



PRICE CLUSTERING IN THE NORD POOL ELECTRICITY MARKET

Duarte José Araújo Pinto



Dissertation

Master in Finance



Supervised by

Júlio Fernando Seara Sequeira da Mota Lobão, PhD



2020

Price Clustering in the Nord Pool Electricity Market

Duarte José Araújo Pinto

Supervisor: Júlio Fernando Seara Sequeira da Mota Lobão, PhD

Master in Finance, School of Economics and Management, University of Porto

July 2020

The Author

Duarte José Araújo Pinto, born on March 14th, 1997, and natural from Vila Nova de Famalicão, Portugal, obtained his Bachelor's in Economics by Faculdade de Economia da Universidade do Porto. A rising interest in financial markets and the corporate world led him to pursue a Master's in Finance at Faculdade de Economia da Universidade do Porto. On a personal level, the study of Behavioural Economics and Behavioural Finance raises a particular interest, due to the implications these have in everyone's daily lives and the insights provided on how anyone can improve their cognitive reasoning.

Acknowledgments

From the entire family and FEP community, a special thanks must be given to my supervisor and professor Júlio Lobão for kindly introducing and guiding me through such an interesting topic, allowing to incorporate both my interests in the fields of Behavioural Finance and financial markets. Also, credits should be given to my family and friends, who were and continue to be an essential pillar to my growth as a person and as a professional.

Abstract

Utility markets in general, and energy markets in particular, involve a large number of stakeholders that depend on its efficiency and well-functioning. But similarly to other financial markets, as the number of participants grows the entire system becomes more complex, often leading to the appearance of inefficiencies such as price clustering. Observing a consistent growth, Nord Pool has become the reference electricity market within Europe, motivating the study of price clustering within this in order to assess its efficiency level and retrieve potential implications to its stakeholders.

The study of price clustering is included in the financial anomaly's literature, which has become increasingly more important among the academics. Price clustering can be defined as the tendency for prices to accumulate around specific values, and several studies have already been done regarding this phenomenon finding supportive evidence in a wide array of asset classes, that go from the stock market to the real estate and even cryptocurrencies. Nevertheless, research focused on utility markets is still relatively small in volume, thus making studies like the present important to develop this field within the academic literature.

To test the hypothesis of price clustering at Nord Pool as well as its nature, public prices available at the institution's website were used, comprehending hourly, daily and weekly timeframes. The Chi-squared goodness-of-fit test evidence the presence of price clustering in hourly prices for 11 out of 21 bidding areas, for a 90% confidence level, while the system price does not show statistically significant evidence that supports this phenomenon. The results from the multivariate regression applied to the bidding areas do not support in general the existent price clustering theories. Practical implications will depend on the perspective adopted, as this market involves a wide group of stakeholders with different interactions and levels of exposure.

Key-words: Nord Pool, price clustering, electricity market, market efficiency.

JEL-Codes: G12, G14, G40

Resumo

Mercados de *utilities* em geral, e mercados de energia em particular, envolvem um grande número de *stakeholders* que dependem da sua eficiência e bom funcionamento. Mas, de forma semelhante a outros mercados financeiros, à medida que o número de participantes cresce o sistema torna-se mais complexo, por vezes levando ao surgimento de ineficiências como *price clustering*. Observando um crescimento constante, o Nord Pool tornou-se no mercado de eletricidade de referência da Europa, motivando o estudo de *price clustering* dentro deste de forma a averiguar o seu nível de eficiência e retirar potenciais implicações para os seus *stakeholders*.

O estudo de *price clustering* está incluído na literatura das anomalias financeiras, que se tem tornado sucessivamente mais importante junto dos investigadores. *Price clustering* pode ser definido como a tendência para os preços de acumularem junto de certos números, e vários estudos já foram realizados relativamente a este fenómeno encontrando evidência que o suporta num vasto grupo de classes de ativos, desde o mercado acionista, mercado imobiliário e até criptomoedas. No entanto, pesquisa focada no mercado de *utilities* é ainda relativamente escassa em volume, por isso fazendo estudos como o presente importantes para desenvolver este campo dentro da literatura académica.

Para testar a hipótese de *price clustering* no Nord Pool bem como a sua natureza, preços públicos disponíveis no website da instituição foram usados, para períodos horários, diários e semanais. O teste do *Chi-squared goodness-of-fit* evidencia a presença de *price clustering* em preços diários para 11 das 21 áreas de licitação, para um intervalo de confiança de 90%, enquanto the o *system price* não apresenta evidência estatisticamente significativa que suporta este fenómeno. Os resultados da regressão multivariada aplicados não suportam em geral as teorias de *price clustering* existentes. Implicações práticas irão depender da perspetiva adotada, sendo que este mercado envolve um grande número de *stakeholders* com diferentes interações e níveis de exposição.

Palavras-chave: Nord Pool, clustering de preços, mercados de electricidade, eficiência de mercado

Códigos JEL: G12, G14, G40

Table of Contents

The Author.....	i
Acknowledgments	ii
Abstract.....	iii
Resumo	iv
List of Figures	vii
List of Tables.....	viii
1. Introduction	1
2. Literature Review.....	4
2.1. Nord Pool Structure	4
2.2. Price Dynamics of Nord Pool.....	6
2.3. Price Clustering in Financial Markets	12
2.3.1. Economic-based Explanations	13
2.3.2. Behavioural Explanations.....	14
2.3.3. Empirical Findings	16
3. Data Base and Methodology.....	19
3.1. Data Base	19
3.2. Methodology.....	20
4. Empirical Results and Analysis.....	24
4.1. Tests on Price Clustering Existence.....	24
a) Hourly Prices	24
b) Daily Prices	29
c) Weekly Prices.....	34

4.2. Test on Price Clustering Nature	40
5. Conclusion.....	48
Glossary	51
Appendix.....	52
References.....	56

List of Figures

Figure 2.1 - Nord Pool Bidding Areas map, by region	8
Figure 4.1 - Cumulative number of occurrences of each number, for Nord Pool day-ahead hourly prices of the entire sample	29
Figure 4.2 - Cumulative number of occurrences of each number, for Nord Pool day-ahead daily prices of the entire sample	34
Figure 4.3 - Cumulative number of occurrences of each number, for Nord Pool day-ahead weekly prices of the entire sample	39

List of Tables

Table 4.1 - Frequency distribution for the hundredth decimal number and whole digits of Nord Pool day-ahead market hourly prices.....	25
Table 4.2 - Frequency distribution for the hundredth decimal number and whole digits of Nord Pool day-ahead market daily prices.....	30
Table 4.3 - Frequency distribution for the hundredth decimal number and whole digits of Nord Pool day-ahead market weekly prices	36
Table 4.4 - Multivariate Regression applied to Nord Pool day-ahead hourly prices, without $\log(Volatility)$ as an explanatory variable.....	41
Table 4.5 - Multivariate Regression applied to Nord Pool day-ahead hourly prices, incorporating $\log(Volatility)$ as an explanatory variable.....	44

1. Introduction

At the time of write, the fields of Economics and Finance are divided into two divergent lines of thought, as the traditional knowledge which makes the foundation of such fields is being challenged by new findings for which the first cannot provide an answer. A multidisciplinary approach to these questions, relying namely upon insights from other social sciences such as Psychology, has created a new way of developing science, commonly mentioned as behavioural Economics and behavioural Finance. The strength of these has grown as research evolves, supported namely by findings made regarding anomalies in the financial markets, whether they are more abstract temporal anomalies (such as the January effect), or more concrete ones such as the excessively high premium demanded by investors of the stock market, not explained by any apparent risk factor.

Price clustering is precisely one of such anomalies. This phenomenon can be defined as the tendency for the price of an asset (stocks, commodities, currencies and others) to stick around some specific set of values, most often round numbers such as whole digits (XX.00) or half digits (XX.50). One can also observe a more specific type of clustering, namely numbers with the right digit equal to zero (XX.X0) or five (XX.X5) for example.

Extensive literature around this particular anomaly has already been developed, from which several explanatory hypothesis surged. Based in both economical or behavioural explanations, have already surged, with the most relevant ones being: Ball, Torous, and Tschoegl (1985) price resolution hypothesis, Harris (1991) negotiation hypothesis, and Christie and Schultz (1994) collusion hypothesis based in economic-incentives explanations; (4) Curcio and Goodhart (1991) attraction hypothesis which suggests that the origin of price clusters come from human biases.

