7.1.1. Estimated Static Hedge Ratios

As previously mentioned, a static OLS hedge model is a model that minimizes the variance of a portfolio (Hull, 2012, p. 57). The hedge ratio of the static OLS model in each period is estimated according to the regression model presented in equation (11). Table 11 presents the static OLS estimated hedge ratio for monthly contracts in each sub-period, along with the estimated hedge ratio for the full in-sample estimation period.

Table 11 *– OLS estimated hedge ratios with monthly futures*

	Sub-period 1	Sub-period 2	Sub-period 3	Full in-sample period
Hedge ratio (β^*)	0.626***	1.559***	0.886***	0.960***
	(0.000)	(0.001)	(0.000)	(0.000)
Constant	0.007	0.002	0.005	0.005
	(0.456)	(0.851)	(0.682)	(0.237)
Observations	163	163	163	489
R^2	0.159	0.176	0.233	0.119

*Parentheses represent p-values. *, **, and *** indicate significance at the 10%, 5% and 1% level, respectively.*

Table 11 indicates a great deal of variation for the estimated hedge ratios across periods. This relates to the discussions in sub-section 6.2 on time-varying volatility in electricity markets. In three of the four periods, the optimal hedge is to enter a short position in futures that is smaller than the current spot position. Sub-period 2 stands out as the only period in which the hedger applies a hedge ratio above 1, in contrast to the other sub-periods. For a hedger in sub-period 2, this implies taking a short position in futures contracts that is relatively larger than the size of the hedger's long position in the spot market. From equation (10) displaying the optimal hedge ratio, it can be seen that the results in sub-period 2 are largely driven by the relatively high standard deviation in the spot market during this sample period, which is evident in Table 2.

The estimated slope coefficients, which represents the hedge ratios of the portfolios, are statistically significant at the 1% level. Intuitively, this means that the minimum variance of the hedged portfolio is obtained by taking a position in futures contracts. The intercept (constant) is close to zero for all periods and statistically insignificant at all conventional significance levels. As previously discussed, keeping the intercept ensures that the first OLS assumption is not violated, and for that reason, the intercept is kept in the model. The R^2's range from 0.159 to 0.233 for the sub-periods, and the R^2 is 0.119 for the full in-sample period. As previously described for the OLS model, the R^2 is a measure of the variance reduction from using the model compared to an unhedged portfolio. The hedging performance of the models in-sample will be further discussed in sub-section 7.1.4.

Table 12 presents the static OLS estimated hedge ratios for quarterly contracts for all sub-periods and the full in-sample period. Again, the hedge ratio in each period is estimated according to the regression model in equation (11).

***Table 12** – OLS estimated hedge ratios with quarterly futures*

	Sub-period 1	**Sub-period 2**	**Sub-period 3**	**Full in-sample period**
Hedge ratio (β^*)	0.580***	1.621***	0.882***	0.943***
	(0.000)	(0.003)	(0.008)	(0.000)
Constant	0.005	-0.003	0.003	0.003
	(0.617)	(0.773)	(0.817)	(0.570)
Observations	163	163	163	489
R^2	0.107	0.158	0.047	0.091

*Parentheses represent p-values. *, **, and *** indicate significance at the 10%, 5% and 1% level, respectively.*

The results for monthly and quarterly contracts for the same periods can be compared, but comparisons across periods do not make sense. The reason is that the dependent variable, in this case the spot returns, needs to be the same for the hedges being compared (Lindahl, 1989). The results in Table 12 are very similar to the results for monthly contracts in Table 11. The optimal hedge ratio is close to the naïve hedge ratio over the full sample and is above 1 in sub-period 2, which is also the case for the monthly contracts. In addition, the hedge ratios for both contract types are close to 0.6 in the sub-period 1 and close to 0.9 in the sub-period 3.

Like the results from monthly contracts, the hedge ratios are statistically significant at the 1% level. What is also evident from Table 12 is that the intercept is close to zero in all the regressions, and none of them is found significantly different from zero based on the p-values. Just like for the monthly contracts, the intercept is kept to ensure that the first OLS assumption is satisfied. Another takeaway is that the R^2's are relatively lower for the quarterly contracts than for the monthly contracts. As mentioned, this will be further discussed in sub-section 7.1.4, where the hedging results of the models will be compared.

7.1.2. Estimated GARCH Parameters

This sub-section will present the estimated parameters from the multivariate GARCH models. Table 13 below summarizes the results from the estimation of both the CCC-GARCH model and the DCC-

GARCH model. The estimated coefficients are only applicable for the in-sample analysis as the coefficients in the out-of-sample analysis vary in each period following the forecasting method described in sub-section 6.3.

As can be seen from Table 13, the AR coefficient for the ARMA(1,1) model for spot returns, ϕ, is estimated to almost 0.16. The AR coefficient represents the positive effect of the return in the previous week on this week's return (Brooks, 2008, p. 215). The moving-average (MA) coefficient, γ, for the same model is estimated to almost -0.37. The MA term is a measure of the effect from a white noise disturbance term in the previous week (Brooks, 2008, p. 211). Both futures series are modeled as white noise processes around a constant, as discussed in sub-section 6.2.4, and AR or MA terms are therefore not included in the conditional mean models.

The estimated coefficients in the univariate GARCH models show notable differences between the return series. For instance, the parameter α measuring the ARCH effects in the conditional variance models shows distinct differences across the series. The parameter is notably higher for the spot series, which is estimated to 0.72, than for both futures series, which are estimated to approximately 0.11 and 0.14 for the monthly and quarterly series, respectively. Intuitively, this implies that volatility shocks in the spot returns have a larger effect on the volatility in the next period than for a similar shock in the futures returns.

Table 13 - Estimated GARCH parameters

	Conditional mean		
Parameter	Spot	Monthly futures	Quarter futures
μ	0.005454	-0.001674	-0.001020
	(0.27277)	(0.517869)	(0.599223)
ϕ	0.158974	-	-
	(0.64146)		
γ	-0.365402	-	-
	(0.27410)		
	Univariate GARCH		
Parameter	Spot	Monthly futures	Quarterly futures
ω	0.006782	0.000647	0.000306
	(0.00018)	(0.006062)	(0.017645)
α	0.721215	0.109323	0.136473
	(0.00622)	(0.001058)	(0.002580)
β	0.275485	0.768355	0.776642
	(0.00791)	(0.00000)	(0.00000)
$\alpha + \beta$	0.996700	0.877678	0.913115
	Multivariate GARCH		
Parameter	Spot and monthly futures		Spot and quarterly futures
$\bar{\rho}$	0.401088		0.321626
θ_1	0.004839		0.003491
	(0.206228)		(0.384253)
θ_2	0.990020		0.991797
	(0.00000)		(0.00000)
Log-likelihood CCC	1174.862		1307.978
Log-likelihood DCC	1175.485		1308.708

The parentheses denote p-values based on the standard errors from the maximum likelihood estimation of the parameters. $\bar{\rho}$ represents the estimated constant conditional correlation between both the spot and monthly futures returns and the spot and quarterly futures returns. θ_1 and θ_2 denote the DCC coefficients as shown in equation (38). The values for the log-likelihood are the maximized values of the likelihood functions. The models are estimated using the statistical software R with the package 'rmgarch' (Ghalanos, 2019). For the R code used, see Appendix 2.

The estimated coefficients for the GARCH term, β, display the opposite of the ARCH effect parameters. For the monthly and quarterly futures models, the parameters are estimated to be approximately 0.77 and

0.78, respectively, while it is estimated to only 0.28 for the spot model. This implies that shocks to the conditional volatility are more persistent in the futures market than in the spot market, as the shocks taper off more slowly. This is visible in Appendix 3.1, displaying the conditional volatility for all three univariate GARCH(1,1) models. This result is also intuitive when considering the previous discussions on the spot price characteristics in sub-section 2.3.3, as large temporary deviations between the spot and futures prices tend to occur due to the complex nature of electricity markets.

All the nine estimated GARCH parameters are statistically significant at the 1% level, suggesting that volatility clusters are present in both the spot and futures market for power. This confirms the results from the LM test for ARCH effects in sub-section 6.2.3. It can also be seen that the sum of the ARCH and GARCH effects, $\alpha + \beta$, is less than one for all series, meaning that all the models fulfill the criterion of stationarity in the variance process (Bollerslev, 1986). This sum is almost one for the spot model, while it is 0.88 and 0.91 for the monthly and quarterly futures models, respectively. Given that the results are close to one, this indicates that the volatility processes are slowly reverting to the mean.

The estimated constant conditional correlations with respect to the spot returns, $\bar{\rho}$, are 0.40 and 0.32 for monthly and quarterly futures returns, respectively. The correlations are found to be lower than what was presented in Table 2 because the model only considers the in-sample period. The quarterly futures returns are somewhat less correlated to the spot returns than the monthly futures returns. The economic intuition of this is that the quarterly futures apply for a longer delivery period than monthly futures, and the effects of supply and demand shocks could, therefore, become more smoothened in the quarterly futures prices compared to the monthly futures prices. An implication of this could be that hedgers with quarterly commitments achieve lower hedging performance compared to those with monthly commitments. The CCC-GARCH model makes use of these correlations, along with the time-varying conditional variances in estimating the hedge ratio, both in-sample and out-of-sample.

Regarding the DCC coefficients, θ_1 and θ_2, Table 13 shows that θ_1 is almost zero for both the model with spot and monthly futures returns and the model with spot and quarterly futures returns. Also, the estimates are not statistically significantly different from zero. The second DCC coefficient, θ_2, is estimated to 0.99 for both models and it is found to be statistically significantly different from zero at the

1% level. Recalling equation (38), this indicates that the dynamic conditional correlations should not deviate substantially from the constant conditional correlation from the CCC model, given that θ_1 is almost zero and θ_2 is close to one, thus causing a high persistence of lagged values in the conditional correlation (see Appendix 3.3). Nevertheless, the sum of the DCC coefficients is still less than one, so the model satisfies the criterion of mean reversion in the conditional correlation process (Engle, 2002). This implies that there is little time dependency modeled in the conditional correlations in the DCC-GARCH model. Recalling Engle and Sheppard's (2001) test for non-constant correlation (see Table 10), there was found evidence of a non-constant correlation for spot and monthly futures returns. Consequently, the hedging performance of both the CCC-GARCH and the DCC-GARCH model will be examined.

7.1.3. Estimated Dynamic Hedge Ratios

In this sub-section, the in-sample estimated hedge ratios from the time-varying conditional variances and covariances between the spot and futures returns will be presented. The hedge ratios are based on the CCC-GARCH and DCC-GARCH models presented in sub-section 6.2. Figure 15 shows the estimated hedge ratios obtained from the application of the dynamic models for monthly futures.

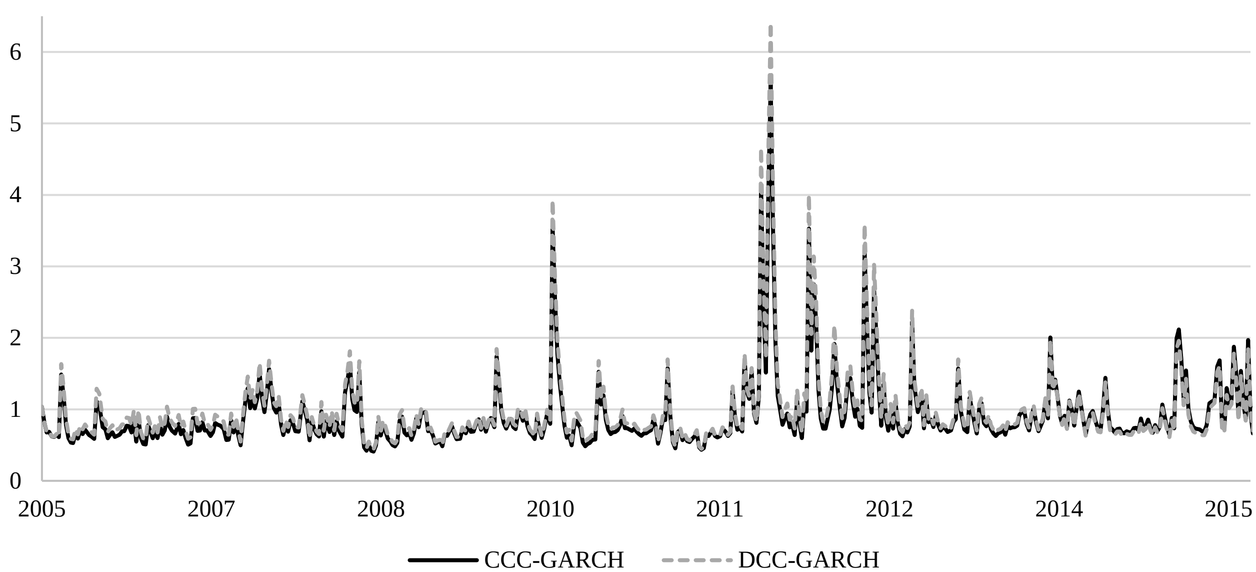


***Figure 15** – Estimated dynamic hedge ratios - monthly futures – in-sample*

The hedge ratios from the dynamic models follow each other relatively closely, and differences between them are minimal. This is not surprising, recalling the discussion on the estimated DCC parameters in sub-section 7.1.2. What can be mentioned is that the DCC-GARCH hedge ratios are generally slightly higher than the CCC-GARCH hedge ratios. This implies that the estimated dynamic conditional correlation from the DCC-GARCH model is often higher than the constant conditional correlation applied to the CCC-GARCH model (see Appendix 3.3). As previously mentioned, the difference between the CCC-GARCH model and the DCC-GARCH model is expected to be more visible out-of-sample.

Figure 16 shows the estimated hedge ratios obtained from the application of the dynamic models with quarterly futures.

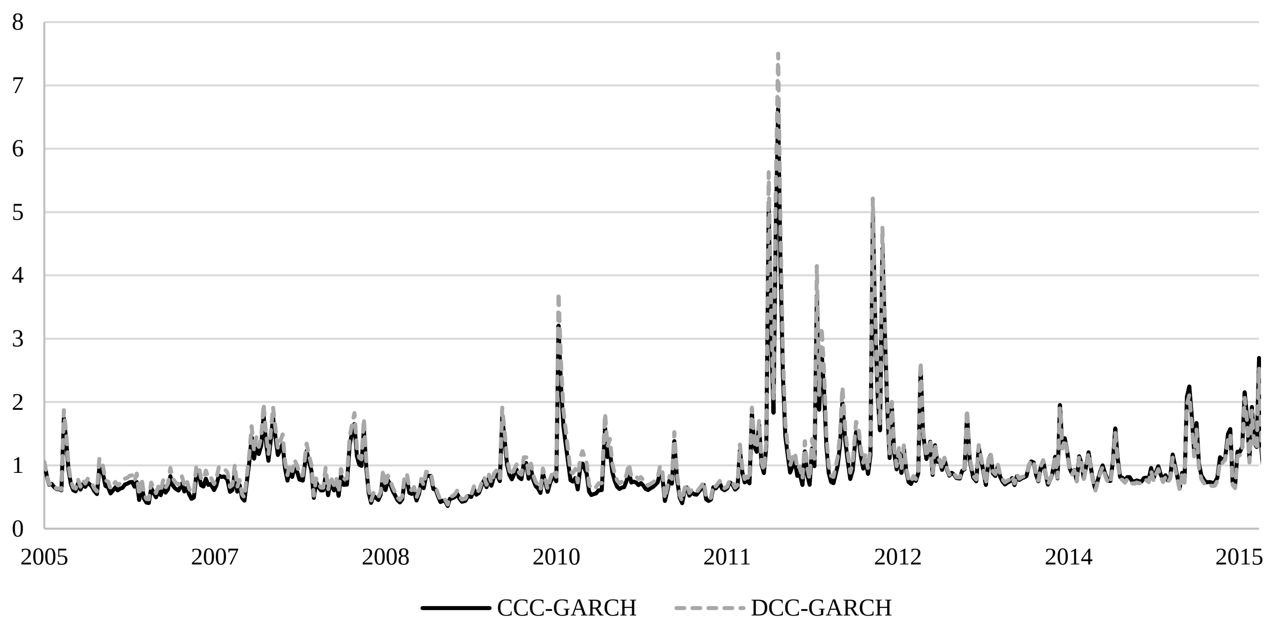


Figure 16 *– Estimated dynamic hedge ratios - quarterly futures – in-sample*

Figure 16, showing the hedge ratios for the quarterly futures contracts, is very similar to the hedge ratios for monthly futures in Figure 15, as the CCC-GARCH and DCC-GARCH hedge ratios seem to be highly correlated (see Appendix 3.3). To further examine the characteristics of the models for both monthly and quarterly contracts, and to be able to make comparisons, descriptive statistics for the time series of the dynamic hedge ratios are presented in Table 14.