Even though most of these hypotheses surged from the study of the American stock market, it is important to understand how results differ from the stock market to other markets. Evidence in favour of the existence of price clustering is already very extensive, and studies have been made for several types of assets. Apart from the two studies regarding the United States stock market (Christie, Harris, & Schultz, 1994; Harris, 1991), research has also been done regarding other exchanges (Hu, Jiang, McInish, & Zhou, 2017), the foreign exchange market (Sopranzetti & Datar, 2002), the real estate (Palmon, Smith, & Sopranzetti, 2004), commodities such as gold

(Ball et al., 1985), cryptocurrencies such as bitcoin (Urquhart, 2017), and even utility markets such as the Australian water market (Brooks, Harris, & Joymungul, 2013).

As it is self-described in the official website, «Nord Pool is Europe's leading power market and offers trading, clearing, settlement and associated services in both day-ahead and intraday markets across nine European countries» (Nord Pool, 2019) at the date of writing with 380 companies from 20 countries trading in the Nordic, Baltic, Germany and United Kingdom markets. Before this open market was created, electricity in the Nordic countries was supplied by the more traditional system of state-owned monopolies, that integrated both production and distribution levels of the vertical chain. Apart from the electricity physical market – Nord Pool Spot – investors can trade derivatives related to the Nord Pool power market via external markets (namely through NASDAQ OMX Commodities Europe). Thus, this exchange not only connects the production of electricity (which corresponds to the supply) with the distributors of electricity (the demand) in several European countries, but also allows investors to follow diverse strategies with exposure to the electricity market by using instruments such as derivatives.

Evidence supporting the hypothesis of clustering in Nord Pool prices can have a significant impact on the stakeholders, namely investors. As Sopranzetti and Datar (2002) state, the existence of price clustering can have two major consequences: (1) regarding the valuation of derivatives instruments, as traditional models assume that the underlying assets show a uniform distribution of prices; (2) investors and stakeholders whose risk hedging and portfolio management strategies depend on this electricity market. Hellström, Lundgren, and Yu (2012) also support such idea: «for electricity market participants, an understanding of electricity price dynamics is crucial for risk management, portfolio management and electricity derivative pricing». The authors continue, referring to the impact in the valuation of derivative instruments: «for example, since electricity is not a storable good the conventional derivative pricing models, e.g., based on the Black–Scholes model, are not appropriate. Instead, the valuation of electricity derivatives is dependent on models that properly describe the dynamics of the underlying electricity spot price» (Hellström et al., 2012).

Thus, additional contributions to this new behavioural economic paradigm are essential not only to strengthen the scientific research, but also to the development of new models that better represent the dynamics of a security, namely when this security is a derivative whose underlying

asset is subject to an anomaly as the price clustering. This work aims to first detect the existence of price clustering in the Nordic electricity market, and second to understand its nature (when and where such phenomenon exists) which should generate important insights for those who wish to incorporate this anomaly into valuation models. This study comes also to complement the growing literature around this financial phenomenon, which is still significantly scarce within the utility's markets in general, and inexistent at the time of writing for electricity markets such as Nord Pool.

For the upcoming empirical analysis, public information regarding prices and volume for the day-ahead market, available at the Nord Pool website, was used. In order to analyse the phenomena in both temporal and cross-sectional dimensions, every bidding area price will be considered for different frequencies: hourly prices for the entire year of 2019; daily prices from January 2017 till December 2019; and monthly prices also from January 2017 till December 2019.

From the price clustering tests applied, evidence found supports the existence of the phenomenon in several bidding areas, mostly at the hourly timeframes. Nevertheless, the special dynamics and characteristics of Nord Pool seem to lead to inconsistent results when the multivariate regression is applied, as no significant insights from the major price clustering theories within the literature support the findings.

The structure of the current dissertation will be as follows. After the current introductory part, the second chapter will involve the literature review that serves as support the upcoming empirical work, analysing more closely the structure of Nord Pool market, the price dynamics of this electricity market, and lastly the overall literature regarding price clustering, from the evidence found in multiple asset classes to explanatory theories already developed by relevant academics. The third chapter will discriminate the data and methodology that will be applied in the empirical exercise. In the fourth chapter, the empirical results from the price clustering tests and the multivariate regression will be analysed. In the fifth and last chapter, a conclusion is made taking into account both existent literature and empirical findings, as well as the positioning of the present study within the academic research.

2. Literature Review

The Nord Pool concept is still relatively unfamiliar to most countries outside the north of Europe. If asked to summarize the main objective/innovation of this system, the answer would inevitably include the word “integration”, since that is precisely what this electricity exchange seeks: the integration of the electrical market between the Nordic countries, with continuous expansion to other European countries.

The following literature review is divided into three major parts. To begin, further analysis of Nord Pool structure and characteristics will be made, with the objective to better understand how this market functions. At a second stage, the price dynamics of Nord Pool will be studied, in order to recognize how demand and supply interact, how the price is formed, and which are the main factors that affect this variable. Finally, the most relevant empirical work about price clustering will be analysed, aiming to detect eventual common characteristics between these findings and to serve as a basis for the following empirical work.

2.1. Nord Pool Structure

The creation of Nord Pool started way back in 1991, when the Norwegian parliament decided to deregulate the market for the trading of electrical energy, with the first stage of Nord Pool ASA resulting in 1996 from a fusion between the Swedish Transmission System Operator (TSO) known as Svenska Kraftnät and the Norwegian TSO called Statnett. Since then, the integration of the Nordic power market kept its growth at a fast pace, being currently expanding to other countries in central Europe. Financial markets have been taking advantage of technologic evolution and Nord Pool is no different, having introduced online trading platforms (both for the day-ahead and intraday markets) in 2015 and opening markets every hour and every day of the year. In the same year of 2015, the institution is appointed as a Nominated Electricity Market Operator (NEMO) by 10 European power markets - which have meanwhile grown to 15 – this way expanding its presence as an official institution. Apart from the day-ahead power trading available in these countries, Nord Pool also offers continuous intraday trading of power products with physical delivery within 13 countries. In addition to these, Nord Pool also provides several services, such as the development of the day-ahead and intraday systems for CropheX and IBEX

(two other electricity market operators) to run the power markets of Croatia and Bulgaria, respectively.

To better understand how Nord Pool system is built, we must first distinguish the two phases that compose every power market, from its producer to its end-user (individuals or companies). In the initial phase, the producers of energy are responsible to generate electricity through a source of energy (such as wind power, hydropower, nuclear reactions or condensing plants), that later provide to a supplier who redirects this power to the final consumer via the distribution grid. Precisely, at this second stage is where the distributor company interacts with the final consumer, to which are applied the traditional costs such as distribution fees and taxes specific to each situation. It is important to notice that this transaction of power is not perfect and faces challenges such as ramping (i.e. change in power flow/generation from one time unit to the following) and losses of energy along the distribution grid (inevitably leading to a difference between the volume sold by a bidding area and the volume received by another).

In practice, the main difference from Nord Pool to the traditional energy market is that, instead of having a small set of monopolistic companies that dominate the vertical chain – producing and distributing electricity to the final consumer – the initial stage previously described is made through a decentralized exchange, with both producers and suppliers vulnerable to changes in the supply and demand, and with the counterparty risk covered by Nord Pool who acts as a clearing house that guarantees the settlement of the contracts and delivery of power. This way, even though energy providers can set their desired price to the final consumer (which can choose among several available energy providers), all of them buy the electrical energy via Nord Pool or directly from a producer.

The Nord Pool Spot allows for the physical delivery of power, both on an intraday or a day-ahead basis. Other financial markets are also connected to this, such as the NASDAQ OMX Commodities Europe market, where both institutional and individual investors can trade several derivative instruments such as futures and options, both for hedging strategies and speculative investments. Note that these two markets are separated, whereas the financial market does not belong to Nord Pool. Regarding the Nord Pool Spot market, we can further divide it into two sub-markets – the day-ahead market or Elspot market, and the intraday market or Elbas market.

Overall, there are six main classes of agents that interact with each other through Nord Pool. On a higher level, we have the Transmission System Operators (TSOs) responsible for the stability and security of supply in each respective bidding area. In the first phase of the market, there are more than 370 electricity producers spread throughout the Nordic and Baltic countries, each one with a different source of energy and consequently different production costs (both fixed and variable). At the next level of the chain there are about 500 distributors, each one responsible for a single geographical area where they ensure that power reaches the end-users, comprehending the central grids and transmission network that connect the power plants with the final consumer. The higher level of competition occurs at the level of electricity suppliers, with about 380 companies buying power directly from a producer or through Nord Pool, and reselling this electricity to households and companies. Placed lastly are the end-users of electricity, which have only one choice regarding the distributor but can choose from a big range of suppliers. Apart from having to pay an amount for the power transmission to the distributor and taxes, the final consumer might choose between different power contracts which can vary for example between a fixed price contract and a market price contract. In the situation of Nord Pool, there are still the traders and brokers who form the financial market. Similarly to other financial markets, brokers serve as mere intermediaries for transactions, and traders involve all those agents which own the power while the trading process is not concluded, often acquiring power from a producer and selling it to a supplier or acquiring it from one supplier to sell to another.

2.2. Price Dynamics of Nord Pool

Despite the differences and the special characteristics of energy, Nord Pool establishes prices using the process known as marginal price setting, also used by other energy markets. Nevertheless, as electricity cannot be stored and must be delivered at the precise moment the final consumer needs it, demand presents high levels of inelasticity. To the supply side, the production of electricity is also more demanding than other types of energy and commodities since it implies high levels of variable costs (depending on the source of energy used). These characteristics justify the usefulness of Nord Pool as a system that establishes the equilibrium between supply and demand, since only by facing the cost of producing one MW/h with the

price the customer is willing to pay for the final MW/h required to satisfy demand it is guaranteed that the price formation process is economically efficient for both parties.

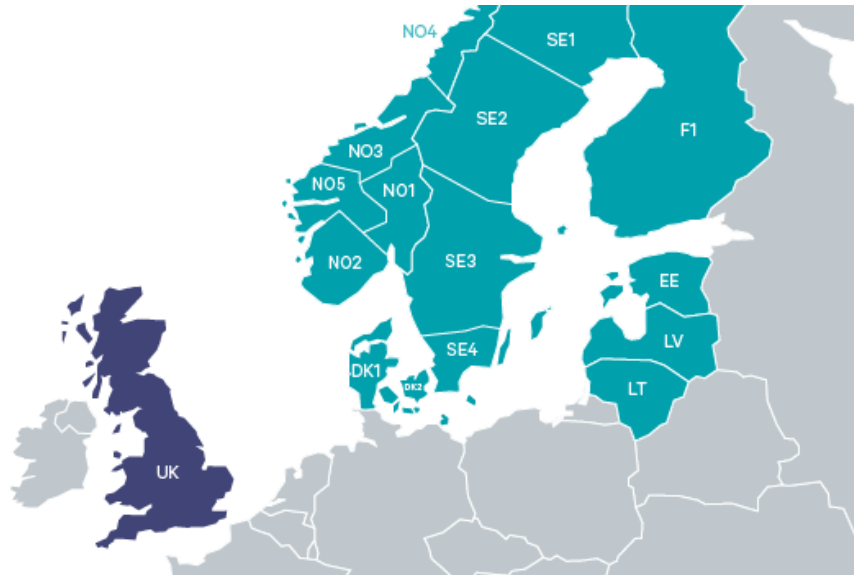
For the market participants, in the day-ahead trading platform Nord Pool offers customers four different order types: single hourly orders; block orders; exclusive groups and flexi orders¹ - with markets open during all the year. In the intraday trading market, customers can trade in 13 intraday markets with high liquidity levels, including the Nordic, Baltic, German, Luxembourg, French, Dutch, Belgian, and Austrian markets (where UK and Germany customers can also participate). Similarly, the intraday markets are also open all year, offering 15 minutes, 30 minutes, hourly and block products that increase the flexibility so that each customer can meet their own needs. To assume a position, the participant can also choose from several order types: limit orders; user-defined block orders; predefined block orders; and iceberg orders².

To better operate the system and handle congestions in the electricity grid more rapidly, each Transmission System Operator divides his respective area into bidding areas - coloured in blue in Figure 2.1. The United Kingdom is also connected but operates as an independent market and not as a bidding area. Additionally, Germany participates in the intraday market and Nord Pool is at the time of writing seeking to expand its day-ahead market to central Europe.

¹ A more detailed overview regarding the day-ahead market orders is provided in Appendix A.

² A more detailed overview regarding the intraday market orders is provided in Appendix B.

Figure 2.1 - Nord Pool Bidding Areas map, by region



Source: Nord Pool (2019)

The major role of a market is to lead supply and demand to converge, from which surges an equilibrium price. This equilibrium is especially important in power markets since electricity cannot be easily stored and failures in supply originate high costs, making it beneficial for both producers and consumers to use Nord Pool services.

This trading of power is performed mainly in the day-ahead market. The exchange functions basically as an auction, where customers submit their orders and an equilibrium between the aggregated supply and demand curves is established for each of the bidding areas. As previously stated, participants can choose among several order types, but the general process is always the same. In summary, the day-ahead market main function is to balance supply and demand in the power market, providing a certain scope for forward planning. To do so, producers and consumers of electricity send their bids/offers to the day-ahead market which then calculates an hourly price that balances both sides. Every morning customers set their bids/offers with respect for the following day, specifying the volume in MW per hour – MW/h - that a customer is willing to buy, or the volume in MW per specific price levels – EUR/MWh – if the party is selling the electricity, discriminated by each hour of the following day. Nord Pool then combines this set of orders and publishes a price for each hour of the following day, properly balanced and adjusted in real-time to face eventual problems.

Analysing now the auction process more closely. At 10:00 CET (Central European Time) the available transmission capacities of the grid are published. In this first phase, until 12:00 CET producers of electricity who want to sell power via Nord Pool set their ask quotes in the day-ahead market, discriminating the quantity of electricity in MW/h as a function of price³ they are willing to sell during the 24 hours of the following day. Similarly, suppliers estimate the end-users' power needs and place their bids, also discriminating the demand for power in MW/h as a function of price. After 12:00 CET bidding offers are aggregated into demand and selling offers into supply, discriminated by hourly intervals and by bidding areas, disregarding the available transmission capacity between these agents in order to anonymize the trades.

The second phase involves the definition of bidding prices. As stated, the Nordic and Baltic areas are divided into bidding areas, where each one can either have a balance, deficit or surplus of electricity. Respecting the law of demand, electricity will then flow from areas where the price offered is lower to the areas when demand is higher. Note that if this transmission of capacity between bidding areas is not sufficient to achieve the full convergence between both areas (i.e. the excess demand on one side remains higher than the excess supply on the other side, or the opposite), this congestion will lead to different prices for each bidding area. Submitted orders are then matched with other orders in the pan-European market coupling process – the Single Day-Ahead Coupling – through a common algorithm known as PCR EUPHEMIA⁴. During this process, the single price for each hour (of the following day) in each bidding zone is set by the intersection of the respective supply and demand curves, taking into account eventual transmission constraints⁵. These will be the prices that the respective producers and consumers will receive and pay, respectively.

³ Even though market participants can price their quotes in EUR, NOK, SEK or DKK, Nord Pool defines the Euro as the official day-ahead market trading currency, so every order needs to be converted to this same currency, at the currency rate valid at 12:00 CET.

⁴ More information relatively to PCR Euphemias algorithm available in Appendix C.

⁵ This way of operating has also external advantages for third parties such as politicians and other stakeholders, by allowing to detect issues along the power chain more efficiently, namely by detecting issues in the power grid when the price goes up. One good example is the identification of production or capacity setbacks, since high prices will

The third and last phase consists in the definition of the Nordic system price, the reference price for the Nordic region. This price is calculated locally at Nord Pool after area prices are calculated and serves as the reference price for most standard financial contracts traded in the Nordic region. Unlike the bidding area prices, the system price is calculated disregarding any network congestion (by artificially setting capacities equal to infinity)⁶.

Hourly clearing prices are then announced to the market, usually at 12:42 CET or slightly later, and each buyer and seller is notified with the respective individual result. At hourly intervals of the second day the delivery will then occur⁷, following an auction model: bidders who have submitted buying orders equal to or higher than the system price will be able to buy electricity at the market price; sellers whose orders are equal or lower than the system price will also be able to sell at the respective market price.

Several studies have already been pursued to assess which are the main factors that impact the system price in Nord Pool. From a fundamental approach, Määttä and Johansson (2011) sought to assess if the traditional factors related to the demand and supply of electricity have statistical significance in explaining the historical prices of the market. Considering both sides of this equation, the results suggest that in fact fundamental variables that form the supply and demand. This capacity to use statistical models to predict, with a certain degree of accuracy, future day-ahead hourly prices based in demand and supply factors is supported by other authors such as Kristiansen (2012), based in the broader work from Weron and Misiorek (2008) that focuses in the electricity markets in general.

indicate an excess of demand and consequently suggest that additional investments should be done in the production side.

⁶ Note that in this phase the delimitation of the areas is different, since Norway, Denmark, Sweden and Finland constitute one block, and the flows now considered are the ones made by this block and the other regions like Netherlands, Germany, Poland and the Baltics.