Table 14 *- Descriptive statistics for in-sample estimated dynamic hedge ratios*

	Monthly futures		**Quarterly futures**	
	CCC-GARCH	DCC-GARCH	CCC-GARCH	DCC-GARCH
Min.	0.409	0.442	0.359	0.382
Max.	5.566	6.405	6.633	7.502
Mean	0.916	0.979	0.969	1.036
No. of times higher than 1	111 / 489	135 / 489	137 / 489	160 / 489
No. of times higher than β^*_{OLS}	129 / 489	149 / 489	154 / 489	185 / 489
ADF	-9.423***	-9.385***	-8.607***	-8.676***
Q^2	410.80***	433.93***	594.32***	591.71***

Min and max are the lowest and highest hedge ratios observed in the time series, respectively. Mean is the average of the dynamic hedge ratios. No. of times above 1 represents all observations that lie above the naïve hedge ratio. β^*_{OLS} *represents the OLS estimated hedge ratio of 0.960 for the monthly contracts and 0.943 for the quarterly contracts. ADF denotes the test statistic for the ADF test[24] for stationarity, with lags selected according to the AIC. The critical value for the ADF test at the 1% (***) significance level is -3.43.* Q^2 *is the Ljung-Box test statistic[25] measuring autocorrelation with 15 lags. The critical value for the Ljung-Box test at the 1% (***) significance level with 15 degrees of freedom is 30.58.*

Overall, the estimated hedge ratios for both contract types display the same characteristics. One difference, however, is that the highest hedge ratio for the DCC-GARCH model is notably higher than the highest hedge ratio from the CCC-GARCH model. The difference between the maximum hedge ratio from CCC-GARCH and DCC-GARCH is 0.839 and 0.869 for the monthly and quarterly contracts, respectively. Also, the average hedge ratio from the DCC-GARCH model is higher than for the CCC-GARCH model, and the DCC-GARCH hedge ratios are also higher than the naïve and OLS hedge ratio on more occasions than the CCC-GARCH model. Recalling the margin requirements set by Nasdaq, this implies that an actor applying the DCC-GARCH model will have more funds bound in the margin account. Compared to the static models, the funds required for the application of the GARCH models are, in some cases, multiple times the amount of the funds required for a static model. As previously mentioned, the only difference between the models is that the CCC-GARCH model assumes a constant correlation while the DCC-GARCH model incorporates a dynamic structure in the correlation. Consequently, the differences in the hedge ratios are solely due to differences in correlation.

[24] For more details on the ADF test, see sub-section 5.5.

[25] For more details on the Ljung-Box test, see sub-section 6.2.4.

By comparing the hedge ratios for the monthly contracts to the hedge ratios for the quarterly contracts, one takeaway is that the hedge ratio series for the quarterly contracts contain lower minimum and higher maximum hedge ratios than those found for the monthly contracts. The mean for the quarterly contracts is also higher than the mean for the monthly contracts. This implies that a hedger following a dynamic model will need to hedge a higher percentage of the portfolio on average if quarterly contracts are used, compared to monthly contracts.

As the null hypothesis of the ADF test and the Ljung-Box test is rejected at the 1% level, all the hedge ratio series are found to be stationary and to exhibit positive autocorrelation. The practical implication of positive autocorrelation in the series is that if the hedger implements a large hedge ratio one week, then the hedger would expect it to remain large also in the following week unless the market is hit by a shock (Kroner & Sultan, 1993). Over time, however, the hedge ratios implemented for the portfolios will converge to the long-run mean as the series are stationary (Enders, 2015, p. 255).

7.1.4. Hedging Performance

In this sub-section, the hedging performance of the in-sample estimated hedge ratios will be presented. As explained in sub-section 5.4, the in-sample estimated hedge ratios are computed and applied to the full in-sample period and the three equally long sub-periods it contains. This is to obtain more robust results and to analyze whether the different characteristics of the sub-periods affect the hedging performance of the models. Table 15 reports the hedging results for all models with the use of monthly futures contracts.

Table 15 ***–*** *Hedging performance – monthly futures – in-sample*

Period	Risk metric	Unhedged	Naïve	OLS	CCC-GARCH	DCC-GARCH
Full in-sample period	Variance	3.36%	2.83%*	2.83%*	2.72%*	**2.72%*†**
	EHE	-	15.86%	15.89%	18.91%	**19.19%**
	VaR (95%)	-30.30%	-27.06%*	-27.08%*	-26.59%*	**-26.50%*†**
	VaR reduction	-	10.68%	10.60%	12.23%	**12.54%**
Sub-period 1	Variance	1.64%	1.46%*	1.35%*	**1.33%***	1.36%*
	EHE	-	11.32%	17.63%	**19.10%**	16.91%
	VaR (95%)	-21.03%	-18.73%*	-18.41%*	**-18.19%*†**	-18.34%*
	VaR reduction	-	10.93%	12.44%	**13.52%**	12.76%
Sub-period 2	Variance	5.05%	4.03%*	3.88%*	3.47%*	**3.40%*†**
	EHE	-	20.28%	23.27%	31.33%	**32.73%**
	VaR (95%)	-37.20%	-32.98%*	-32.21%*	-30.61%*	**-30.26%*†**
	VaR reduction	-	11.35%	13.42%	17.71%	**18.65%**
Sub-period 3	Variance	3.43%	3.03%*	**3.02%*†**	3.41%	3.41%
	EHE	-	11.69%	**11.88%**	0.64%	0.43%
	VaR (95%)	-30.70%	**-27.97%*†**	-28.04%*	-29.47%*	-29.48%*
	VaR reduction	-	**8.90%**	8.67%	4.02%	3.97%

Figures in bold denote the best performing model. The significance levels of the results are obtained with the bootstrapping technique described in sub-section 6.5, and all results are reported in Appendix 1.
***** *indicates significance at the 5% level when comparing each hedging model to the unhedged portfolio.*
† *indicates significance at the 5% level when comparing the performance of the best-performing dynamic model to the best-performing static model.*

By looking at the obtained results from the hedging models in Table 15, one immediate finding is that all the hedging models obtain both variance and VaR reductions of the portfolios' returns compared to a no-hedge strategy in all periods. This indicates that actors operating in the market would have reduced the volatility of their electricity portfolios by hedging with monthly futures contracts during the examined period. The results from the t-tests support this, as it is found that almost all the hedging models obtain significantly lower portfolio variance and VaR compared to an unhedged spot portfolio.

In addition to assessing the statistical significance of the hedged portfolios against the spot portfolio, it is also carried out tests to check whether there are significant differences between the best-performing dynamic model and the best-performing static model in each period. The results show that there is a significant difference between the best-performing dynamic and static model in most periods. The only exception is sub-period 1 in which there is found no significant difference when testing whether the variance from the CCC-GARCH model is significantly lower than the variance of the OLS model.

By first considering the variance reductions, it can be seen that although the hedging models reduce the risk compared to the unhedged portfolio, the hedging performance varies considerably across periods. Specifically, the variance reductions (EHE) from the best-performing hedging model range from 11.88% in sub-period 3 to 32.73% in sub-period 2. What is evident for sub-period 1 is that the CCC-GARCH model obtains the best results based on both risk metrics. This implies that there was no gain from modeling a time-varying structure of the correlation between the return series for a hedger in this sub-period. The results are somewhat expected based on the parameters estimated and discussed in sub-section 7.1.2. Sub-period 2 stands out as the period with best hedging results, where all hedging models obtain variance reductions above 20%. Sub-period 3, however, shows a more interesting finding. During this period, both GARCH models perform worse than the static models, especially when measured by variance reduction. Both models are only barely reducing the variance compared to the unhedged portfolio, and neither of them obtains a variance that is significantly lower than the unhedged portfolio. The static models, on the other hand, obtain a significant variance reduction of almost 12%. Furthermore, the OLS model obtains a significantly lower variance than the best GARCH model, and the naïve model obtains a significantly lower VaR than the best GARCH model. This means that the GARCH models are not accurately modeling the volatility dynamics of the spot and futures market during this period. The reason for this inaccuracy is not apparent, but could be due to the large spikes in the squared residuals of the futures returns at the end of sub-period 3 (see Appendix 3.2). Since the coefficients in-sample are constant, the GARCH models may have difficulties in capturing these changes.

Further considering the VaR reductions, the best-performing model obtains a reduction of 18.65% in sub-period 2, which is the most effective hedging period, while only a reduction of 8.67% is obtained in sub-period 3, which is the least effective hedging period based on this risk metric. This suggests that if a hedger obtains good hedging results in one period, the hedger cannot necessarily expect equally good results in the successive period. The VaR reductions in the full in-sample period are lesser in size compared to the variance reductions, as these are found in the range of 10.60% to 12.54%. This is not surprising as the hedging models are constructed with hedge ratios minimizing the portfolio variance and not the VaR. To put the VaR measure into economic terms, if an electricity producer holds a spot-futures portfolio with a value of €1 million, then hedging with the best performing model (DCC-GARCH) would have reduced the producer's value at risk by €38,000 (€303,000 – €265,000).

As previously explained, no static hedge ratio can outperform the OLS estimated hedge ratio in-sample when it comes to variance reduction, which is clearly the case for all periods examined. Interestingly, however, the naïve hedge manages to obtain a higher VaR reduction compared to the OLS hedge in both the full period and sub-period 3. The OLS estimated hedge ratio is 0.63 and 0.96 in these two periods, ergo lower than the naïve hedge ratio of 1. This suggests that a hedger in the market aiming to reduce downside risk could potentially benefit from applying the naïve hedge when the OLS model suggests a lower hedge ratio.

By considering the full in-sample period, it can be inferred that an actor operating in the Nordic power market could have obtained variance reductions ranging from 15.86% to 19.19% compared to a no-hedge strategy, depending on the model choice. The best-performing model over the full in-sample period is the DCC-GARCH model with a variance reduction of 19.19%. It obtains a significantly lower variance compared to both the unhedged portfolio and the OLS model, but not compared to the CCC-GARCH model. This indicates that GARCH models are more effective than static hedging models when hedging with monthly futures in the Nordic power market, but also that more parsimonious GARCH models, such as the CCC-GARCH model, are just as good as more advanced models, such as the DCC-GARCH model.

As discussed in sub-section 5.4, the volatility in the spot market varies across the periods, and the standard deviation of the returns in the spot market was found to be lowest in sub-periods 1 and 3 (see Table 2). In sub-period 3, the OLS model outperformed both GARCH models, and in sub-period 1, only one of the GARCH models performed better than the OLS model. Sub-period 2 is the most volatile sub-period, and it is evident that the GARCH models obtain a significantly lower variance and value at risk than the static models in this sub-period. This suggests that dynamic hedging models are more effective than static hedging models when the volatility in the spot market is high. This has a practical implication for hedgers in the Nordic power market, as it suggests that dynamic hedging models are most beneficial to apply in times when hedging is most important.

What is also evident from Table 15 is that the risk reductions seem to be higher in periods when the correlation between the spot and futures returns is higher. This is shown by the greatest risk reductions in sub-periods 1 and 3, which are the two periods with the highest correlation between the spot and futures returns (see Table 2). The correlation coefficient is lowest in sub-period 3, and this is the period with the lowest risk reductions. This is an intuitive result, as it is essential that the hedging instrument has a strong correlation with the underlying asset in order to offset the risk of price fluctuations in the spot price.

By using the approach suggested by Byström (2003) described in sub-section 6.4.1, the results in Table 16 are obtained.

Table 16 *– Byström approach – monthly contracts – in-sample*

	Absolute returns: $\lvert r_\pi \rvert < \lvert r_s \rvert$		**Returns:** $r_\pi < r_s$	
Model	Number of times	% of full sample	Number of times	% of full sample
Naïve	249	50.92%	275	43.76%
OLS	251	51.33%	275	43.76%
CCC-GARCH	253	51.74%	275	43.76%
DCC-GARCH	247	50.51%	275	43.76%

The number of times the weekly return of any of the hedged portfolios is less (in both absolute and real terms) than that of the unhedged spot portfolio when hedging with monthly futures. The total number of weekly observations for the in-sample period is 489.

By looking at Table 16, it can be inferred that the DCC-GARCH model performs worse than the other models when using this approach, while the CCC-GARCH achieves the best results. However, the number of times the returns from the OLS model and the GARCH models lies below the returns from a spot position is almost identical. This evaluation criterion benefits the OLS model relatively more compared to the other models, and the results from Table 16 indicate that the variances obtained from the OLS model are affected by a small number of relatively large absolute returns. This shows that even though the GARCH models, overall, outperform the OLS model when evaluated by variance and VaR, it is clear that the OLS model tends to reduce the variance almost as often as the CCC-GARCH model, and more often than the DCC-GARCH model. Although these results provide useful insights on the performance of the models, the OLS model should still be regarded as less favorable compared to the GARCH models given that one of the main purposes of hedging is to protect the portfolio against large price drops. Another remark from the application of Byström's (2003) approach is that the returns from the hedged portfolio are higher than the returns from holding only a spot position in over 56% of the weeks, regardless of what hedging model that is used.

To analyze whether the GARCH models are preferable when accounting for the expected portfolio return and the transaction costs in addition to the portfolio variance, the average weekly utility according to the utility function presented in sub-section 6.4.3 is computed. The average weekly utility and the utility gain compared to the unhedged portfolio is reported for the risk aversion parameters 4 and 6. Both levels of risk aversion are found in the literature. For instance, Grossman and Schiller (1981) estimate the risk-aversion parameter to be 4, while Friend and Hasbrouck (1982) estimate it to be 6. Both risk aversion parameters are applied to discover potential changes in the ranking of the models depending on the hedger's level of risk aversion. The expected portfolio return is set to the mean return of the portfolio, as suggested by Zhou (2016).

As previously mentioned, the transaction costs in the model represent the reduced returns that are caused by the costs of trading futures. In the analysis, the transaction cost is set to 0.5%. This is based on information received from KIKS, and this is assumed to be a good approximation for a typical actor in the market (Håkonsen, 2019). It is, however, important to note that the transaction costs vary from actor to actor, but an approximation is necessary to use in this case. For the dynamic models, the transaction cost is accounted for in the utility function each week as these models require a weekly rebalancing of the hedged portfolio. The static hedging models assume a time-invariant hedge ratio, but the hedger incurs transaction costs when rolling over the monthly futures, which is at the end of each month (or every fourth week). The transaction costs are therefore assumed to be 0.125% (0.5% / 4) on an average weekly basis for the static models. The results are reported in Table 17, for all in-sample sub-periods.