⁷ In order to prevent that some regions have advantages due to transmission constraints, in Nord Pool day-ahead market prices are announced in the day before the delivery day, reducing this way the uncertainty and providing some room for market participants to plan their offers.

Aiming to compare Nord Pool electricity market with other commodity markets, Botterud, Kristiansen, and Ilic (2010) studied the relation between spot and future prices. Results point that futures prices tend to be constantly higher than the respective spot prices – observed by a negative average convenience yield – but the strength of such conclusion is dependent on cyclical factors. Despite the negative convenience yield, that goes against most findings made in commodity markets, the authors claim that such yield is still directly linked to the fundamental drivers of supply (such as type of production) and demand (namely the risk preferences and ability to profit from short-term price fluctuations). These results were some years later challenged by Weron and Zator (2014), in which the authors took into account seasonal factors that weakened the results from the first study.

Analysing now the volatility of Nord Pool hourly system price, for a sample of 12 years Simonsen (2005) found evidence that suggests that this market shares many similarities with others financial and commodity markets, presenting phenomena like «volatility clustering, log-normal distribution, and long-range volatility-volatility correlations» (Simonsen, 2005). Nevertheless, more uncommon features are also evidenced, such as an overall high level of volatility and price level-dependent volatility. Predictability regarding the volatility of Nord Pool day-ahead market was also studied by Haugom, Westgaard, Solibakke, and Lien (2011), where the authors suggest a stronger predictive capacity of models where total variation is segregated in a continuous and a jump component, due to the strong degree of persistence evidenced by the realized volatility.

Having completed the analysis of the main price drivers of Nord Pool system price, it remains relevant to take a closer look at the financial markets, where derivatives can be traded. Here, investors can find contracts with a time horizon of up to ten years, and with maturities that go from daily and weekly to monthly, quarterly and annual periods, which are mainly used for price hedging and risk management. Unlike in the day-ahead and intraday markets, in the financial markets there is no physical delivery, unless the product is a future contract, and cash settlement takes place throughout trading. In terms of pricing, similarly to any other derivative contract the price of electricity-related derivatives depends mainly on the price of the underlying asset – in this case, this reference price is the system price of Nord Pool. Thus, irregularities found in the system price, such as evidence that supports the hypothesis of price clustering, will have a

subsequent impact via the derivatives, and the traditional models used to value these derivatives will have certain limitations that need to be acquainted.

The existence of price clustering in Nord Pool can impact not only market participants but also every investor and stakeholder. As Sopranzetti and Datar (2002) state, one major impact of such anomaly relies on the derivatives valuation level, since traditional models such as the Black-Scholes are only applicable when the underlying asset's returns follow a normal distribution. Additionally, Hellström et al. (2012) recall for the fact that electricity, in particular, is not a storable good, which once again justifies why traditional valuation models are not adequate to this specific type of utility, meaning that other models should be developed so that the dynamics of the underlying electricity spot price is properly incorporated.

2.3. Price Clustering in Financial Markets

The price clustering phenomena can be simply defined as the tendency for prices to accumulate around certain sets of values. As Brown and Mitchell (2008) state, this tendency is usually assumed to result from human bias and/or imprecise beliefs about the underlying value of the asset, but academics have also developed economical-based explanations for such phenomenon.

Several studies have already been made in the field of price clustering, spread throughout multiple asset classes and markets, building this way a significant amount of evidence supporting this hypothesis first introduced by Osborne (1962) in the academic literature.

In a purely theoretical approach, the occurrence of price clustering and market efficiency are two mutually exclusive hypotheses. The efficient market hypothesis, introduced by Fama (1970), continues to have a great number of supporters, from academics to professional investors, and serves as a foundation for most traditional models developed in Finance. Based on the assumption that prices incorporate all the available public information, since the future is unpredictable one can conclude that prices should follow a “random walk”, i.e. are unable to be estimated with precision beforehand. Thus, since prices reflect the intrinsic value of an asset, there is no plausible reason to explain why it can accumulate around certain specific numbers, and what we should observe is a uniform distribution where each price level has the same probability to occur. Plus, if such anomaly is detected, it is expected that arbitrageurs and other informed investors will take advantage of it and lead once again the market to its original state.

To support this empirical irregularity, academics have also developed theories that seek an explanation for this phenomenon, both based on rational or irrational foundations. Most studies point four main theories in the literature: price resolution hypothesis (Ball et al., 1985); negotiation hypothesis (Harris, 1991); attraction hypothesis (Curcio & Goodhart, 1991); and the collusion hypothesis (Christie & Schultz, 1994). In summary, the explanatory hypothesis can be divided into two main groups, with the first one involving theories based on economic rationales, and the second group involving theories more based in behavioural explanations.

2.3.1. Economic-based Explanations

The price resolution hypothesis was first proposed by Ball et al. (1985), following a study made using the London gold market to investigate and develop an explanation for the rounding in transaction prices. According to this theory, the formation of price clusters is directly related with the investors' level of uncertainty regarding the intrinsic value of the asset – measured by the volatility – because in these situations they tend to use coarser price grids, that is, less accurate prices/lower degree of price resolution. Additionally, higher price levels also seem to lead to the use of coarser price grids by investors, and consequently to price clusters. Besides these factors, liquidity seems to have an opposite effect by allowing to reduce the asymmetry of information and uncertainty, motivated by the higher amount of market participants. To sum up, clustering will occur because investors seek a desired level of price accuracy/price resolution, and seems to be positively correlated with the level of uncertainty and level of price, and negatively correlated with the level of liquidity.

A significant contribution to Ball et al. (1985) work can be pointed to Harris (1991). Focusing the analysis in the New York Stock Exchange the author builds what is known as the negotiation hypothesis supported by the evidence found that stock prices tend to cluster on round fractions. More precisely, the conclusions retrieved show that integers numbers are more common than halves, halves are more common than odd quarters ($1/4$ and $3/4$), odd quarters are more common than odd eights ($1/8$, $3/8$, $5/8$ and $7/8$), and other fractions are rarely observed. According to this hypothesis, the formation of price clusters can be explained by two main factors. First, due to exchange regulations, fractions smaller than eights are rarely used in the NYSE, because for most stocks this is not allowed. Second, the author assumes that this tendency is associated with the fact that traders use these common discrete price sets to make

the convergence between bids and offers faster, thus lowering the costs of negotiating. By other words, trading within these defined price sets is assumed to be intentional due to the economic advantages it brings, mainly by limiting the number of bid and ask quotes thus converging buyers and sellers' bids more rapidly. Also, this second factor is also motivated by the exchange regulations, namely by the requirement imposed that «quotes and transaction prices be stated as some multiple of a minimum price variation, or trading tick» (Harris, 1991), which aims to increase the liquidity of the markets and enforces investors' benefits of trading within the discrete price sets established. Lastly, evidence was also found to support the hypothesis that clustering is positively correlated with price level and volatility, and negatively correlated with capitalization and transaction frequency, which goes in line with the overall argument developed by the author and also by Ball et al. (1985). To summarize, the key idea of the negotiation hypothesis is that when there is uncertainty regarding the true value of a security, investors will prefer to use a common discrete set of prices to minimize their negotiation costs, leading to the formation of price clusters.

One other work worth mention was developed by Christie and Schultz (1994) and is recognized as collusion hypothesis. Encompassing a large sample of stocks from the NASDAQ market, the authors disregard Harris (1991) negotiation hypothesis based on the findings that market makers seemed to avoid odd-eight quotes, which is not efficient nor competitive since it implies a minimum spread of at least \$0.25 (instead of a spread of only \$0.125). In addition, evidence showed that larger trades – with higher costs of negotiation - seemed to occur more on odd eights than trades of 1,000 shares or less, suggesting also that the lack of odd-eight quotes cannot be rationally explained by the hypothesis of investors seeking to minimize costs of negotiation. Thus, instead of defending that clustering results mainly from investors seeking to minimize their transaction costs and speed, Christie and Schultz (1994) raise the hypothesis that this clustering resulted from a collusion of market makers, which have precisely the opposite effect of the one pointed by the negotiation hypothesis, as NASDAQ stocks showed higher trading costs.

2.3.2. Behavioural Explanations

From a behavioural perspective, the first relevant study was developed by Curcio and Goodhart (1991). From the work surged what is known as the attraction hypothesis, which suggests that

investors are naturally attracted to numbers ending in certain values, with each number having a different “gravitational” strength – the number 0 is pointed to be the one preferred by investors, followed by the number 5. In addition to these, investors also seem to be attracted to even numbers more than odd numbers, namely the ones ending in 2 and 8. Numbers ending in 1 or 9 are according to the results the least observed, supporting the hypothesis that investors prefer to round these numbers to the nearest integer.