Table 17 *– Average weekly utility - monthly futures – in-sample*

Period			**Unhedged**	**Naïve**	**OLS**	**CCC-GARCH**	**DCC-GARCH**
Full in-sample period	$\lambda = 4$	Utility	-0.1358	-0.1084	-0.1086	-0.1084	**-0.1075**
		Utility gain		0.0275	0.0272	0.0275	**0.0283**
	$\lambda = 6$	Utility	-0.2030	-0.1649	-0.1651	-0.1629	**-0.1618**
		Utility gain		0.0381	0.0379	0.0402	**0.0412**
Sub-period 1	$\lambda = 4$	Utility	-0.0652	-0.0483	**-0.0482**	-0.0504	-0.0509
		Utility gain		0.0169	**0.0170**	0.0148	0.0143
	$\lambda = 6$	Utility	-0.0980	-0.0774	**-0.0752**	-0.0770	-0.0782
		Utility gain		0.0206	**0.0228**	0.0210	0.0198
Sub-period 2	$\lambda = 4$	Utility	-0.2044	-0.1621	-0.1546	-0.1435	**-0.1403**
		Utility gain		0.0423	0.0498	0.0608	**0.0640**
	$\lambda = 6$	Utility	-0.3054	-0.2426	-0.2321	-0.2129	**-0.2083**
		Utility gain		0.0628	0.0733	0.0925	**0.0971**
Sub-period 3	$\lambda = 4$	Utility	-0.1396	**-0.1159**	-0.1166	-0.1324	-0.1325
		Utility gain		**0.0237**	0.0230	0.0072	0.0071
	$\lambda = 6$	Utility	-0.2081	**-0.1764**	-0.1770	-0.2005	-0.2007
		Utility gain		**0.0318**	0.0311	0.0077	0.0074

Utility gain is the utility obtained from each hedging model compared to the unhedged portfolio. Figures in bold denote the best performing model.

It can be inferred that each hedging model leads to gain in the utility compared to the no-hedge strategy in each period, suggesting that hedging, in general, is beneficial for an actor with the proposed utility function and a risk aversion above 4. Considering the differences in performance between the models, the rankings in the different periods are overall the same as when examining the variance and VaR, but there are a few exceptions. The most noteworthy is from sub-period 1 in which both static models outperform the GARCH models. This is in sharp contrast to the ranking of the models based on variance and value at risk in Table 15. The reason for this change is that the transaction costs are added more frequently to the dynamic models as they require rebalancing every week.

The next part of this sub-section will present the hedging results for the quarterly futures and a brief summary of the results for both contracts along with comparisons. Table 18 reports the hedging results when measured by variance and VaR reduction.

***Table 18** – Hedging performance – quarterly futures – in-sample*

Period	**Risk metric**	**Unhedged**	**Naïve**	**OLS**	**CCC-GARCH**	**DCC-GARCH**
Full in-sample period	Variance	3.36%	3.06%*	3.05%*	2.87%*	**2.85%*†**
	EHE	-	9.05%	9.08%	14.72%	**15.30%**
	VaR (95%)	-30.30%	-28.43%*	-28.45%*	-27.60%*	**-27.49%*†**
	VaR reduction	-	6.17%	6.10%	8.89%	**9.27%**
Sub-period 1	Variance	1.64%	1.56%*	1.47%*	**1.41%*†**	1.43%*
	EHE	-	5.12%	10.71%	**14.25%**	12.99%
	VaR (95%)	-21.03%	-19.71%*	-19.42%*	**-19.19%*†**	-19.28%*
	VaR reduction	-	6.29%	7.66%	**8.74%**	8.29%
Sub-period 2	Variance	5.05%	4.37%*	4.26%*	3.84%*	**3.76%*†**
	EHE	-	13.45%	15.76%	24.02%	**25.62%**
	VaR (95%)	-37.20%	-34.65%*	-34.20%*	-32.31%*	**-31.95%*†**
	VaR reduction	-	6.87%	8.08%	13.13%	**14.11%**
Sub-period 3	Variance	3.43%	3.27%*	**3.27%*†**	3.38%	3.38%
	EHE	-	4.61%	**4.69%**	1.28%	1.24%
	VaR (95%)	-30.70%	**-29.33%*†**	-29.40%*	-29.77%*	-29.78%*
	VaR reduction	-	**4.47%**	4.25%	3.03%	3.00%

Figures in bold denote the best performing model. The significance levels of the results are obtained with the bootstrapping technique described in sub-section 6.5, and all results are reported in Appendix 1.
***** *indicates significance at the 5% level when comparing each hedging model to the unhedged portfolio.*
† *indicates significance at the 5% level when comparing the performance of the best-performing dynamic model to the best-performing static model.*

The best-performing model according to each risk metric in each sub-period is the same for both contract lengths, and it can be seen that the hedging results vary considerably across the different periods also for the quarterly contracts. Just as for the monthly futures, the variance and VaR obtained by the different hedging models are significantly lower than those of the unhedged spot portfolio except for the GARCH models in sub-period 3. Regarding the difference in the performance of the best-performing dynamic and static model in each sub-period, it is found that the CCC-GARCH model in sub-period 1 has a significantly lower variance than the OLS model. It is also found that the sub-periods in which the hedging effectiveness is highest are the periods with the strongest correlation between the spot and futures returns, just as for the monthly futures.

As such, the qualitative differences between the models' results in the sub-periods are the same when analyzing both monthly and quarterly futures. However, there seems to be a general tendency that the hedging performance, overall, is relatively better for monthly futures. As an example, the hedger obtains a variance reduction of 19.19% for the DCC-GARCH model when hedging with monthly futures, but a variance reduction of only 15.10% is obtained for the same model when hedging with quarterly futures. One reason for this can be the relatively lower correlation between the spot and quarterly futures returns compared to the correlation between the spot and monthly futures returns, as shown in Table 2. The small differences between the results obtained from the two GARCH models can be explained by the results from the test for non-constant correlation in sub-section 6.2.5.2, in which no evidence of a time-varying correlation structure between the spot and quarterly futures returns was found.

Just as with the hedging effectiveness measure (EHE), the VaR metric shows worse results than those obtained with monthly futures. The greatest VaR reduction from each period range from 4.47% for the naïve hedge in sub-period 3 to 14.11% for the DCC-GARCH model in sub-period 2. Considering the full in-sample period, the implication is that if a hedger has an amount of €1 million exposed in the electricity market, hedging with quarterly futures using the best performing model (DCC-GARCH) would reduce the value at risk by €28,100 (€303,000 – €274,900).

By using the approach suggested by Byström (2003) described in sub-section 6.3.1, the results in Table 19 are obtained for the quarterly futures.

***Table 19** – Byström approach – quarterly contracts – in-sample*

	Absolute returns: $|r_\pi| < |r_s|$		**Returns:** $r_\pi < r_s$	
Model	Number of times	% of full sample	Number of times	% of full sample
Naïve	244	49.90%	261	46.63%
OLS	245	50.10%	261	46.63%
CCC	247	50.51%	261	46.63%
DCC	243	49.69%	261	46.63%

The number of times the weekly return of any of the hedged portfolios is less (in both absolute and real terms) than that of the unhedged spot portfolio when hedging with quarterly futures. The total number of weekly observations for the in-sample period is 489.

The ranking of the models according to this approach is the same as in the case of the monthly contracts. Again, this indicates that the relatively higher variances and VaR estimates reported for the static hedging models compared to the dynamic models are due to a few observations with high absolute returns. Another takeaway is that the returns from the hedged portfolio are higher than the returns from holding only a spot position in over 53% of the weeks, regardless of what hedging model that is used.

It is again examined whether the models with time-varying hedge ratios are preferred over models with time-invariant hedge ratios when applying the mean-variance utility function. The only difference in the setup is that the transaction costs accounted for in the static hedges are now lower as the quarterly contracts have a longer contract length compared to the monthly contracts. As a result, they only require rebalancing every three months (or every 12th week). The weekly transaction costs for the static models when hedging with quarterly futures are therefore estimated to 0.04167% (0.5% / 12). The obtained results are given in Table 20.

Table 20 *- Average weekly utility - quarterly contracts – in-sample*

Period			**Unhedged**	**Naïve**	**OLS**	**CCC-GARCH**	**DCC-GARCH**
Full in-sample period	$\lambda = 4$	Utility	-0.1358	-0.1193	-0.1197	-0.1172	**-0.1162**
		Utility gain		0.0165	0.0162	0.0186	**0.0196**
	$\lambda = 6$	Utility	-0.2030	-0.1804	-0.1808	-0.1745	**-0.1731**
		Utility gain		0.0226	0.0223	0.0285	**0.0299**
Sub-period 1	$\lambda = 4$	Utility	-0.0652	**-0.0508**	-0.0578	-0.0581	-0.0584
		Utility gain		**0.0144**	0.0075	0.0072	0.0068
	$\lambda = 6$	Utility	-0.0980	**-0.0801**	-0.0889	-0.0862	-0.0870
		Utility gain		**0.0179**	0.0091	0.0118	0.0111
Sub-period 2	$\lambda = 4$	Utility	-0.2044	-0.1732	-0.1780	-0.1594	**-0.1560**
		Utility gain		0.0312	0.0264	0.0450	**0.0484**
	$\lambda = 6$	Utility	-0.3054	-0.2583	-0.2654	-0.2362	**-0.2311**
		Utility gain		0.0472	0.0400	0.0692	**0.0743**
Sub-period 3	$\lambda = 4$	Utility	-0.1396	**-0.1270**	-0.1279	-0.1355	-0.1356
		Utility gain		**0.0126**	0.0117	0.0041	0.0040
	$\lambda = 6$	Utility	-0.2081	**-0.1923**	-0.1932	-0.2031	-0.2033
		Utility gain		**0.0158**	0.0149	0.0050	0.0049

Utility gain is the utility obtained from each hedging model compared to the unhedged portfolio. Figures in bold denote the best performing model.

The results show that the DCC-GARCH model still outperforms the other models in sub-period 2 and the full period when the transaction costs are considered. However, the results are more mixed compared to when variance and VaR were the performance measures as the naïve hedge is now the preferred model in sub-period 1 and 3. Additionally, the results from the utility analysis shows smaller differences in performance between the models compared to the risk reduction analysis. This shows how the transaction costs can affect the choice of hedging model.

7.1.4.1. Summary of In-Sample Analysis

This sub-section has presented the hedging results from the in-sample analysis. The results indicate that the dynamic models are the most effective hedging strategies as they show the best performance among the models, overall. It can be inferred that an actor operating in the Nordic power market could have obtained a total variance reduction of 19.19% and 15.30% over the full in-sample period by using the DCC-GARCH model when hedging with monthly and quarterly futures, respectively. The variance obtained by the DCC model is significantly lower compared to both the unhedged portfolio and the OLS model, but not compared to the CCC-GARCH model. The same is shown for VaR reduction in the same period. One interesting finding for both contracts is that the OLS model significantly outperforms the GARCH models in sub-period 3. Hence, the GARCH models struggle to capture the volatility characteristics in the return series during this period.

It is found that the differences in the hedging results are smaller when using the mean-variance utility function to assess the models' performances. A notable finding from the results is that there is a substantial variation in the hedging effectiveness of the models across the sub-periods. One potential reason for this could be that the high and time-varying volatility in electricity markets makes it challenging to obtain good hedging results over time (Hanly et al., 2017). It was also found indications that the dynamic models obtain relatively better hedging results compared to the static models in periods with higher volatility in the spot market. Comparing the results for the monthly and quarterly futures, the ranking of the models is practically the same. However, the risk reductions with quarterly futures are significantly lower than the risk reductions obtained with monthly futures. This could be explained by the fact that the spot returns are more correlated to the monthly futures returns than the quarterly futures returns. An implication of this is that actors hedging a quarterly delivery of electricity obtain lower risk reductions in comparison to actors hedging a monthly delivery. The next sub-section will present the out-of-sample analysis to increase the robustness of the hedging results.

7.2. Out-of-Sample Analysis

The optimal hedge ratios from sub-sections 7.1.1 and 7.1.3 are computed ex-post. In reality, however, hedgers need to decide on their hedge ratio ex-ante. This sub-section will present the findings from the out-of-sample analysis as described in sub-section 6.3. As previously mentioned, this is to increase the robustness of the results in the analysis.

7.2.1. Forecasted Hedge Ratios

In this sub-section, the hedge ratios applied in the out-of-sample analysis will be presented. The time-varying hedge ratios from the GARCH models are based on one-step-ahead forecasts for the conditional covariances and variances using the rolling window method explained in sub-section 6.3. Figure 17 shows the obtained hedge ratios from the forecasts for the dynamic models when using monthly futures.

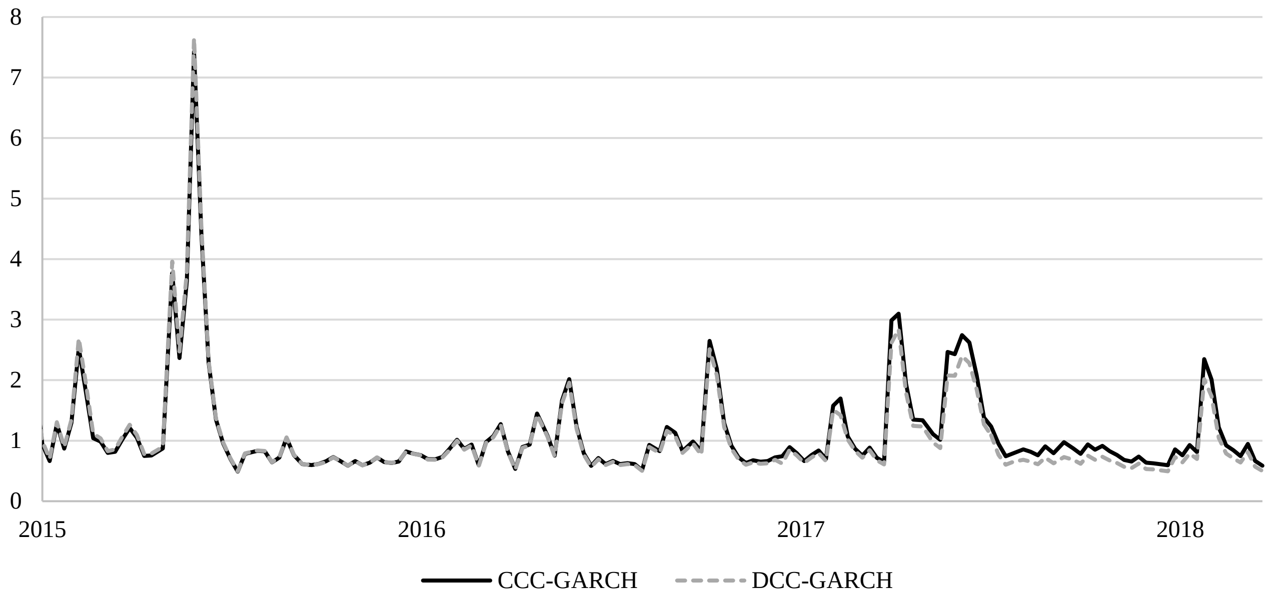


***Figure 17** – Forecasted dynamic hedge ratios - monthly futures – out-of-sample*

Figure 17 shows that the hedge ratios from the CCC-GARCH and the DCC-GARCH model are very similar. This is found even though the correlation in the CCC-GARCH model out-of-sample is based on historical data up to the last week in the in-sample period, while the DCC-GARCH model forecasts the conditional correlation for each week by adding additional observations from the out-of-sample period. The first noticeable difference between the models is that the CCC-GARCH hedge ratio is higher than the DCC-GARCH hedge ratio in the spike at the beginning of 2016. After the spike, the differences between the two models are negligible until mid-2017 where the CCC-GARCH model starts to produce somewhat higher hedge ratios. The mathematical reason for this is that forecasted conditional correlations for the DCC-GARCH model in this period become lower than the constant conditional correlation that is assumed for the CCC-GARCH model (see Appendix 3.3.1).