Another study worth mentioning was developed by Mitchell (2001), where the author seeks to explain why psychological barriers such as price clusters occur in financial markets, finding evidence that factors such as cultural conventions and even the decimal place value system can lead people to become more attracted to certain numbers. Also, individuals are pointed to use heuristics and other methods to simplify the information obtained when using mental calculus, supporting the idea that mental limitations are important and can have practical impacts. Apart from these conclusions, Mitchell (2001) also raises a possible economic explanation for the occurrence of price clustering, based in the idea that «people may select numbers that they believe others will recognize (...) to facilitate the decision-making process» (Mitchell, 2001), in a sense similar to the negotiation hypothesis first presented by Harris (1991).

Narayan and Smyth (2013) examined evidence for the South Pacific Stock Exchange, finding there a tendency for stocks to cluster at the decimals 0 and 5. In addition, the authors examined how political instability induced this phenomenon, results that led to the development of the self-intituled panic trading hypothesis. According to this, panic-motivating factors such as political instability are expected to be positively correlated with the observation of price clusters, theoretically justified by the fact that during these periods investors’ main goal is to close their positions quickly, which can be facilitated if bids are placed using common and discrete sets of values.

Still in the field of behavioural finance, a recent study from Blau (2019) aimed to, instead of finding a theoretical rationale to justify the formation of price clusters, find a correlation between price clustering, specific characteristics of the stocks - such as volatility and spread - and an investor sentiment indicator (which aims to assess if the investors are dominantly bullish or bearish). Based on the assumption (supported by previous studies) that in periods characterized by high investor sentiment (i.e. investors are optimistic) securities are more likely to be mispriced,

the evidence found by Blau (2019) suggests that there is a positive correlation between the level of price clustering and periods of high investor sentiment. Thus, by connecting both previous conclusions, one can infer that price clusters are more likely to occur when securities are mispriced, this way complementing Ball et al. (1985) hypothesis by a different procedure.

2.3.3. Empirical Findings

Although there is still no agreement regarding what justifies the formation of price clusters, the fact that these clusters exist is at this stage unquestionable. In fact, it is not only a phenomenon specific of the American stock market – in which most of the previous theories were based – but it is also detectable within different time frames, assets classes and market structures. Regarding other asset classes, apart from the American stock market it is possible to find evidence of price clustering related to other exchanges (Hu et al., 2017), the foreign exchange market (Sopranzetti & Datar, 2002), the real estate (Palmon et al., 2004), commodities such as gold (Ball et al., 1985), and even about cryptocurrencies such as bitcoin (Urquhart, 2017).

Taking the stock market in particular, price clustering shows to be transversal both through industries, geographies and the country's level of development. Regarding the geography, several studies of price clustering were already developed for the most relevant exchanges, such as: Osborne (1962), Harris (1991), Huang and Stoll (2001) and Comerton-Forde and Putniņš (2013) for the American exchange; Huang and Stoll (2001) and ap Gwilym and Alibo (2003) for the United Kingdom; Sonnemans (2006) and Geoffrey Booth, Kallunki, Lin, and Martikainen (2000) for Netherlands and Finland, respectively; Aitken, Brown, Buckland, Izan, and Walter (1996) for Australia; and Aşcıoğlu, Comerton-Forde, and McInish (2007) for Japan.

Developing and emerging countries are also included within the price clustering literature, even though in a smaller number, but also leading in general to the similar conclusions supporting this phenomenon. To mention some, it is possible to highlight the work developed by: Brown, Chua, and Mitchell (2002) using data from six Asia-Pacific stock markets; Booth and Yuksel (2006) for Istanbul; and Narayan, Narayan, Popp, and D'Rosario (2011) for Mexico.

Despite the wide research regarding price clustering made for many developed and developing financial markets, the nature of such phenomenon often raises conflicting conclusions among academics. For example, Lobão, Pacheco, and Alves (2019) analysed the stock market during

the 2008 financial crisis, more precisely the European and United States banking industry, and from their analysis found that even though these assets showed significant levels of price clustering, this crisis did not have a significant impact on the magnitude of price clustering observed. Such conclusion partially contradicts the expectation from Ball et al. (1985) and Narayan and Smyth (2013) according to which the higher investors' uncertainty, characteristic of the financial crisis, should have resulted in a higher presence of price clusters.

Talking about commodity and utility markets in particular, apart from impact in the valuation models, another indirect but important impact price clustering can have is associated with the adequate management model of exchanges and its regulation. An interesting case to illustrate this occurred in 1994. Following the study already mentioned by Christie and Schultz (1994), which served as the foundation of the collusion hypothesis and where the authors suggested that market makers were colluding to avoid odd-eight quotes in order to increase the spread, Christie et al. (1994) found that almost from a sudden four of the largest and most traded stocks in NASDAQ – Amgen, Cisco, Microsoft and Apple – seemed to stop avoiding odd-eight quotes, with «no evidence that the reduction in spreads is explained by changes in the costs of market making» (Christie et al., 1994), implicitly supporting the initial idea that market makers were in fact colluding between each other. The articles of Christie and Schultz (1994) «closely followed a reported threat of regulation by the NASD in a meeting with market makers» (Christie et al., 1994), once again showing the potential impact of such studies, in this example as a way to partially solve a potential market inefficiency.

For the specific case of our study, focused in Nord Pool, apart from eventual political and regulatory purposes it is plausible to assume that the major utility of the upcoming conclusions reached will be related to the valuation of electricity derivatives for hedging and portfolio management objectives. Thus, it is still relevant to analyse similar studies made for other utility markets – with structure similar to the Nord Pool – since from all the previous studies mentioned it is possible to conclude that, in general, the type of price clustering and underlying motives can widely vary from one asset class to another.

The utilities sector includes companies that provide basic goods and services such as water, electricity and natural gas. A study made by Brooks et al. (2013) analyses the latter, more precisely the occurrence of price clustering in the Australian water market. The reform in the Australian

water sector began in the 80's decade, intending to not only develop water markets in every region but also benefit from the market mechanism, leading to flexibility and efficiency gains namely in the control of producer and consumer surplus and even in terms of water quality. Similarly to Nord Pool electricity market, the Australian water market also allows to build risk hedging strategies for those entities whose activity is dependent on this utility, namely farmers who are now able to manage their intra-seasonal risk. This increasing participation has also led to the creation of an electronic clearing house in 2002 called Watermove (closed in 2013 but replaced by other entities such as Waterpool) whose functions are closely related to the ones of Nord Pool: reduce search and negotiation costs; act as a clearing house; and provide an online exchange platform (Waterpool, 2019). Relatively to the price formation, the process also consists in an auctioning with bidders and sellers sorted accordingly to their offers (the bidders who are willing to pay more and the sellers who accept lower values are first in the position), and the pool price for a certain trading zone will be the one that maximizes the volume of water traded. Studying the prices for three different trading zones, the authors concluded in a first phase that there was significant evidence suggesting clustering at integers dollar amounts. Furthermore, for a deeper analysis the authors explored for which whole digits buy and sell offers have more probability to cluster, having found that these are most concentrated around values ending in 0, followed by a smaller concentration around whole numbers ending in 5. Besides, similarly to previous studies Brooks et al. (2013) also tried to find correlated variables that might explain how and why price clusters are formed, finding for example that with greater activity the tendency for price clusters decreased.

3. Data Base and Methodology

3.1. Data Base

The following empirical study will involve two sequential parts, according to the results obtained. On a first moment, the price clustering hypothesis will be tested, through the procedures that will be later described. To analyse this phenomenon in its temporal and cross-sectional dimensions, the system price and the bidding area prices will be considered for different frequencies: hourly prices for the year of 2019, daily prices from January 2017 till December of 2019, and monthly prices also from January 2017 till December of 2019. Since Nord Pool declared itself that information regarding bid and ask quotes is not available, the system prices will be used as a reference price for the exchange. Since the day-ahead market is the predominant choice by market participants to transact power, the price of this market will be the dependent variable analysed. Additionally, the hourly system price also assumes a special relevance, as it is the one defined directly by market participants and used in most derivative instruments.

Nord Pool assumes the Euro as the official day-ahead market trading currency, so this will be the currency analysed. The most favourable scenario would be the one where bid and ask quotes were available, to analyse if market participants show a preference for defined sets of values. Nevertheless, as one important impact being studied is associated with the implications on the valuation of derivatives, the final settlement prices are the ones relevant, so the conclusions are still valid.