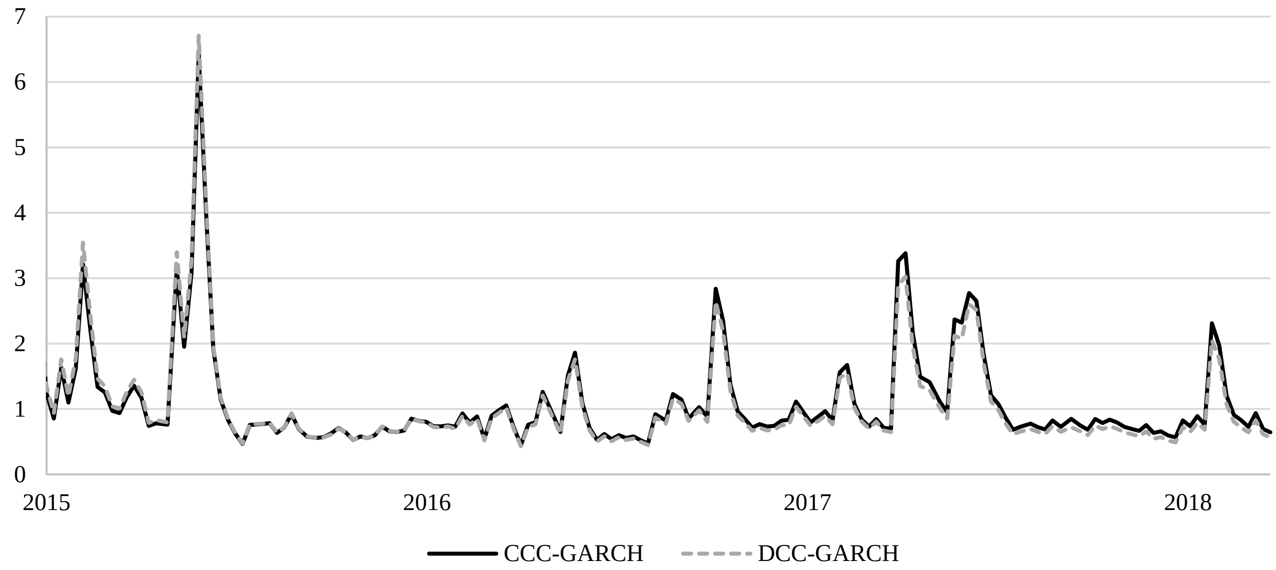


Figure 18 *– Forecasted dynamic hedge ratios - quarterly futures – out-of-sample*

Figure 18 shows the hedge ratios for the quarterly futures. Just like for the monthly futures, the forecasted dynamic hedge ratios for the two GARCH models are closely linked, with small differences between them. This implies that the forecasted conditional correlations for the DCC-GARCH model do not deviate much from the constant correlation for the CCC-GARCH model (see Appendix 3.3.2). To further examine the characteristics of the models, and to be able to make comparisons, descriptive statistics for the time series of the dynamic hedge ratios are presented in Table 21.

Table 21 *- Descriptive statistics for out-of-sample forecasted dynamic hedge ratios*

	Monthly futures		**Quarterly futures**	
	CCC-GARCH	DCC-GARCH	CCC-GARCH	DCC-GARCH
Min	0.488	0.493	0.448	0.431
Max	7.428	7.651	6.424	6.708
Mean	1.110	1.054	1.095	1.056
No. of times higher than 1	49 / 163	47 / 163	49 / 163	49 / 163
No. of times higher than β_{OLS}^{*}	53 / 163	52 / 163	55 / 163	51 / 163
ADF	-5.591***	-5.424***	-5.839***	-5.617***
Q^2	124.82***	132.97***	119.00***	134.62***

Min and max are the lowest and highest hedge ratios observed in the time series, respectively. Mean is the average of the dynamic hedge ratios. No. of times above 1 represents all observations that lie above the naïve hedge ratio. β_{OLS}^{*} *represents the OLS estimated hedge ratio of 0.960 for the monthly contracts and 0.943 for the quarterly contracts. ADF denotes the test statistic for the ADF test[26] for stationarity, with lags selected according to AIC. The critical value for the ADF test at the 1% significance level is -3.43.* Q^2 *is the Ljung-Box test statistic[27] measuring autocorrelation with 15 lags. The critical value for the Ljung-Box test at the 1% significance level with 15 degrees of freedom is 30.58. *** indicates rejection of the null hypothesis at the 1% significance level.*

In contradiction to the in-sample estimated hedge ratios, it is now shown that the CCC-GARCH model produces higher hedge ratios than the DCC-GARCH for both contract types, shown by the mean in Table 21. By comparing the hedge ratio for the monthly contracts to those for the quarterly contracts, one takeaway is that hedge ratio series for the quarterly contracts contain lower minimum and higher maximum hedge ratios than those found for the monthly contracts. This is the opposite results of those obtained from the in-sample analysis. All the hedge ratio series in the out-of-sample analysis are found to be stationary and to exhibit positive autocorrelation. These two properties imply that a high hedge ratio decided by the hedger in one week will lead to a high hedge ratio the next week in the absence of shocks, and that the hedge ratio will converge to the long-term mean.

7.2.2. Hedging Performance

This sub-section will present the performance of the hedging models out-of-sample. The hedging performance based on variance and value at risk is reported in Table 22.

[26] For more details on the ADF test, see sub-section 5.5.
[27] For more details on the Ljung-Box test, see sub-section 6.2.4.

***Table 22** – Hedging performance - monthly futures – out-of-sample*

Period	Risk metric	Unhedged	Naïve	OLS	CCC-GARCH	DCC-GARCH
Out-of-sample (sub-period 4)	Variance	4.81%	4.63%	4.63%	**4.28%*†**	4.28%*
	EHE	-	3.61%	3.72%	**11.02%**	10.93%
	VaR (5%)	-35.36%	-35.65%	-35.59%	-34.24%*	**-34.14%*†**
	VaR reduction	-	-0.81%	-0.64%	3.18%	**3.45%**

Figures in bold denote the best performing model. Figures in red denote results worse than the unhedged portfolio. The significance levels of the results are obtained with the bootstrapping technique described in sub-section 6.4., and all results are reported in Appendix 1.
** indicates significance at the 5% level when comparing each hedging model to the unhedged portfolio.*
† indicates significance at the 5% level when comparing the performance of the best-performing dynamic model to the best-performing static model.

Table 22 shows that both dynamic hedging models obtain a significantly lower variance compared to the unhedged spot portfolio, suggesting that they are effective in reducing the portfolio risk. In contrast, neither of the static hedging models manage to obtain a significantly lower variance or VaR compared to the unhedged portfolio. This contradicts the findings from the in-sample analysis in which all the hedging models in the analysis obtained a significantly lower variance and VaR compared to a no-hedge strategy. The poor performance of the static models out-of-sample could be due to the time-varying volatility found for the Nordic power market, implying that an optimal hedge ratio in one period could be suboptimal when applied in another period. The best-performing model based on variance is the CCC-GARCH model, which obtains a significantly lower variance than the best-performing static hedge, which is the OLS model.

Regarding the VaR metric, the DCC-GARCH model obtains the best result of the models with a VaR that is significantly lower than both the unhedged portfolio and the best static hedging model. An interesting finding is that both static models fail in reducing the VaR compared to the unhedged portfolio. Although there is not found evidence of a statistically significantly higher VaR for either static model compared to the unhedged portfolio based on the bootstrapping and t-tests, the results could suggest that static hedging strategies struggle to perform in an out-of-sample context.

In Table 2, the descriptive statistics show that the volatility in the spot market during sub-period 4 is the second highest after sub-period 2. The out-of-sample analysis therefore further suggest that dynamic hedging models are recommended in times of relatively high market volatility. This is a result that substantiates the findings from the in-sample analysis, in which the GARCH models are found to provide better hedging results than the static models in periods with higher volatility in the underlying asset.

Compared to the results from the full in-sample period, it is evident that the risk reductions are lower in the out-of-sample analysis. Three factors could explain this. First, the hedging results from the out-of-sample analysis rely on forecasted time-varying hedge ratios, and it is possible that these are somewhat less accurate than those that are estimated and applied in-sample. Second, the hedge ratio applied from the OLS model is estimated in a different period than the evaluation period it is tested in. It is therefore expected that this model performs worse out-of-sample than in-sample due to the time-varying volatility in the market. Third, it can be seen from Table 2 that the correlation between the spot and futures returns are lowest in sub-period 4, and this would generally suggest a lower hedging performance.

The results from the approach suggested by Byström (2003) is reported in Table 23 to further investigate the performance of the models.

Table 23 *– Byström approach – monthly futures – out-of-sample*

Absolute returns: $|r_\pi| < |r_s|$		
Model	Number of times	% of full sample
Naïve	79	48.47%
OLS	79	48.47%
CCC	84	51.35%
DCC	85	52.15%

The number of times the weekly return of any of the hedged portfolios is less (in absolute terms) than that of the unhedged spot portfolio when hedging with monthly futures. The total number of weekly observations for the out-of-sample period is 163.

It can now be inferred that the naïve hedge and the OLS model perform worse than the dynamic models. Compared to the in-sample analysis, the GARCH models are now reducing the variance more often than the static models. This makes intuitive sense as the GARCH models are estimated based on forecasts, while the OLS model is unconditional and depends only on historical data up to the point of selecting the hedge ratio.

The results from the mean-variance utility function are presented in Table 24. As previously mentioned, this approach incorporates three aspects that the portfolio variance and VaR do not account for: the portfolio returns, the transaction costs, and the risk aversion of the hedger.

Table 24 *- Average weekly utility – monthly futures – out-of-sample*

Period			**Unhedged**	**Naïve**	**OLS**	**CCC-GARCH**	**DCC-GARCH**
Out-of-sample	$\lambda = 4$	Utility	-0.1853	-0.1890	-0.1884	-0.1783	**-0.1773**
		Utility gain		-0.0037	-0.0032	0.0070	**0.0079**
	$\lambda = 6$	Utility	-0.2814	-0.2816	-0.2809	-0.2638	**-0.2629**
		Utility gain		-0.0003	0.0004	0.0176	**0.0184**

Utility gain is the utility obtained from each hedging model compared to the unhedged portfolio. Figures in bold denote the best performing model.

The results reveal that the dynamic GARCH models still outperform the other models when using the utility approach. This is not surprising considering the substantial differences between the model types when measured by risk reduction. An interesting finding from this analysis is that the static hedges obtain an average weekly utility that is lower than that of the no-hedge strategy in three out of four cases. The OLS model manages to obtain a higher utility than the unhedged portfolio only when the risk aversion parameter is set to 6. This means that a hedger with a relatively high risk aversion would prefer the OLS model over a no-hedge strategy. However, the DCC-GARCH model would be preferred over all the other models for both risk aversion parameters in the analysis.

The alternative version of Kroner and Sultan’s (1993) utility framework presented in sub-section 6.4.3 has also been adopted in the analysis. This utility framework provides the hedger with the choice of rebalancing in each week. The results are reported in Table 25.

Table 25 *– Total utility when deciding to rebalance or not in each week - monthly futures – out-of-sample*

	Unhedged	**Naïve**	**OLS**	**DCC-GARCH**
Utility	-33.10	-30.16	-30.17	**-28.38**
Number of rebalances	0	38	38	61

The number of rebalances are based on the utility benefit compared to the transaction costs. The potential number of rebalances are 163. Figures in bold denote the best performing model.

As described in sub-section 6.4.3, the ranking of the models is based on the total utility from each week during the full out-of-sample period. The results show that a utility-maximizing market participant will choose to rebalance the portfolio in 36.8% (61/163) of the weeks in the sample. The naïve hedge and the OLS model incur transaction costs at the end of each month when they roll over to the next futures contract, but the hedge ratio in the portfolios stays the same during the entire period for both models. As shown in Table 25, the hedger obtains a higher total utility when following the rebalancing strategy from the DCC-GARCH models compared to both static models and the unhedged strategy. Therefore, giving the hedger a choice of either rebalancing the hedged portfolio or keeping the current hedge ratio based on the expected utility in the following period benefits the DCC-GARCH model. This shows the advantage of a dynamic hedging model compared to a strict static model in an applied context.

The remainder of this sub-section will present and discuss the hedging results in the out-of-sample period for quarterly futures. Table 26 shows the hedging results based on the risk reduction metrics applied in the thesis.

Table 26 *- Hedging performance - quarterly futures – out-of-sample*

Period	**Risk metric**	**Unhedged**	**Naïve**	**OLS**	**CCC-GARCH**	**DCC-GARCH**
Out-of-sample (sub-period 4)	Variance	4.81%	4.73%	4.72%	**4.64%*** †	4.65%*****
	EHE		1.62%	1.82%	**3.35%**	3.16%
	VaR (5%)	**-35.36%**	-35.86%	-35.78%	-35.56%	-35.53% †
	VaR reduction		-1.42%	-1.18%	-0.58%	-0.49%

Figures in bold denote the best performing model. Figures in red denote results worse than the unhedged portfolio. The significance levels of the results are obtained with the bootstrapping technique described in sub-section 6.4., and all results are reported in Appendix 1.

***** *indicates significance at the 5% level when comparing each hedging model to the unhedged portfolio.*

† *indicates significance at the 5% level when comparing the performance of the best-performing dynamic model to the best-performing static model.*

Just like for the monthly futures, neither of the static hedging models obtains a variance or a VaR that is significantly lower than the unhedged portfolio. The CCC-GARCH model performs best when it comes to variance reduction, and the variance is also significantly lower compared to both the OLS model and the unhedged portfolio. The DCC-GARCH model is almost just as good as the CCC-GARCH model when it comes to variance reduction, and it also obtains a variance that is significantly lower than that of the unhedged portfolio.

When considering VaR, none of the hedged portfolios manages to outperform the unhedged portfolio. This is interesting and shows that a hedged portfolio does not guarantee a risk reduction compared to an unhedged portfolio, although the analysis show that this is usually the case. Furthermore, it is again evident that the quarterly contracts lead to lower risk reduction compared to the monthly contracts, overall.

These results are also worse than the overall results from the in-sample analysis, probably due to the same three reasons as described for the monthly contracts. To further examine the performance of the models, the results from the Byström (2003) approach is presented in Table 27.

Table 27 *– Byström approach – quarterly futures – out-of-sample*

Absolute returns: $\lvert r_\pi \rvert < \lvert r_s \rvert$		
Model	Number of times	% of full sample
Naïve	73	44.79%
OLS	74	45.40%
CCC	77	47.24%
DCC	77	47.24%

The number of times the weekly return of any of the hedged portfolios is less (in absolute terms) than that of the unhedged spot portfolio when hedging with quarterly futures. The total number of weekly observations for the out-of-sample period is 163.

The results in Table 26 show that the static hedging models still perform worse than the dynamic hedging models. This contradicts the findings from the in-sample analysis in which it was found that the OLS model produced weekly absolute returns that were lower than the unhedged portfolio on most occasions. Again, the reason for this could be that the OLS hedge ratio is estimated in-sample.

The average weekly utility from the models is presented in Table 28.

Table 28 *- Average weekly utility – quarterly futures – out-of-sample*

Period			Unhedged	Naïve	OLS	CCC-GARCH	DCC-GARCH
Out-of-sample	$\lambda = 4$	Utility	**-0.1853**	-0.1901	-0.1900	-0.1919	-0.1916
		Utility gain		-0.0048	-0.0048	-0.0067	-0.0064
	$\lambda = 6$	Utility	**-0.2814**	-0.2845	-0.2846	-0.2848	-0.2847
		Utility gain		-0.0031	-0.0032	-0.0035	-0.0034

Utility gain is the utility obtained from each hedging model compared to the unhedged portfolio. Figures in bold denote the best performing model.

It is found that none of the hedging models manages to produce a higher average weekly utility than the unhedged portfolio. This result suggests that the variance reductions obtained from the hedging models in Table 28 are not enough to outweigh the transaction costs of adjusting the portfolios. To further examine the utility comparisons, the rebalancing approach (Kroner & Sultan, 1993) is included in the utility framework. Table 29 reports the results with the utility values representing the total utility for the hedger during the full out-of-sample period.

Table 29 *–Total utility when deciding to rebalance or not in each week - quarterly futures – out-of-sample*

	Unhedged	**Naïve**	**OLS**	**DCC-GARCH**
Utility	-33.10	-31.08	-31.09	**-30.00**
Number of rebalances	0	13	13	37

The number of rebalances are based on the utility benefit compared to the transaction costs. The potential number of rebalances are 163. Figures in bold denote the best performing model.

The results in Table 29 show that a utility-maximizing market participant hedging with quarterly futures will choose to rebalance the portfolio in only 22.70% (37/163) of the weeks, which is less than in the analysis of the monthly futures. The number of rebalances can be viewed in relation to the hedge ratio dynamics discussed in sub-section 7.2.1. The presence of autocorrelation in the hedge ratio series imply that next week’s hedge ratio often is close to this week’s hedge ratio. Rebalancing the portfolio in these weeks is therefore in most cases not worthwhile when considering the impact of transaction costs. For the naïve hedge and the OLS model, the transaction costs are accounted for at the end of each quarter when the hedger rolls over to the next contract. Just like for the monthly futures, this setup benefits the GARCH model and leads to a higher total utility when adding the weekly utility functions together. As a result, a hedger maximizing the expected utility for the coming period would prefer the GARCH model over the static hedges also when hedging with quarterly futures contracts.

7.2.2.1. Summary of Out-of-Sample Analysis

The findings from the out-of-sample analysis provide essential insights for answering the research question with the corresponding sub-questions of the thesis, and the key takeaways will be summarized in the following.

Overall, the results from the out-of-sample analysis show that the GARCH models are preferred over the static hedging models. This is shown by a variance reduction for the best-performing GARCH model that is significantly lower than both the OLS model and the unhedged portfolio for both contract types. However, one interesting finding is that all hedging models fail to reduce the VaR when hedging with quarterly futures. It is found that both GARCH models obtain a higher average weekly utility than all other hedges for the monthly contracts, but none of the hedging models manage to obtain a higher utility than the unhedged portfolio when examining the quarterly contracts. By applying the rebalancing model of Kroner and Sultan (1993), it was found that the DCC-GARCH model produced a higher utility when summing the utility functions for each week for both contract types. Thus, a dynamic hedging model is shown to be preferred by a mean-variance utility maximizing hedger for both contract types.

The out-of-sample period is characterized by a relatively high volatility in the spot market and a low correlation between the spot and futures returns. As the dynamic models are, overall, found to be the preferred models for this period, this further indicates that dynamic models are beneficial in periods with high volatility in the spot market. The hedging results from monthly contracts out-of-sample are also found to produce a significantly lower variance and VaR, compared to quarterly contracts.

8. Discussion

This section will extend the discussion from the empirical results and provide interpretations and implications for the main findings of the thesis. The results will also be compared to findings from existing research, and it will be discussed how the results of this thesis add value to the literature. Limitations of the results as well as the aspects of implementing the hedging models in practice will also be discussed.

The main choices made for the analysis in this thesis are the frequency of the data, the contract length for the futures, how often to rebalance, the rollover procedure, the type of hedging models, and the transaction costs. These are all justified throughout the thesis, based on research and literature on the field. It should be noted that adjusting these choices will possibly affect the outcome of the analysis.

A general finding from the in-sample analysis is that all hedging models significantly reduce the variance and value at risk of the portfolio compared to an unhedged portfolio for both contract lengths considered. This suggests that an actor in the Nordic power market seeking to reduce the risk of spot price fluctuations could benefit from hedging with power futures. However, the results indicate that differences in the hedging performance are dependent on the choice of hedging model, the contract length, and the sub-period examined.

Overall, it is found that the dynamic hedging models outperform the static models when measured by risk reduction. By considering the full in-sample period, it is evident that the best-performing dynamic model obtains a significantly lower variance and value at risk compared to the best-performing static model. Regarding sub-period 1, there is found no evidence of a significantly lower variance for the most effective dynamic model compared to the OLS model for monthly contracts, while the value at risk metric shows significant differences for both contracts. Furthermore, sub-period 3 shows unexpected results as both dynamic hedging models perform worse than the static models. Moreover, it is found that the most effective dynamic model achieves a significantly higher variance and value at risk for the same sub-period compared to the OLS model and the naïve hedge. This poor performance by the GARCH models could be because the parameters from the maximum likelihood estimation are not able to adequately capture the volatility dynamics in this period, as described in sub-section 6.3. Andersen et al. (2003) highlight that the constant coefficients are a drawback of GARCH models. Consequently, the large spikes in the squared residuals of the futures returns (see Appendix 3.2) at the end of sub-period 3 could be one factor contributing to these poor results.

The results from sub-periods 2 and 4 show the most considerable differences in the risk reductions between the dynamic and the static hedging models. Furthermore, it is found that the most effective GARCH model in these periods obtains a significantly lower variance and value at risk when compared to those of the most effective static model. By considering the market characteristics in the different sub-periods, it was observed that the two most volatile periods in the spot market were sub-periods 2 and 4, while sub-periods 1 and 3 were the least volatile periods. The results therefore indicate, but do not prove, that the relative advantage of hedging with time-varying hedge ratios in the Nordic power market is greater when the spot market volatility is high, that is, when hedging is most important.

Compared to the static models, the GARCH models have the advantage that they allow the hedger to adjust the number of futures contracts relative to the spot position in the portfolio based on changing market conditions. This is particularly relevant in the Nordic power market as it is a highly volatile market due to the non-storability of the commodity. The result that the GARCH models, in general, tend to outperform the static models is therefore intuitive. Comparing the GARCH models, the DCC-GARCH model allows for time dependency in the conditional correlation between spot and futures returns while the CCC-GARCH assumes a constant conditional correlation. Intuitively, this would suggest that the hedge ratios from the DCC-GARCH would produce better results than the CCC-GARCH model. However, when testing whether the DCC-GARCH achieves lower variance or value at risk than the CCC-GARCH, significant evidence is only found in three out of 20 cases (see Appendix 1). It should be noted that one of these cases is for the full in-sample period for the monthly contracts in which the DCC-GARCH significantly outperforms the CCC-GARCH model, but no evidence of differences is found in the out-of-sample period. Overall, the results indicate that there are no or modest gains for a hedger in the Nordic power market to incorporate a time-varying structure between the correlation of spot and futures returns in the hedging model. This implies that more advanced models do not necessarily increase hedging performance compared to more parsimonious models.

The results further indicate that hedging effectiveness depends on the correlation between spot and futures returns. The highest correlation is found in sub-period 2 in which all hedges achieve relatively high risk reductions. The out-of-sample period shows the lowest correlation between spot and futures returns. The risk reductions out-of-sample are considerably lower than those in sub-period 2. This could be driven by the low correlation (see Appendix 3.3), but could also be a result of the hedge ratios being forecasted out-of-sample. Furthermore, the forecasting could also explain the poor results for the value at risk metric in this period.

The results show that hedging with monthly contracts leads to significantly higher risk reductions compared to quarterly contracts, but the rankings of the models are generally the same across sub-periods. This can be explained by the relatively higher correlation between spot and monthly futures returns in comparison to the correlation between spot and quarterly futures returns. This is a natural result as contracts with shorter delivery periods are expected to be more affected by changes in the spot market than those with longer delivery periods.

Even though variance and value at risk are widely used risk metrics in hedging literature (Berk & DeMarzo, 2014, p. 317; Bodie et al., 2011, p. 138), the main drawback of these metrics is that they do not account for the transaction costs associated with the models. Specifically, the dynamic hedging models are more costly compared to the static models as they require weekly rebalancing. Consequently, the models are placed in a mean-variance utility framework that accounts for this. Additionally, this framework considers the expected return and differences in risk aversion for the hedger. Application of this framework was expected to decrease the relative differences between the static and dynamic models. The results suggest that the GARCH models still outperform the static models both in-sample and out-of-sample for monthly contracts, while more mixed results were obtained for the quarterly contracts. By further including the hedger's choice of rebalancing based on the expected utility in the successive period, the dynamic models obtained the best results also for the quarterly contracts. These results indicate that GARCH models are adequate for actors in the market that are utility-maximizing and not only risk-minimizing. It should, however, be highlighted that these results are highly dependent on the transaction costs of the market participants. The transaction costs applied in this thesis are based on an approximation provided by KIKS. Even though this approximation is meant to be representative for a typical actor in the market, the exact transaction costs typically varies across actors. The results from the mean-variance utility framework therefore serve as an example of how transaction costs could affect the choice of hedging model for a hedger in the market.

Previous studies on the field present contradictory findings regarding the hedging performance of futures in the Nordic power market. Byström (2003) finds that the OLS model outperforms two versions of the multivariate GARCH model based on variance reduction but reports further that there are no significant differences between the performances of the models. Zanotti et al. (2009) find that GARCH models are superior to traditional static hedging models when it comes to variance reduction in the Nordic power market. They further highlight that for EEX, this is particularly the case when the volatility in the spot market is relatively high. This is consistent with what the results of this thesis suggests for the Nordic power market. Hanly et al. (2017) report inconclusive results as to whether GARCH models are more effective in reducing risk than the OLS model based on both variance and value at risk. Furthermore, they find no gains of extending the CCC-GARCH to a DCC-GARCH model, which is also in line with most of the cases examined in this thesis.

The results of this thesis add value to the literature on mainly three areas. First, a more extended sample period is examined, including the most recent price data, thus increasing the relevance and robustness of the results. Second, the hedging performance of the quarterly contracts is examined. Including these in the analysis is of high relevance as the traded volume of these contracts is of considerable size (Botterud et al., 2009). The results from the analysis of quarterly contracts can provide valuable insights for actors in the market, such as KIKS, who frequently trade quarterly futures for hedging purposes. Third, this thesis includes utility maximization as one of the performance measures. Byström (2003) notes that transaction costs are of high importance when choosing a hedging strategy but he does not analyze utility in his paper. Examining the hedging models in a mean-variance utility framework that accounts for expected return, transaction costs, and the risk aversion of the hedger therefore fills another gap in the existing literature.

Although transaction costs are accounted for, there are other aspects associated with the dynamic hedging models that are more difficult to include in an analytical framework. Implementation of dynamic strategies in practice requires higher analytical skills compared to simpler static models. In some periods, the volatility in the spot market is very high, and the estimated hedge ratios are multiple time the size of the static hedge ratios. In these periods the dynamic strategies require higher margin account deposits compared to the static models. The hedger is therefore required to have sufficient funds available when following these strategies. There are also higher operational risks associated with more complex models as they require weekly forecasts of the hedge ratio and a weekly rebalancing of the hedged portfolio (Byström, 2003). A limitation of this thesis is that the analysis does not account for this. Even though dynamic models can be too extensive for general market participants, the results of this thesis provide important insights on the potential benefits of following a dynamic model and how it relates to volatility clusters in the Nordic power market. However, it should be noted that static models also show significant reductions in portfolio risk in most cases, suggesting that futures hedging in general reduces the portfolio risk for actors operating in the market.

9. Conclusion

This thesis has examined the hedging effectiveness in the Nordic power market over the period from December 2006 to November 2018, by analyzing the performance of well-known hedging models from the literature. The full data sample was divided into four sub-periods, and both an in-sample and an out-of-sample analysis were performed to increase the robustness of the results. In addition, a bootstrapping technique was conducted in order to make statistical inferences of the hedging models' results. The following aim to answer the research question of the thesis and the corresponding sub-questions.

1. *Are hedging models with dynamic hedge ratios more effective in reducing risk compared to models with static hedge ratios?*

The empirical results reveal notable differences in the hedging performance of static and dynamic hedging models, both across periods and across different performance measures. It is found that the dynamic models obtain a significantly lower variance and value at risk compared to the static models in both the full in-sample and out-of-sample period. Additionally, the thesis finds no significant gains from extending the CCC-GARCH model to include a dynamic conditional correlation between the spot and futures returns.

2. *Does the performance of the hedging models depend on the volatility in the spot market?*

The varying results from the sub-periods indicate that the dynamic models do not always outperform the static models, and different market conditions seem to affect the hedging performance. Specifically, the results indicate that the relative advantage of the GARCH models compared to the static models is greater when the volatility in the spot market is high, which is when hedging is most important. However, a general trend of improved performance for all hedging models in periods of higher volatility is not found. The results suggest that hedging effectiveness varies substantially depending on the market situation and that the hedging effectiveness, in general, is higher when the correlation between the spot and futures returns is stronger.

3. *How do dynamic and static hedging models compare when accounting for transaction costs in a mean-variance utility framework?*

Measuring the hedging performance by applying the mean-variance utility framework shows smaller differences in performance between the models. However, the dynamic models are still outperforming the static models for both contract lengths overall. This is also found in the out-of-sample analysis when including the choice of rebalancing. Consequently, the results suggest that dynamic models are more effective than static models also for hedgers that are utility-maximizing and not only risk-minimizing.

4. *How does the hedging performance between the use of monthly and quarterly futures compare?*

It is found that the hedging models obtain a significantly lower variance and value at risk when hedging with monthly contracts compared to quarterly contracts, both in-sample and out-of-sample. The results across different sub-periods, as well as the results from the mean-variance utility framework, further suggest that greater hedging results are obtained with monthly futures compared to quarterly futures.

The answers from the sub-questions provide a solid basis for answering the main research question of the thesis:

How can actors in the Nordic power market most effectively hedge electricity price risk with futures?

The main conclusion that can be drawn from this thesis is that the most effective way actors in the Nordic power market can hedge electricity price risk with futures is to dynamically adjust the hedging position according to a multivariate GARCH model. However, it should be highlighted that the results from the sub-periods indicate that the hedging performance is period-specific, considering the substantial changes in the hedging results across sub-periods. Specifically, the advantage of hedging dynamically with a multivariate GARCH model seems to be greater in periods when the spot market volatility is relatively high.

For a market participant, the suggested hedging model based on the empirical results in the thesis is to apply a dynamic model when hedging a spot position with futures. As there was seldom a significant difference in the results of the CCC-GARCH model and the DCC-GARCH model, the CCC-GARCH model is recommended to be used in practice for parsimony reasons and simpler implementation.

10. Reflections and Suggestions for Further Research

In this section, alternative approaches will be presented along with suggestions for further research on the Nordic power market and power markets in general.

A natural extension of the analysis conducted in this thesis would be to analyze the performance of the hedging models from this thesis in different electricity markets around the world. Besides, it is also possible to conduct a similar analysis from the viewpoint of retailers or large consumers that have committed to buying electricity in the future, and therefore needs to take a long position in futures to mitigate the risk of spot price fluctuations. Furthermore, evaluating the hedging performance of other power derivatives traded at Nasdaq Commodities, such as options or EPADs, could also provide valuable insights for hedgers in the market.