For those bidding areas where the hypothesis of price clustering shows statistical support, further analysis will be made to study the nature of this phenomenon. As seen in the literature, price clustering seems to have different potential explanatory motives, so a multivariate analysis should allow to better understand these. For such analysis, three main variables will be used: (1) the price level; (2) the level of volatility implicit in prices; (3) and the volume of energy traded, as a measure of liquidity.

As stated previously, other variables would be also relevant for a deeper study. Namely, the bid and ask quotes would allow to infer the bid-ask spread, which would also measure the transaction costs and the market liquidity, but this information is not publicly available.

3.2. Methodology

To answer both questions proposed by this study, the methodology will be divided into two sequential parts. The first phase will consist of an analysis of the price data from Nord Pool, in order to assess whether it shows evidence of price clustering or not. The second phase will surge if the result from the first phase is positive, i.e. if the hypothesis of price clustering is valid. In this second phase, the nature of this phenomenon will be analysed, so that a following comparison with the literature is implemented. Due to the completeness and similarities of the questions analysed, the study performed by Brooks et al. (2013) regarding the Australian water markets will be used as a reference throughout this chapter, as well as the research made by Chung and Chiang (2006) and Grossman, Miller, Cone, Fischel, and Ross (1997) whose methodological techniques also served as a foundation for the first study mentioned.

For the first phase of the analysis, the hypothesis to be tested is to assess if Nord Pool price data shows evidence of price clustering or not.

A first assessment of price clustering passes through the analysis of a frequency table, based in Brooks et al. (2013), Chung and Chiang (2006), Grossman et al. (1997) and Lobão et al. (2019), where prices from the sample are segregated by their ending number. This test will involve two phases: at first the number of occurrences of whole digits versus decimal digits will be compared; second, within the prices that comprehend decimal values, a relative frequency for the figure in the hundredth place will be extracted. In the presence of price clustering, the threshold for the first test will be one percent – i.e. in an uniform distribution, numbers ending in XX.00 should represent close to one percent of the observations – while for the second test this threshold is ten percent – i.e. prices ending in XX.X0 should have the same probability than those ending in XX.01; XX.02; ...; XX.09. In order to perform a both time-series and cross-sectional analysis, every bidding area data was analysed for different timeframes. Nevertheless, the hourly system price will assume a major relevance, as it serves as the main reference price for most derivative instruments related to Nord Pool.

The second test will rely upon the Chi-squared goodness-of-fit statistics, once again taking as a reference the work from Brooks et al. (2013), Chung and Chiang (2006) and Lobão et al. (2019), applied on the frequencies of the last hundredth digit of each price. This test aims to compare

the observed level of price clustering with the expected level of price clustering under a uniform distribution, where the difference between the observed and expected number of occurrences are studied as follows:

$$W = \sum \frac{(O_i - A_i)^2}{A_i}$$

Equation 3.1 - Chi-Squared Goodness-of-Fit test

where O_i represents the observed absolute frequency and A_i the expected absolute frequency under a uniform distribution, for each ending digit. Note that for k possible digits there will be $(k - 1)$ degrees of freedom under the null hypothesis of no price clustering. In our case, k will therefore assume the value of 10, as there are 10 possible ending digits for the hundredth decimal place. Under the null hypothesis H_0 each outcome should have the same probability to occur (one-tenth) while under the alternative hypothesis H_1 there is evidence that the sample is concentrated around certain values. A higher divergence from the observed and expected frequencies will lead to a higher value of the W statistic, suggesting a higher divergence from the uniform distribution hypothesis.

The third and final test for this first phase consists in the computation of the Herfindahl-Hirschman Index (HHI), a measure widely used to assess the concentration of a certain market by computing the sum of the squared market shares of all market participants. When applied to assess the price clustering of the hundredth decimal number, such as in the works from Brooks et al. (2013), Chung and Chiang (2006) and Lobão et al. (2019), it will be calculated as follows:

$$HHI = \sum (f_i)^2$$

Equation 3.2 - Herfindahl-Hirschman Index calculation

where f_i is equal to the frequency (in percentage) of prices that occur for each ending digit i . Thus, if there is not price clustering it is expected that HHI takes a value close to $1/k$, in our case $1/10$, meaning that prices ending in a certain digit should represent close to 10% of the entire observations. If values obtained for the HHI are higher than 0.10, evidence of price clustering is found.

After this first phase of price clustering identification, a multivariable analysis is implemented to Nord Pool day-ahead hourly prices, in order to better understand the nature of this phenomena and connect the results with the theories already existent within the literature. Literature points that factors such as the level of price, volatility (a proxy of uncertainty) and the volume can be correlated with the tendency for price clusters to form. To act as a comparison with the study developed by Brooks et al. (2013), the dependent variable for the regression will be the difference (in percentage terms) between the observed level of price clustering and the expected level of price clustering under a uniform price distribution. The multivariate regression estimated comprehend the following parameters:

$$\begin{aligned} Clustering - E(Clustering)_t = & \beta_0 + \beta_1 Price_t \\ & + \beta_2 \log(Buy Volume)_t \\ & + \beta_3 \log(Sell Volume)_t + \varepsilon_t \end{aligned}$$

Equation 3.3 - Multivariable Regression Model to study price clustering nature, without considering volatility

where the dependent variable $Clustering - E(Clustering)_t$ is the difference between the observed and the expected relative frequency of price clustering for observation t , expressed in percentage. $Price_t$ represents the respective system price or bidding area price of the observation t , expressed in Euros/MWh. As Nord Pool provides data for both buy and sell volume of each area (due to the transactions that occur between these) the two measures expressed in MWh was segregated in $\log(Buy Volume)_t$ and $\log(Sell Volume)_t$. The choice of these variables is based in the work from Brooks et al. (2013) and Chung and Chiang (2006), adjusted to the public information provided by Nord Pool.

Regarding the incorporation of volatility, it is important to recall that quotes from Nord Pool day-ahead market are placed in the previous day, discriminated by monetary units per megawatt/hour. In addition, prices in electricity markets have a daily cyclical component, where demand and price decreases during “sleep hours”. These special characteristics from the market create the necessity to use another proxy different than the ones available to assess the stock market intraday hourly volatility such as the Parkinson (1980) extreme value estimator used by Chung and Chiang (2006). To overcome such problem and still consider the effect of volatility in the formation of price clusters, apart from the linear regression model presented in Equation

3.3, another regression with the log-transformed daily standard deviation was also computed, as presented in Equation 3.4:

$$\begin{aligned} Clustering - E(Clustering)_t = & \beta_0 + \beta_1 Price_t \\ & + \beta_2 \log(Volatility)_d \\ & + \beta_3 \log(Buy Volume)_t \\ & + \beta_4 \log(Sell Volume)_t + \varepsilon_t \end{aligned}$$

Equation 3.4 - Multivariable Regression Model to study price clustering nature, considering volatility

The log-transformed daily standard deviation $\log(Volatility)_d$ is computed, for each day d , as:

$$\log(Volatility)_d = \log \left[\sqrt{\frac{1}{m-1} \sum_{i=1}^m (P_{n-i} - \bar{P})^2} \right]_d \quad [5]$$

Equation 3.5 - Calculation for the log-transformed daily standard deviation

where m is equal to the total n observations for day d (thus m is equal to 24 hourly observations), P_n corresponds to the bidding area price of the hourly observation n of day d , and $\bar{P}$ represents the average of the hourly price for day d . The introduction of $\log(Volatility)_d$ acts only as an external addition to the model, since despite the incorporation of this dimension in previous works namely from Chung and Chiang (2006), the special characteristics of Nord Pool bidding system make it necessary to structure a different proxy measure of the market uncertainty, different than the ones used by previous authors when applied to the stock market.

Focusing in the more extensive regression presented in Equation 3.4, and taking into consideration the literature review presented in the previous chapters, namely the most relevant price clustering hypothesis that dominate the literature, *ex-ante* insights can be retrieved regarding the signal of each estimated parameter. According to the price resolution hypothesis by Ball et al. (1985) estimates for coefficients β_1 and β_2 – level of price and volatility, respectively – are expected to present a positive signal, while estimates for β_3 and β_4 – associated with liquidity – should contribute negatively to the formation of price clusters, thus are expected to have a negative signal.

4. Empirical Results and Analysis

The following chapter includes the results from the empirical application of the tests presented in the previous section, as well as the respective analysis. Two sequential sections are presented, where the first tests the existence of price clustering (segregated by hourly, daily and weekly timeframes) and the latter studies the nature of price clustering in hourly prices, since these present, as it is going to be seen, more evidence of price clustering and more financial implications namely in the valuation of derivatives.