When adding the monetary transaction costs and the costs of time spent on development and implementation of the dynamic hedging models, the actual cost of applying a dynamic hedging model can become substantially higher. A further analysis of operational risks or transaction costs associated with dynamic hedging can be subject for further research.

Based on the literature review in section 4, the chosen volatility models in this thesis were the constant conditional correlation (CCC) GARCH and the dynamic conditional correlation (DCC) GARCH model. However, numerous volatility models exist in the literature, and examination of other models could therefore be a topic for further research.

In September 2018, the Nordic power market reached the headlines of newspapers all around the world. One of the most successful traders in the market, Einar Aas, incurred a massive loss of €117 million (Karagiannopoulos, 2018). Aas had entered a spread position between German and Nordic power prices that was too big in relation to the liquidity in the market, and when he could not cover the losses, he was declared default by Nasdaq Clearing (Nasdaq, n.d.-d). As a consequence of this event, Nasdaq increased the margin requirements to reduce the risk profile in the market (Nasdaq, n.d.-d), which would suggest lower liquidity in the market. Also, large power companies highlighted that the absence of Aas in itself would damage the liquidity and that it has been an asymmetry in the risk-taking in the market for some time (Sagmoen & Jordheim, 2018). It would therefore be interesting to examine the long-term effects of this on hedging performance in the market, and this could be a direction for further research.

11. Bibliography

Abubaker, R. (2016). Consumption and Money Uncertainty at the Zero Lower Bound. *Economics Bulletin, Vol. 34 (1)*, 449-463.

Andersen, T., Bollerslev, T., Diebold, F., & Labys, P. (2003). Modeling and Forecasting Realized Volatility. *Econometrica, Vol. 71 (2)*, 579-625.

Andersen, T., Davis, R., Kreiss, J.-P., & Mikosch, T. (2009). *Handbook of Financial Time Series.* Berlin: Springer-Verlag Berlin Heidelberg.

Anderson, R., & Danthine, J.-P. (1980). Hedging and Joint Production: Theory and Illustrations. *The Journal of Finance, Vol. 25 (2)*, 487-498.

Ankirchner, S., Pigorsch, C., & Schweizer, N. (2013). *Estimating Residual Hedging Risk With Least-Squares Monte Carlo.* Technical report, University of Bonn.

Asteriou, D., & Hall, S. (2011). *Applied Econometrics 2ed.* New York: Palgrave MacMillan.

Baillie, R., & Myers, R. (1991). Bivariate GARCH Estimation of the Optimal Commodity Futures Hedge. *Journal of Applied Economics, Vol. 6.*, 109-124.

Bauwens, L., Laurent, S., & Rombouts, J. V. (2006). Multivariate GARCH Models: A Survey. *Journal of Applied Econometrics, Vol. 21.*, 79-109.

Benet, B. A. (1992). Hedge Period Length and Ex-Ante Future Hedging Effectiveness: The Case of Foreign-Exchange Risk Cross Hedges. *Journal of Futures Markets, Vol.12.*, 163-175.

Berk, J., & DeMarzo, P. (2014). *Corporate Finance.* Boston: Pearson Education Inc.

Bessembinder, H., & Lemmon, M. L. (2002). Equilibrium Pricing and Optimal Hedging in Electricity Forward Markets. *The Journal of Finance, Vol. 57 (3)*, 1347-1382.

Bodie, Z., Kane, A., & Marcus, A. J. (2011). *Investments.* McGraw-Hill.

Bollerslev, T. (1986). Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics, Vol. 31*, 307-327.

Bollerslev, T. (1990). Modelling the Coherence in Short-run Nominal Exchange Rates: A Multivariate Generalized ARCH Model. *The Review of Economics and Statistics, Vol. 72 (3)*, 498-505.

Botterud, A., Kristiansen, T., & Ilic, M. D. (2009). The Relationship Between Spot and Futures Prices in the Nord Pool Electricity Market. *Energy Economics, Vol. 32.*, 967–978.

Brooks, C. (2008). *Introductory Econometrics for Finance.* New York: Cambridge University Press.

Brooks, C., & Burke, S. P. (2003). Information Criteria for GARCH Model Selection. *The European Journal of Finance, Vol. 9 (6)*, 557-580.

Brooks, C., Henry, O., & Persand, G. (2002). The Effect of Asymmetries on Optimal Hedge Ratios. *The Journal of Business, Vol. 75 (2)*, 333-352.

Brown, G., Crabb, P., & Haushalter, D. (2006). Are Firms Successful at Selective Hedging? *Journal of Business, Vol. 79 (6)*, 2925-2950.

Burns, P. (2002). Robustness of the Ljung-Box Test and Its Rank Equivalent.

Byström, H. (2003). The Hedging Performance of Electricity Futures on the Nordic Power Exchange. *Applied Economics, Vol 35 (1)*, 1-11.

Bøhren, Ø., Michalsen, D., & Norli, Ø. (2012). *Finans: Teori og praksis.* Bergen: Fagbokforlaget.

Chang, C.-L., McAleer, M., & Tansuchat, R. (2011). Crude oil hedging strategies using dynamic multivariate GARCH. *Energy Economics, Vol. 33 (5)*, 912-923.

Charnes, J. M., & Koch, P. (2003). Measuring Hedge Effectiveness for FAS 133 Compliance. *Journal of Applied Corporate Finance, Vol. 15 (4)*, 95-103.

Chevallier, J. (2010). The impact of Australian ETS news on wholesale spot electricit yprices: An exploratory analysis. *Energy Policy, Vol. 38*, 3910-3921.

Chiulli, R. M. (1999). *Quantitative Analysis: An Introduction.* Los Angeles: CRC Press.

Choudhry, M. (2007). *The Futures Bond Basis 2nd ed.* Chichester: John Wiley & Sons Ltd.

de Jong, A., De Roon, F., & Veld, C. (1997). Out-of-Sample Hedging Effectiveness of Currency Futures for Alternative Models and Hedging Strategies. *The Journal of Futures Markets, Vol. 17*, 817-837.

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

EA. (n.d.). *Papers.* Retrieved February 13, 2019, from EA-Energianalyse: http://www.ea-energianalyse.dk/papers/1667_Analysis%20of%20future%20electricity%20prices.pdf

Ederington, L. H. (1979). The Hedging Performance of the New Futures Markets. *The Journal of Finance, Vol. 34 (1)*, 157-170.

Ederington, L. H., & Salas, J. M. (2008). Minimum Variance Hedging When Spot Price are Partially Predictable. *Journal of Banking & Finance, Vol. 32.*, 654-663.

EEX. (2017, November 16). *EEX acquires 100% of Powernext*. Retrieved from EEX: https://www.eex.com/en/about/newsroom/news-detail/eex-acquires-100--of-powernext--/75784

Efron, B. (1979). Bootstrap Methods: Another Look at the Jackknife. *The Annuals of Statistics, Vol. 7 (1)*, 1-26.

EIA. (2002). *Derivatives and Risk Management in the Petroleum. Natural Gas, and Electricity Industries.* Washington DC: Energy Information Administration.

Enders, W. (2015). *Applied Econometrics Time Series.* New York: John Wiley & Sons Inc.

Energifakta Norge. (n.d.). *Kraftmarkedet*. Retrieved January 26, 2019, from Energifakta Norge: https://energifaktanorge.no/norsk-energiforsyning/kraftmarkedet/

Engle, R. (2002). Dynamic Conditional Correlation: A Simple Class of Multivariate Generalized Autoregressive Conditional Heteroskedasticity Models. *Journal of Business & Economics Statistics, Vol. 20 (3)*, 339-350.

Engle, R. F. (1982). Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdon Inflation. *Econometricia, Vol. 40 (4)*, 987-1008.

Engle, R. F., & Sheppard, K. (2001). Theoretical and Empirical properties of Dynamic Conditional Correlation Multivariate GARCH. *NBER Working Papers 8554, National Bureau of Economic Research, Inc.*

European Union. (n.d.). *EU Action*. Retrieved April 10, 2019, from European Union: https://ec.europa.eu/clima/policies/ets_en

Europex. (n.d.). *Association of European Energy Exchanges*. Retrieved March 10, 2019, from Nord Pool: https://www.europex.org/members/nord-pool/

Eurostat. (2015, February 9). *Statistics Explained - Glossary: In-sample vs. out-of-sample forecasts*. Retrieved May 9, 2019, from Eurostat: https://ec.europa.eu/eurostat/statistics-explained/index.php/Glossary:In-sample_vs._out-of-sample_forecasts

Fabozzi, F. J., Focardi, S. M., & Kolm, P. N. (2010). *Quantitative Equity Investing: Techniques and Strategies.* New Jersey: Wiley.

Fama, E. F., & French, K. R. (1987). Commodity Futures Prices: Some Evidence on Forecast Power, Premiums, and the Thory of Storage. *Journal of Business, Vol. 60.*, 55-74.

Fiorenzani, S. (2006). *Quantitative Methods for Electricity Trading and Risk Management: Advanced Mathematical and Statistical Methods for Energy Finance.* New York: Palgrave Macmillan Ltd.

Garcia, P., Roh, J., & Leuthold, R. (1995). Simulateneously Determined, Time-Varying Hedge Ratios in the Soybean Complex. *Applied Economics, Vol, 27 (12)*, 1127-1134.

Ghalanos, A. (2019). *rmgarch: Multivariate GARCH models.* Retrieved from R package version 1.3-6.: https://cran.r-project.org/web/packages/rmgarch/vignettes/The_rmgarch_models.pdf

Ghosh, A. (1993). Hedging with Stock Index Futures: Estimation and Forecasting with Error Correction Model. *The Journal of Futures Markets, Vol. 13 (7)*, 743-752.

Gjolberg, O., & Brattestad, T.-L. (2011). The Biased Short-Term Futures Price at Nord Pool: Can it Really Be a Risk Premium? *The Journal of Energy Markets, Vol. 4 (1)*, 3-19.

Grossman, S. J., & Shiller, R. J. (1981). The Determinants of the Variability of Stock Market Prices. *American Economic Review, Vol. 71*, 222-227.

Haldrup, N., & Jansson, M. (2005). Improving Size and Power in the Unit Root Testing. *Aarhus University Economics Paper, Vol. 2*.

Hanly, J., Morales, L., & Cassells, D. (2017). The Efficacy of Financial Futures as a Hedging Tool in Electricity Markets. *International Journal of Finance & Economics, Vol. 23 (1)*, 29-40.

Huisman, R., & Kilic, M. (2012). Electricity Futures Prices: Indirect Storability, Expectations, and Risk Premiums. *The Journal of Energy Economics, Vol. 34 (4)*, 892-898.

Huisman, R., Michels, D., & Westgaard, S. (2014). Hydro Reservoir Levels and Power Price Dynamics. Empirical Insight on the Nonlinear Influence of Fuel and Emission Cost on Nord Pool Day-Ahead Electricity Prices. *Journal of Energy and Development, Vol. 40 (2)*, 149-187.

Hull, J. C. (2012). *Options, Futures, and Other Derivatives. Eight Edition.* Prentice Hall.

Håkonsen, G. Q. (2019, April 7). KIKS market situation. (B. Kjevik, Interviewer)

Isogai, T. (2015). *An empirical Study of the Dynamic Correlation of Japanese Stock Returns.* Tokyo: Bank of Japan.

Jani, P. N. (2014). *Business Statistics - Theory and Applications.* New Delhi: PHI Learning.

Jarque, C. M., & Bera, A. K. (1980). Efficient Tests for Normality, Homoscedasticity and Serial Independence of Regression Residuals. *Economics Letters* , 255-259.

Johnson, L. L. (1960). The Thory of Hedging and Speculation in Commodity Futures. *The review of Economic Studies, Vol. 27 (3)*, 139-151.

Jovanovic, S. (2014). *Hedging Commodities: A Practical Guide to Hedging Strategies with Futures and Options.* Petersfield: Harriman House Ltd.

Karagiannopoulos, L. (2018, September 13). *Nordic power trader's loss costs Nasdaq and members 114 million euros*. Retrieved from Reuters: https://www.reuters.com/article/us-nordic-power-nasdaq/nordic-power-traders-loss-costs-nasdaq-and-members-114-million-euros-idUSKCN1LT28G

Kavussanos, M., & Nomikos, N. (2000). Futures Hedging when the Structure of the Underlying Asset Changes: The Case of the BIFFEX Contract. *The Journal of Futures Markets, Vol. 20 (8)*, 775-801.

Koekebakker, S., & Ollmar, F. (2001). Forward Curve Dynamics in the Nordic Electricty Market. *Agder University College & Norwegian School of Economics and Business Administration*.

Konsesjonskraft IKS. (2017). *Rapport krafthandel desember.* Retrieved from Konsesjonskraft IKS: http://konsesjonskraft.no/media/1509/rapport-krafthandel-desember-2017.pdf

Konsesjonskraft IKS. (n.d.). *KIKS*. Retrieved January 25, 2019, from Konsesjonskraft IKS: http://konsesjonskraft.no/

Kroner, K. F., & Sultan, J. (1993). Time-Varying Distributions and Dynamic Hedging with Foreign Currency Futures. *Journal of Financial and Quantitative Analysis, Vol. 28 (4)*, 535-551.

Lee, J., & Zhang, Y. (2009). Evidence on Normal Backwardation and Forecasting Theory in Futures Markets. *Journal of Derivatives & Hedge Funds, Vol. 15 (2)*, 158-170.

Lien, D.-H. D. (1996). The Effect of the Cointegrating Relationship on Futures Hedging: A Note. *The Journal of Futures Markets, Vol. 16 (7)*, 773-780.

Lin, J.-L., & Granger, C. (1994). Forecastin From Non-Linear Models in Practice. *Jornal of Forecasting, Vol. 13 (1)*, 1-9.

Lindahl, M. (1989). Measuring Hedging Effectiveness With R^2: A Note. *The Journal of Futures Markets, Vol. 9 (5)*, 469-475.

Lovdata. (1917, December 14). *Vassdragsreguleringsloven*. Retrieved from Lovdata: https://lovdata.no/dokument/NL/lov/1917-12-14-17

Malliaris, A. G., & Urrutia, J. L. (1991). The Impact of the Lengths of Estimation Periods and Hedging Horizons on the Effectiveness of a Hedge: Evidence from Foreign Currency Futures. *The Journal of Futures Markets, Vol. 11 (3)*, 271-289.

Masteika, S., Rutkauskas, A. V., & Alexander, J. A. (2012). Continuous futures data series for back testing and technical analysis. *IPEDR, Vol. 29*, 265-269.

Medel, C., & Salgado, S. (2012). Does BIC Estimate and Forecast Better than AIC? *Revista de Analisis Economico, Vol. 28 (1)*, 48-64.

Miffre, J. (2004). Conditional OLS Minimum Variance Hedge Ratios. *Journal of Futures Markets, Vol. 24 (10)*, 945-964.

Moon, G.-H., Yu, W.-C., & Hong, C.-H. (2009). Dynamic Hedging Performance with the Evaluation of Mulitvariate GARCH Models: Evidence From KOSTAR Index Futures. *Applied Economics Letters, Vol. 16 (9)*, 913-919.

Moschini, G., & Myers, R. J. (2002). Testing for constant hedge ratios in commodity markets: a multivariate GARCH approach. *Journal of Empirical Finance, Vol. 9*, 589-603.

Myers, R. J. (1991). Estimating Time-Varying Optimal Hedge Ratios on Futures Markets. *Journal of Futures Markets, Vol. 11 (1)*, 39-53.

Mäntysaari, P. (2015). *EU Electricity Trade Law*. New York: Springer.

Nasdaq. (2008, October 21). *Investor Relations.* Retrieved from Nasdaq: http://ir.nasdaq.com/static-files/fa4fc30a-f049-4b70-b0e7-3eb0e21890f3

Nasdaq. (2016). *Nordic Power Products.* Retrieved from Nasdaq: https://business.nasdaq.com/media/nordic-power-fs_tcm5044-53598.pdf

Nasdaq. (n.d.-a). *Power Futures*. Retrieved March 1, 2019, from Nasdaq: https://business.nasdaq.com/trade/commodities/products/power-derivatives/power-futures.html

Nasdaq. (n.d.-b). *Power Derivatives*. Retrieved March 1, 2019, from Nasdaq: https://business.nasdaq.com/trade/commodities/products/power-derivatives/index.html

Nasdaq. (n.d.-c). *Transactions costs*. Retrieved March 11, 2019, from Nasdaq: https://www.nasdaq.com/investing/glossary/t/transaction-costs

Nasdaq. (n.d.-d). *Nasdaq Clearing*. Retrieved March 4, 2019, from Nasdaq: https://business.nasdaq.com/updates-on-the-Nasdaq-Clearing-Member-Default/index.html

Nasdaq OMX. (2003, February 7). *Changes in the Nord Pool financial product structure.* Retrieved from Nasdaq: https://cns.omxgroup.com/cdsPublic/viewDisclosure.action?disclosureId=405740&messageId=488873

Nasdaq OMX. (2018, June 11). *Contract Specifications - Commodity Derivatives.* Retrieved from Nasdaq: http://www.nasdaqomx.com/digitalAssets/109/109228_180611-joint-appendix-2---contract-specifications.pdf

Nasdaq OMX. (n.d.). *Indexes - Settlement value.* Retrieved February 5, 2019, from Nasdaq: https://indexes.nasdaqomx.com/docs/SettlementValueGIW.pdf

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

Nord Pool. (n.d.-a). *Global assets.* Retrieved February 11, 2019, from Nord Pool: https://www.nordpoolgroup.com/globalassets/download-center/rules-and-regulations/the-nordic-electricity-exchange-and-the-nordic-model-for-a-liberalized-electricity-market.pdf

Nord Pool. (n.d.-b). *About Us*. Retrieved February 23, 2019, from Nord Pool: https://www.nordpoolgroup.com/About-us/

Nord Pool. (n.d.-c). *Trading*. Retrieved March 11, 2019, from Nord Pool: https://www.nordpoolgroup.com/trading/

Nord Pool. (n.d.-d). *The Power Market*. Retrieved February 5, 2019, from Nord Pool: https://www.nordpoolgroup.com/the-power-market/

Nord Pool. (n.d.-e). *Price Calculation*. Retrieved April 29, 2019, from Nord Pool: https://www.nordpoolgroup.com/trading/Day-ahead-trading/Price-calculation/

Nord Pool. (n.d.-f). *The Nordic Electricity Exchange and the Nordic Model for a Liberalized Electricity Market.* Retrieved from Nord Pool: https://www.nordpoolgroup.com/globalassets/download-center/rules-and-regulations/the-nordic-electricity-exchange-and-the-nordic-model-for-a-liberalized-electricity-market.pdf

Nord Pool. (n.d.-g). *History*. Retrieved March 3, 2019, from Nord Pool: https://www.nordpoolgroup.com/About-us/History/

Olje- og energidepartementet. (2006). *Kraftmarkedet.* Retrieved from Olje- og energidepartementet: (https://www.regjeringen.no/globalassets/upload/kilde/oed/bro/2000/0002/ddd/pdfv/110237-kap._7.pdf).

Olje- og energidepartementet. (2008). *Kraftmarkedet.* Retrieved from Olje- og energidepartementet: https://www.regjeringen.no/globalassets/upload/oed/pdf_filer/faktaheftet/evfakta08/evfakta08_kap07_no.pdf

Olje- og energidepartementet. (2015, January). *Energi- og vannressurser i Norge.* Retrieved from Olje- og energidepartementet: https://www.regjeringen.no/contentassets/fd89d9e2c39a4ac2b9c9a95bf156089a/1108774830_897155_fakta_energi-vannressurser_2015_nett.pdf

Olje- og energidepartementet. (n.d.). *Global Assets.* Retrieved February 5, 2019, from Olje- og energidepartementet: (https://www.regjeringen.no/globalassets/upload/kilde/oed/bro/2000/0002/ddd/pdfv/110237-kap._7.pdf).

Park, T. H., & Switzer, L. N. (1995). Bivariate GARCH Estimation of Optimal Hedge Ratios for Stock Index Futures: a note. *The Journal of Futures Markets, Vol. 15 (1)*, 61-67.

Poomimars, P., Cadle, J., & Theobald, M. (2003). Futures Hedging Using Dynamic Models of the Variance/Covariance Structure. *The Journal of futures Markets, Vol. 23 (3)*, 241-260.

Posner, B. (2018). *PennState*. Retrieved from The Fundamentals of Electricity Markets: https://www.e-education.psu.edu/ebf200/node/151

R Core Team. (2013). R: A language and environment for statistical computing. Vienna, Austria. Retrieved from http://R-project.org/

Redl, C., & Bunn, D. W. (2013). Determinants of premium in forward contracts. *Journal of Regulatory Economics, Vol. 43 (90)*, 90-111.

Rohatgi, V., & Ehsanes Saleh, A. (2015). *An Introduction to Probability and Statistics.* New Jersey: Wiley.

Sagmoen, I., & Jordheim, H. (2018, September 19). *Kraftselskapene: Børsen trenger en som Einar Aas*. Retrieved from E24: https://e24.no/boers-og-finans/kraftproduksjon/kraftselskapene-mener-boersen-trenger-en-som-einar-aas/24443714

Samuelson, P. (1965). Proof that Properly Anticipated Prices Fluctuate Randomly. *Industrial Management Review, Vol. 6*, 41-49.

Sanda, G., Olsen, E., & Fleten, S.-E. (2013). Selective Hedging in Hydro-based Electricity Companies. *Energy Economics, Vol. 40*, 326-338.

Simonsen, I. (2005). Volatility of power markets. *Physica A*, 10-20.

Skantze, P., & Ilic, M. (2001). *Valuation, Hedging and Speculation in Competitive Electricity Markets: A Fundamental Approach.* Boston: Kluwer Academic Publishers.

Skjevrak, S. (2019, March 28). Futures and forwards on the Nordic power exchange. (B. Kjevik, Interviewer)

Statnett. (2018, November 12). *Regulerkraftmarkedet*. Retrieved from Statnett: https://www.statnett.no/for-aktorer-i-kraftbransjen/systemansvaret/kraftmarkedet/reservemarkeder/tertiarreserver/regulerkraftmarkedet/

Stock, J. H. (1994). Unit Roots, Structural Breaks and Trends. In R. F. Engle, & D. L. McFadden, *Handbook of Econometrics, Vol. 4* (pp. 2739-2841). New York: North Holland.

Stock, J. H., & Watson, M. W. (2015). *Introduction to Econometrics.* Harlow: Pearson.

Stoll, H., & Whaley, R. (1993). *Futures and Options: Theory and Applications.* Cincinnati: South-Western Publishing Co.

Torró, H. (2009, November 29). *Munich Personal RePEc Archive.* Retrieved from Assessing the Influence of Spot Price Predictability on Electricity Futures Hedging: https://mpra.ub.uni-muenchen.de/18892/

Tuckman, B. (2002). *Fixed Income Securities: Tools for Today's Markets 2nd ed.* Hoboken, New Jersey: John Wiley & Sons, Inc.

Vehviläinen, I., & Keppo, J. (2003). Managing Electricity Market Price Risk. *European Journal of Operational Research, Vol. 145*, 136-147.

Wang, Y.-H., Yeh, T.-J., & Chuang, S.-L. (2015). Hedging effeectiveness of the Hedged Portfolio: The Expected Utility Maximization Subject to the Value-at-Risk Approach. *Applied Economics, Vol. 47 (20)*, 2040-2052.

Weron, R., & Zator, M. (2014). Revisiting the relationship between spot and futures prices in the Nord. *Energy Economics, Vol. 44,* 178-190.

Wimschulte, J. (2010). The Futures and Forward Price Differential in the Nordic Electricity Market. *Energy Policy, Vol. 38,* 4731-4733.

Wooldridge, J. M. (2012). *Introductory Econometrics - A Modern Approach.* Cengage Learning.

Working, H. (1953). Futures Trading and Hedging. *The American Economic Review, Vol. 43 (3)*, 314-343.

Yang, J., & Awoke, T. O. (2003). Assets Storability and Hedging Effectiveness in Commodity Futures Markets. *Applied Economics Letters, Vol. 10*, 487-491.

Zanotti, G., Gabbi, G., & Geranio, M. (2009). Hedging with futures: Efficacy of GARCH correlation models to Euroepan electricity markets. *Journal of International Financial Markets, Institutions & Money, Vol 20 (2)*, 135-148.

Zhou, J. (2016). Hedging performance of REIT index futures: A comparison of alternative hedge ratio estimation methods. *Economic Modelling, Vol. 52*, 690-698.

12. Appendix

Appendix 1: Two-Sample t-Tests of Bootstrapped Distributions

This appendix covers the two-sided two-sample t-tests for the mean values of the bootstrapped (1000x) variance and VaR distributions as desribed in sub-section 6.4.

Appendix 1.1. Hedging Models vs. Unhedged Portfolio

Testing whether the variance and VaR of each hedging model are significantly less compared to that of the unhedged spot portfolio.

		Monthly futures				**Quarterly futures**			
		Variance		VaR		Variance		VaR	
Model	Period	*t*-stat	p-value	*t*-stat	p-value	*t*-stat	p-value	*t*-stat	p-value
Naïve	Full period	-20.494	0.000	28.005	0.000	-11.258	0.000	15.925	0.000
	Sub-period 1	-17.264	0.000	28.419	0.000	-8.016	0.000	18.366	0.000
	Sub-period 2	-12.704	0.000	16.536	0.000	-8.978	0.000	8.091	0.000
	Sub-period 3	-14.057	0.000	21.227	0.000	-5.104	0.000	11.081	0.000
	Sub-period 4	-0.612	0.270	-0.340	0.633	-1.478	0.069	-3.070	0.999
OLS	Full period	-20.776	0.000	29.048	0.000	-11.902	0.000	15.438	0.000
	Sub-period 1	-28.001	0.000	33.678	0.000	-15.530	0.000	20.871	0.000
	Sub-period 2	-17.695	0.000	20.330	0.000	-11.334	0.000	11.095	0.000
	Sub-period 3	-13.809	0.000	21.861	0.000	-3.163	0.001	9.762	0.000
	Sub-period 4	-0.612	0.056	-0.338	0.632	-0.697	0.243	-2.054	-0.979
CCC	Full period	-25.416	0.000	33.116	0.000	-20.125	0.000	22.708	0.000
	Sub-period 1	-29.587	0.000	38.486	0.000	-22.285	0.000	26.493	0.000
	Sub-period 2	-23.879	0.000	26.857	0.000	-17.439	0.000	20.344	0.000
	Sub-period 3	-1.040	0.149	10.147	0.000	-0.283	0.389	7.042	0.000
	Sub-period 4	-6.924	0.000	3.367	0.000	-2.482	0.007	0.810	0.209
DCC	Full period	-27.157	0.000	34.010	0.000	-20.528	0.000	24.098	0.000
	Sub-period 1	-25.287	0.000	33.932	0.000	-21.499	0.000	24.052	0.000
	Sub-period 2	-24.735	0.000	29.559	0.000	-21.044	0.000	21.169	0.000
	Sub-period 3	-0.527	0.299	10.501	0.000	0.664	0.747	6.407	0.000
	Sub-period 4	-6.403	0.000	3.846	0.000	-2.858	0.002	0.238	0.406

Appendix 1.2. Best Dynamic Hedge vs. Best Static Hedge

Testing whether the best-performing dynamic model in each period has a statistically significantly lower variance and VaR compared to the best-performing static hedging model.

		Variance			VaR		
Futures	Period	Models tested	t-stat	p-value	Models tested	t-stat	p-value
Monthly	Full period	DCC vs. OLS	-5.710	0.000	DCC vs. naïve	4.853	0.000
	Sub-period 1	CCC vs. OLS	-1.514	0.065	CCC vs. OLS	4.134	0.000
	Sub-period 2	DCC vs. OLS	-6.799	0.000	DCC vs. OLS	9.771	0.000
	Sub-period 3	CCC vs. OLS	13.402	1.000	DCC vs. naïve	-11.355	1.000
	Sub-period 4	CCC vs. OLS	-5.433	0.000	DCC vs. OLS	4.194	0.000
Quarterly	Full period	DCC vs. OLS	-8.151	0.000	DCC vs. naïve	7.930	0.000
	Sub-period 1	CCC vs. OLS	-6.260	0.000	CCC vs. OLS	4.757	0.000
	Sub-period 2	DCC vs. OLS	-9.794	0.000	DCC vs. OLS	10.321	0.000
	Sub-period 3	CCC vs. OLS	2.859	0.998	CCC vs. OLS	-3.118	0.999
	Sub-period 4	CCC vs. OLS	-2.858	0.002	DCC vs. OLS	2.089	0.018

Appendix 1.3. DCC-GARCH vs. CCC-GARCH

Testing whether the DCC-GARCH has a lower variance and VaR compared to that of the CCC-GARCH model.