4.1. Tests on Price Clustering Existence

a) Hourly Prices

Starting our tests by detecting the presence of price clustering in hourly data, Table 4.1 Panel A summarizes the absolute and relative frequency for the number of occurrences of each figure in the hundredth decimal place, highlighting also the number of prices ending in XX.X0 and XX.05 due to its relevance within the existent literature. Frequencies of prices ending in XX.00 are also displayed, to allow for comparisons between whole versus decimal values, similarly to Brooks et al. (2013). Both Herfindahl-Hirschman Index (HHI) and Chi-squared goodness-of-fit test are displayed in Panel B, where the p-value of the latter statistic is presented under brackets below the respective value. For this study, the system price (SYS) and bidding areas' prices were analysed, as well as the clustering in the total sample (excluding the system price) which is presented in the last row denominated as "All".

Table 4.1 - Frequency distribution for the hundredth decimal number and whole digits of Nord Pool day-ahead market hourly prices

Panel A: Distribution of the last digit or last two digits of the price.															Panel B: Clustering tests and HHI	
Last Digit	0	1	2	3	4	5	6	7	8	9	Total	0 & 5	00	X ²	HHI (%)	
SYS	#	909	905	939	868	836	864	879	863	883	8,760	1,773	110	13.44521	10.02	
	%	10.38	10.33	10.72	9.91	9.54	9.86	10.03	9.85	10.08		20.24	1.26	(0.143474)		
SE1	#	944	899	865	881	839	815	888	917	838	8,760	1,759	139	15.5959	10.02	
	%	10.78	10.26	9.87	10.06	9.58	9.30	10.14	10.47	9.57		20.08	1.59	(0.0758**)		
SE2	#	943	900	864	881	840	815	888	917	838	8,760	1,758	139	15.4384	10.02	
	%	10.76	10.27	9.86	10.06	9.59	9.30	10.14	10.47	9.57		20.07	1.59	(0.0796**)		
SE3	#	971	903	881	876	839	811	891	898	835	8,760	1,782	142	20.7808	10.02	
	%	11.08	10.31	10.06	10.00	9.58	9.26	10.17	10.25	9.53		20.34	1.62	(0.0137**)		
SE4	#	960	915	895	883	833	832	870	863	848	8,760	1,792	149	15.9658	10.02	
	%	10.96	10.45	10.22	10.08	9.51	9.50	9.93	9.85	9.68		20.46	1.70	(0.0676*)		
FI	#	952	874	855	890	847	832	877	880	859	8,760	1,784	149	11.2146	10.01	
	%	10.87	9.98	9.76	10.16	9.67	9.50	10.01	10.05	9.81		20.37	1.70	(0.2613)		
DK1	#	1251	842	816	840	810	821	797	848	883	8,760	2,072	280	184.5982	10.21	
	%	14.28	9.61	9.32	9.59	9.25	9.37	9.10	9.73	10.08		23.65	3.20	(0.0000***)		
DK2	#	1129	887	840	865	799	824	815	861	878	8,760	1,953	229	89.4132	10.10	
	%	12.89	10.13	9.59	9.87	9.12	9.41	9.30	9.83	10.02		22.29	2.61	(0.0000***)		
Oslo	#	919	897	901	873	805	844	883	870	910	8,760	1,763	137	12.0479	10.01	
	%	10.49	10.24	10.29	9.97	9.19	9.63	10.08	9.79	10.39		20.13	1.56	(0.2106)		

Panel A: Distribution of the last digit or last two digits of the price.															Panel B: Clustering tests and HHI		
Last Digit		0	1	2	3	4	5	6	7	8	9	Total	0 & 5	00	X ²	HHI (%)	
Kr.sand	#	916	909	894	869	816	838	882	867	859	910	8,760	1,754	135	11.0365	10.01	
	%	10.46	10.38	10.21	9.92	9.32	9.57	10.07	9.90	9.81	10.39		20.02	1.54	(0.2732)		
Bergen	#	913	897	893	875	806	838	889	871	863	915	8,760	1,751	133	11.7900	10.01	
	%	10.42	10.24	10.19	9.99	9.20	9.57	10.15	9.94	9.85	10.45		19.99	1.52	(0.2254)		
Molde	#	908	906	873	866	861	833	893	887	873	860	8,760	1,741	132	5.4589	10.01	
	%	10.37	10.34	9.97	9.89	9.83	9.51	10.19	10.13	9.97	9.82		19.87	1.51	(0.7926)		
Tr.heim	#	908	906	873	866	861	833	893	887	873	860	8,760	1,741	132	5.4589	10.01	
	%	10.37	10.34	9.97	9.89	9.83	9.51	10.19	10.13	9.97	9.82		19.87	1.51	(0.7926)		
Tromsø	#	893	895	863	858	870	859	902	880	874	866	8,760	1,752	122	2.5845	10.00	
	%	10.19	10.22	9.85	9.79	9.93	9.81	10.30	10.05	9.98	9.89		20.00	1.39	(0.9785)		
EE	#	917	908	871	893	839	828	877	904	866	857	8,760	1,745	142	9.0616	10.01	
	%	10.47	10.37	9.94	10.19	9.58	9.45	10.01	10.32	9.89	9.78		19.92	1.62	(0.4316)		
LV	#	932	898	874	907	850	823	867	900	857	852	8,760	1,755	140	11.0320	10.01	
	%	10.64	10.25	9.98	10.35	9.70	9.39	9.90	10.27	9.78	9.73		20.03	1.60	(0.2735)		
LT	#	933	903	865	897	851	824	869	894	863	861	8,760	1,757	141	9.8584	10.01	
	%	10.65	10.31	9.87	10.24	9.71	9.41	9.92	10.21	9.85	9.83		20.06	1.61	(0.3621)		
AT	#	674	417	409	404	377	378	385	439	406	456	4,345	1,052	210	159.9321	10.37	
	%	15.51	9.60	9.41	9.30	8.68	8.70	8.86	10.10	9.34	10.49		24.21	4.83	(0.0000***)		

Panel A: Distribution of the last digit or last two digits of the price.														Panel B: Clustering tests and HHI	
Last Digit	0	1	2	3	4	5	6	7	8	9	Total	0 & 5	00	X ²	HHI (%)
BE	#	806	408	389	384	394	353	401	366	443	4,345	1207	222	365.0783	10.84
	%	18.55	9.39	8.95	8.84	9.07	8.12	9.23	8.42	10.20		27.78	5.11	(0.0000**)	
DE-LU	#	697	433	400	391	384	354	431	393	443	4,345	1116	178	191.1818	10.44
	%	16.04	9.97	9.21	9.00	8.84	8.15	9.92	9.04	10.20		25.68	4.10	(0.0000**)	
FR	#	722	403	388	382	388	365	443	373	472	4,345	1131	229	233.5339	10.54
	%	16.62	9.28	8.93	8.79	8.93	8.40	10.20	8.58	10.86		26.03	5.27	(0.0000**)	
NL	#	1045	366	377	370	367	347	376	340	386	4,345	1416	289	957.0046	12.20
	%	24.05	8.42	8.68	8.52	8.45	7.99	8.65	7.83	8.88		32.59	6.65	(0.0000**)	
All	#	19,333	16,366	15,886	15,951	15,276	15,785	16,118	15,852	16,070	161,885	34,581	3,569	746.1811	10.05
	%	11.94	10.11	9.81	9.85	9.44	9.75	9.96	9.79	9.93		21.36	2.20	(0.0000**)	

Table 4.1 Panel A presents the frequency table regarding Nord Pool hourly prices decimal values, for all the available bidding areas. The symbol # represents the absolute frequency, while % represents the relative frequency.

Names of the bidding areas are displayed by their initials, where: SYS stands for the system price; SE1, SE2, SE3 and SE4 refer to Sweden; FI refers to Finland; DK1 and DK2 refer to Denmark; Osl, Kr, and, Bergen, Molde, Tr, heim and Trondheim refer to cities in Norway; EE stands for Estonia; LV refers to Latvia; LT refers to Lithuania; AT refers to Austria; BE refers to Belgium; DE-LU refers to Germany; FR refers to France; NL refers to Netherlands.

Panel B of Table 4.1 includes the tests price clustering tests applied for each bidding area. HHI stands for Hirschman-Herfindahl Index, while X² is the Chi-squared goodness-of-fit test, with the respective p-value displayed below in brackets.

*** represents values statistically significant for a level of confidence of 99%.

** represents values statistically significant for a level of confidence of 95%.

* represents values statistically significant for a level of confidence of 90%.