		Variance		VaR	
Futures	Period	t-stat	p-value	t-stat	p-value
Monthly	Full in-sample period	-2.087	0.019	-0.320	0.626
	Sub-period 1	4.152	1.000	-4.460	0.000
	Sub-period 2	-0.696	0.243	2.850	0.998
	Sub-period 3	0.536	0.704	0.505	0.693
	Sub-period 4	0.318	0.625	0.465	0.679
Quarterly	Full in-sample period	0.115	0.546	1.064	0.144
	Sub-period 1	1.001	0.842	-2.080	0.981
	Sub-period 2	-3.667	0.000	1.025	0.153
	Sub-period 3	0.927	0.827	-0.776	0.781
	Sub-period 4	-0.392	0.348	-0.528	0.701

Appendix 1.4. Monthly vs. Quarterly Full In-Sample Period

Testing whether the variance and VaR from the hedging models with monthly futures are significantly lower than those obtained with quarterly futures

	Variance		**VaR**	
Full period	t-stat	p-value	t-stat	p-value
Naïve	-9.622	0.000	11.543	0.000
OLS	-9.105	0.000	12.250	0.000
CCC	-5.122	0.000	9.495	0.000
DCC	-7.448	0.000	9.084	0.000

Appendix 2: R Code for Multivariate GARCH Modeling and Forecasting

The package used for the GARCH modeling and forecasting is the ‘rmgarch’ package by Ghalanos (2019)

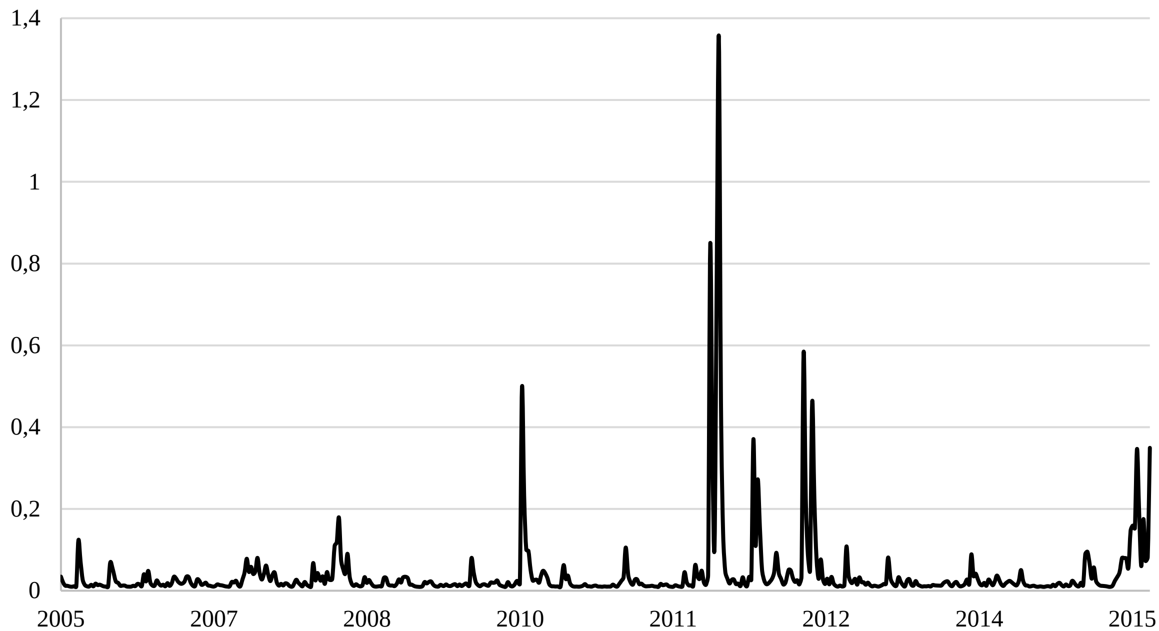
```
# Load the packages
lapply(c("dynlm", "readxl", "rmgarch", "graphics", "quantmod"), library,
character.only = TRUE)

# Reading in the data (return series)
data <- read_excel(“spot_futures_data.xlsx")
data.xts <- xts(weeklytest[,2:3], order.by = as.Date(data$date, "%m/%d/%Y"))

# Specifying the multivariate DCC-GARCH model
spec1 = ugarchspec(mean.model = list(armaOrder = c(0,0)))
spec2 = ugarchspec(mean.model = list(armaOrder = c(1,1)))
uspec.n = multispec(c(spec1, spec2))
multf = multifit(uspec.n, data.xts)
spec1 = dccspec(uspec = uspec.n, dccOrder = c(1,1), distribution = 'mvnorm')
fit1 = dccfit(spec1, data = data.xts, fit.control = list(eval.se = TRUE), fit = multf)

# Specifying the multivariate CCC-GARCH model:
spec3 = cgarchspec(uspec = uspec.n)
fit2 = cgarchfit(spec2, data = data.xts, fit.control = list(eval.se = TRUE, stationarity =
TRUE, scale = FALSE))

# One-step-ahead forecasts with a rolling window for the conditional covariance matrix for
the multivariate GARCH models:
out.xts <- xts(matrix(NA, nrow=(nrow(data.xts)*0.25), ncol = 3), order.by =
index(data.xts[(nrow(data.xts)*0.75+1):nrow(data.xts),]))
colnames(out.xts) <- c("variance futures", "covariance", "variance spot")
xIndexEnd <- (nrow(data.xts)*0.75)

for (i in 1:(nrow(data.xts)*0.25)){
  multfTemp = multifit(uspec.n, data.xts[i:(xIndexEnd+i)])
  fitTemp = dccfit(spec1, data = data.xts[i:(xIndexEnd+i)],
fit.control = list(eval.se = TRUE), fit = multfTemp)
  dccfTemp <- dccforecast(fitTemp, n.ahead = 1)
  out.xts[i,1] <-  as.matrix(dccfTemp@mforecast$H[[1]])[1,1]
  out.xts[i,2] <-  as.matrix(dccfTemp@mforecast$H[[1]])[2,1]
  out.xts[i,3] <-  as.matrix(dccfTemp@mforecast$H[[1]])[4,1]
}
```

Appendix 3: Volatility and Correlation from the GARCH Models

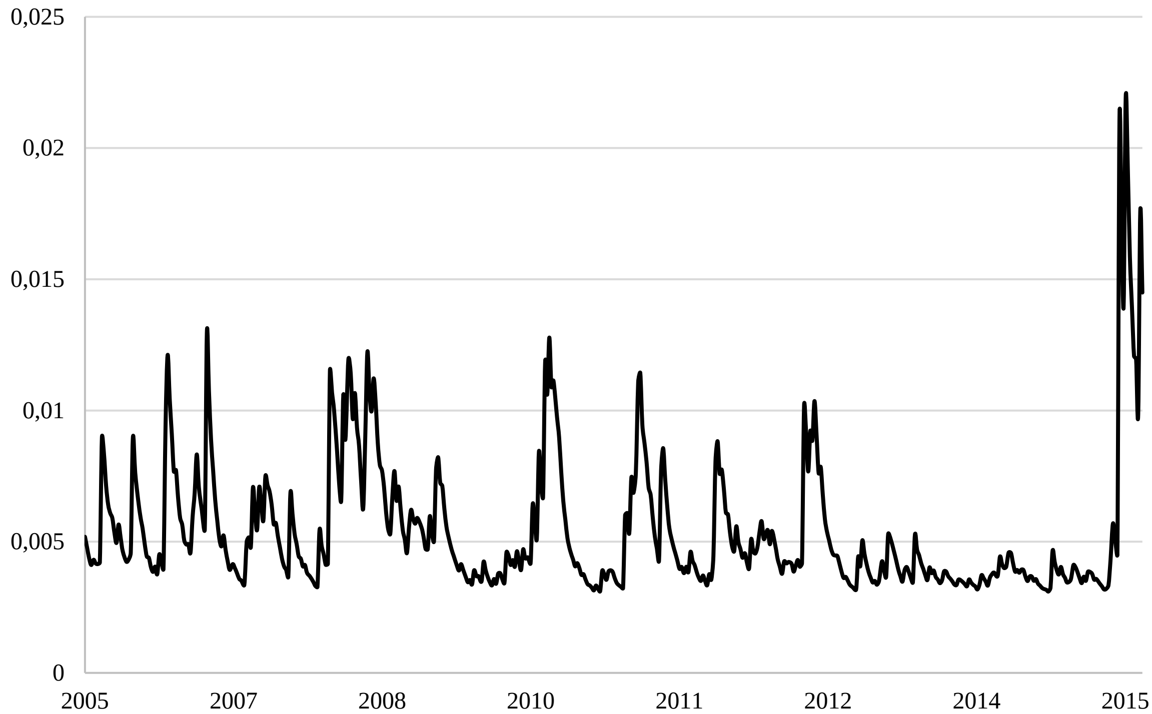
This appendix includes graphical representations of the estimated conditional volatility and correlation from the multivariate GARCH models.

Appendix 3.1. Conditional Volatility from Univariate GARCH Models – in-sample

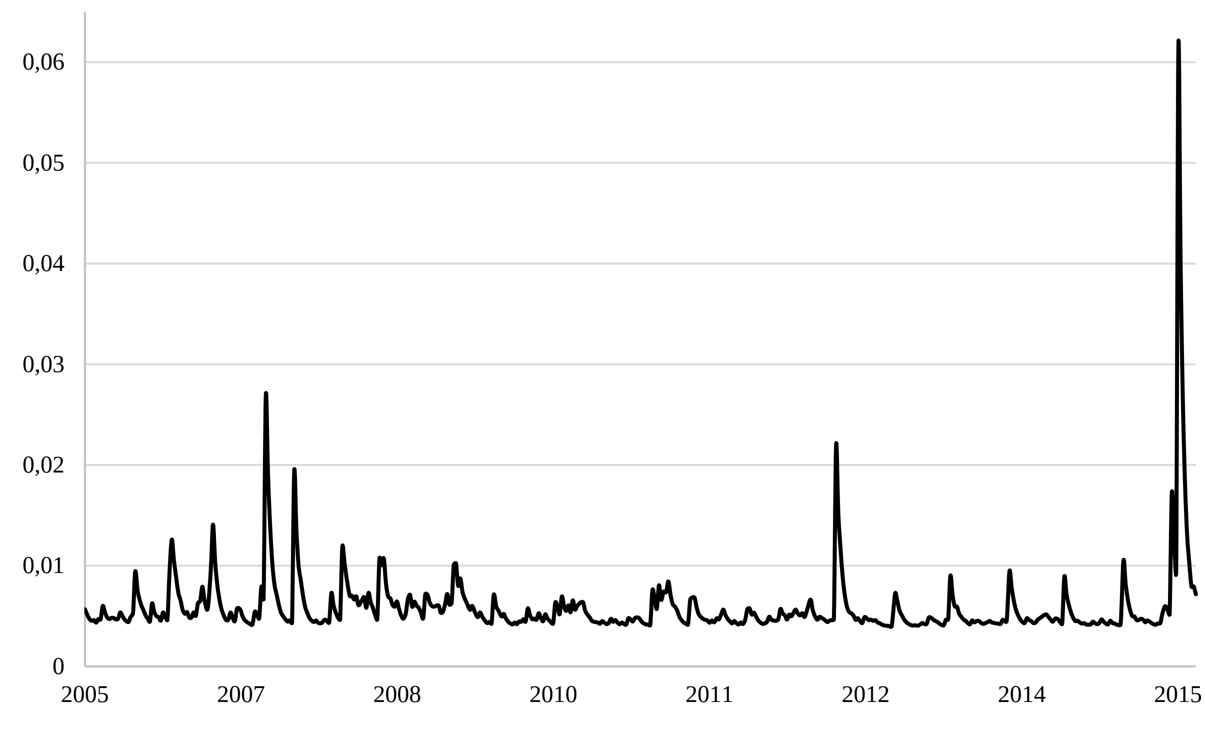
Appendix 3.1.1. Conditional Volatility Spot Returns



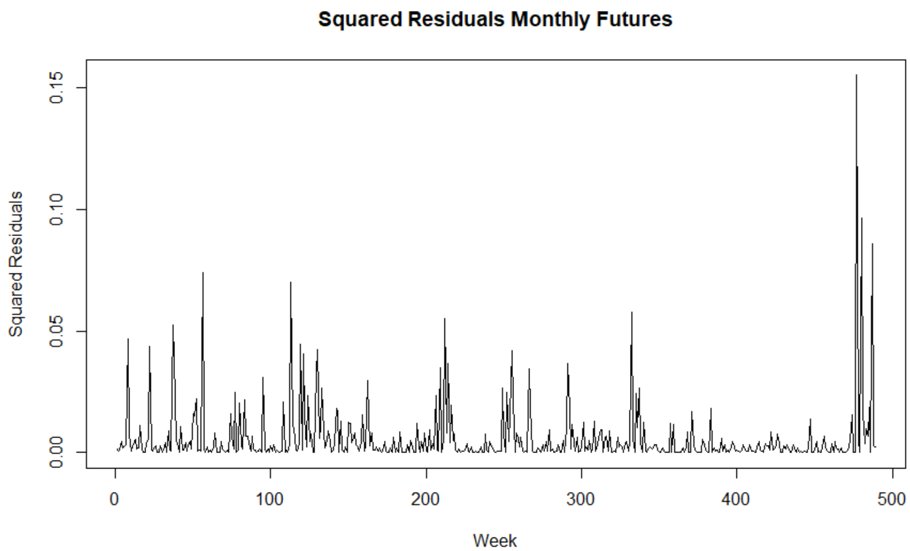
Appendix 3.1.2. Conditional Volatility - Monthly Futures Returns

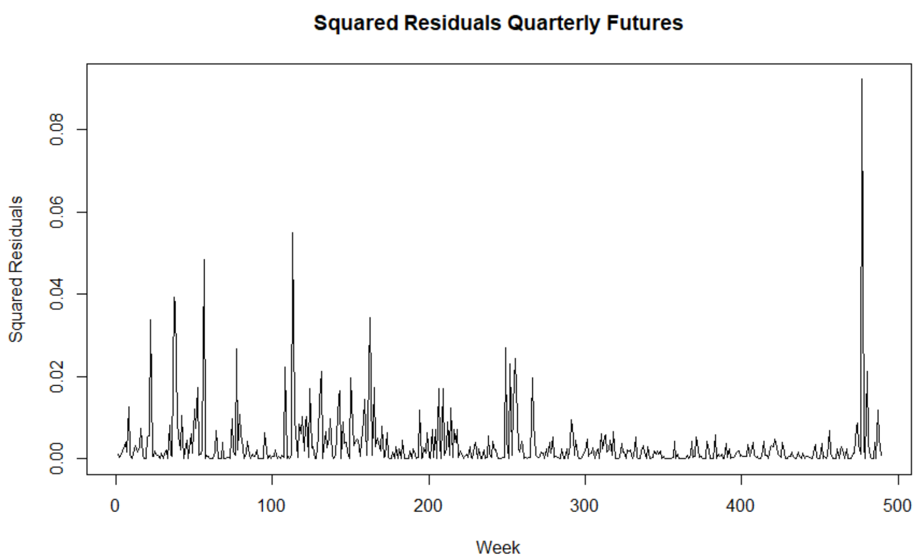


Appendix 3.1.3. Conditional Volatility - Quarterly Futures Returns



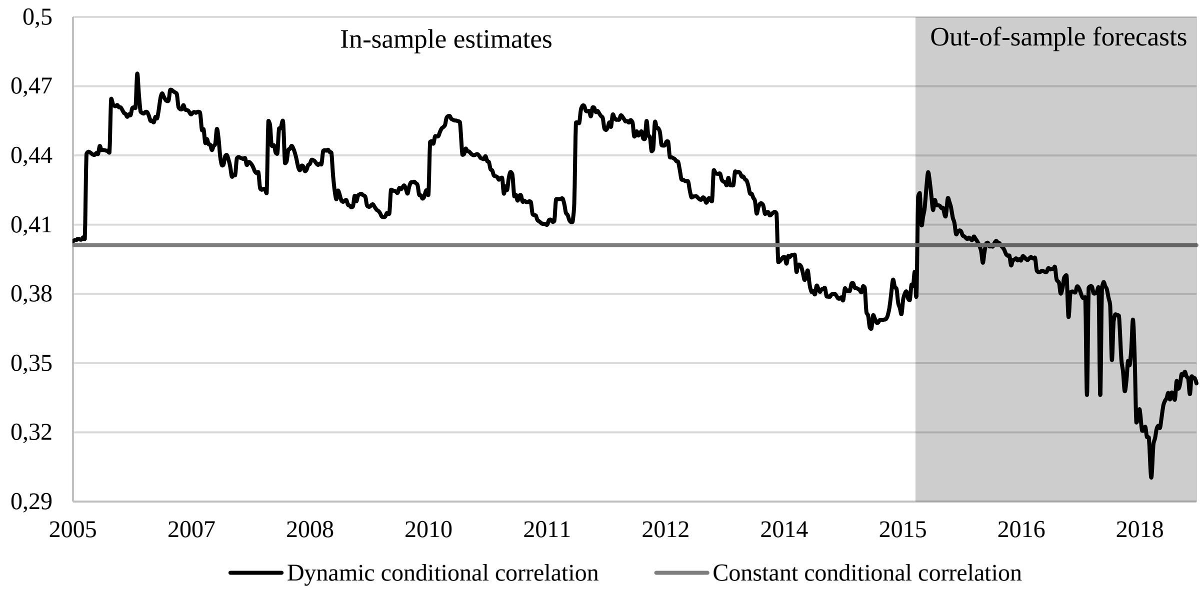
Appendix 3.2. Squared Residuals of the Conditional Mean Models – in-sample





Appendix 3.3. Conditional Correlations from Multivariate GARCH Models

Appendix 3.3.1. – Conditional correlations (monthly contracts)



Appendix 3.3.2 – Conditional correlations (quarterly contracts)