On a first analysis of the frequency table presented in Panel A of Table 4.1, the occurrence of price clustering at the hundredth decimal place does not seem to be a generalized phenomenon throughout every bidding area. Nevertheless, as expected by some authors such as Curcio and Goodhart (1991) in their attraction hypothesis, the number 0 shows a relatively higher frequency relatively to others, with such conclusion being more salient for the bidding areas of DK1, DK2, AT, BE, DE-LU, FR and NL. On the other hand, the number 5, pointed as the second most frequent number by the authors, does not present an abnormal number of occurrences in any bidding area. The areas of DK1, DK2, AT, BE, DE-LU, FR and NL show also an abnormal frequency of whole digits, taking into consideration that under a uniform distribution the frequency of numbers ending in XX.00 should be close to 1%. The system price does not show significant evidence that suggests the existence of price clustering, neither in the hundredth decimal number or for whole values. When analysing the whole sample, the previous conclusions retrieved for the mentioned bidding areas are still applicable.

Observing now Panel B from Table 4.1, the statistical tests for the detection of price clustering allow to conclude that the previous conclusions are statistically significant, namely for the bidding areas of *DK1*, *DK2*, *AT*, *BE*, *DE-LU*, *FR* and *NL* for a level of confidence of 99%, and for a level of confidence of 90% the group includes also the bidding areas of *SE1*, *SE2*, *SE3* and *SE4*. The system price shows once again no sufficient evidence that suggests the presence of price clustering in the sample, resulting in a p-value of 14.35%. For the entire sample, results remain statistically significant for a 99% confidence of level.

A temporal analysis of the cumulative number of occurrences for each hundredth number, in the hourly prices for the entire sample, is displayed in the chart of Figure 4.1.

Figure 4.1 - Cumulative number of occurrences of each number, for Nord Pool day-ahead hourly prices of the entire sample

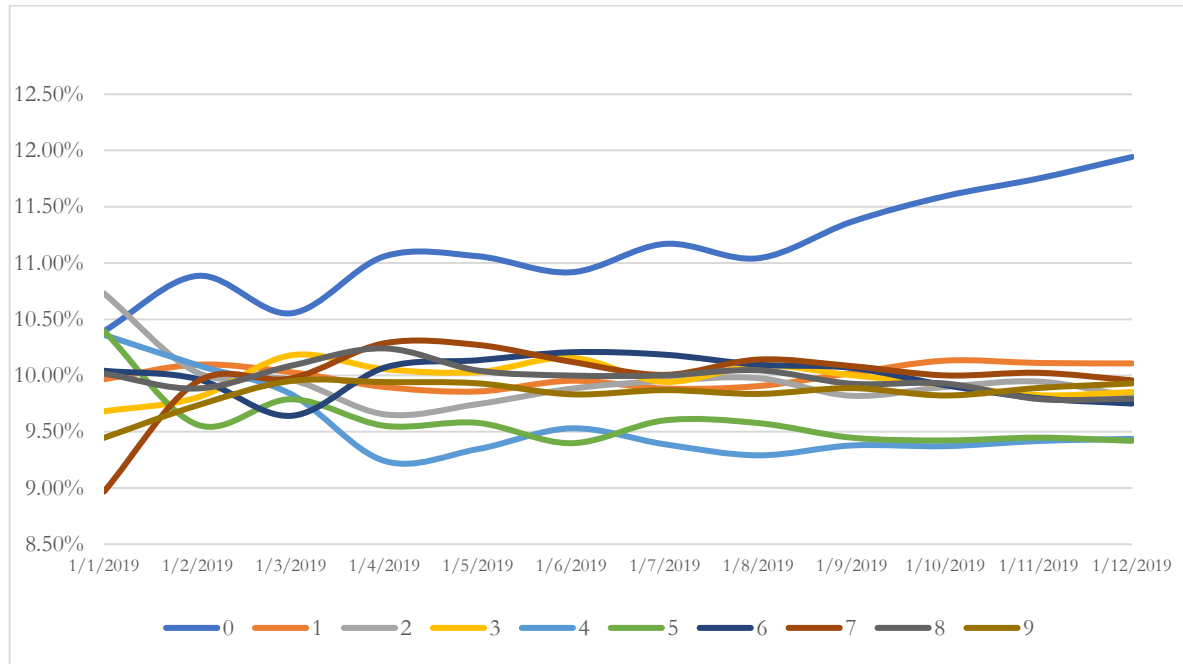


Figure 4.1 displays the cumulative number of occurrences for each number in the hundredth place of Nord Pool's day-ahead hourly prices, for the entire sample, except the system price, from January 1st to December 31st of 2019.

As observed in Figure 4.1, the dominance of prices ending in XX.X0 is relevant throughout the entire sampling period. Although the majority of other numbers fluctuate close to the expected value of 10% of the total occurrences, prices ending in numbers 4 and 5 present the lowest relative frequency, which goes in accordance with the results from Table 4.1.

b) Daily Prices

Looking now to daily prices, Table 4.2 Panel A displays both absolute and relative frequency for the number of occurrences of each digit in the hundredth decimal place. The difference between whole digits versus decimal prices is also assessed through the column with the number of occurrences of numbers with the last two digits being 0. Panel B presents both Herfindahl-Hirschman Index (HHI) and Chi-squared goodness-of-fit tests, where the p-value of the latter statistic is presented under brackets below the respective value. In addition to the system price (SYS) and bidding areas' prices, clustering in the total sample is also analysed in the last row denominated as "All".

Table 4.2 - Frequency distribution for the hundredth decimal number and whole digits of Nord Pool day-ahead market daily prices

<i>Panel A: Distribution of the last digit or last two digits of the price.</i>														<i>Panel B: Clustering tests and HHI</i>		
Last Digit		0	1	2	3	4	5	6	7	8	9	Total	0 & 5	00	X²	HHI (%)
SYS	#	115	106	112	103	104	115	96	113	108	123	1095	230	11	4.844749	10.04%
	%	10.50	9.68	10.23	9.41	9.50	10.50	8.77	10.32	9.86	11.23		21.00	1.00	(0.8476)	
SE1	#	127	97	112	95	114	126	92	107	113	112	1095	253	16	11.8950	10.11
	%	11.60	8.86	10.23	8.68	10.41	11.51	8.40	9.77	10.32	10.23		23.11	1.46	(0.2193)	
SE2	#	126	98	113	95	114	126	92	107	113	111	1095	252	16	11.3836	10.10
	%	11.51	8.95	10.32	8.68	10.41	11.51	8.40	9.77	10.32	10.14		23.01	1.46	(0.2503)	
SE3	#	120	91	118	105	114	124	90	115	108	110	1095	244	15	10.8539	10.10
	%	10.96	8.31	10.78	9.59	10.41	11.32	8.22	10.50	9.86	10.05		22.28	1.37	(0.2859)	
SE4	#	105	119	109	92	117	124	106	120	97	106	1095	229	14	8.8995	10.08
	%	9.59	10.87	9.95	8.40	10.68	11.32	9.68	10.96	8.86	9.68		20.91	1.28	(0.4466)	
FI	#	112	122	121	99	114	116	103	108	103	97	1095	228	9	6.4886	10.06
	%	10.23	11.14	11.05	9.04	10.41	10.59	9.41	9.86	9.41	8.86		20.82	0.82	(0.6902)	
DK1	#	123	118	112	107	112	109	103	104	103	104	1095	232	19	3.8219	10.03
	%	11.23	10.78	10.23	9.77	10.23	9.95	9.41	9.50	9.41	9.50		21.19	1.74	(0.9227)	
DK2	#	119	102	116	104	99	121	114	102	104	114	1095	240	9	5.3744	10.05
	%	10.87	9.32	10.59	9.50	9.04	11.05	10.41	9.32	9.50	10.41		21.92	0.82	(0.8005)	
Oslo	#	110	117	114	112	98	114	114	101	114	101	1095	224	16	3.8402	10.04
	%	10.05	10.68	10.41	10.23	8.95	10.41	10.41	9.22	10.41	9.22		20.46	1.46	(0.9216)	

Panel A: Distribution of the last digit or last two digits of the price.															Panel B: Clustering tests and HHI	
Last Digit		0	1	2	3	4	5	6	7	8	9	Total	0 & 5	00	X ²	HHI (%)
Kr.sand	#	111	112	123	108	97	106	116	104	114	104	1095	217	18	4.4247	10.04
	%	10.14	10.23	11.23	9.86	8.86	9.68	10.59	9.50	10.41	9.50		19.82	1.64	(0.8813)	
Bergen	#	121	115	125	107	109	111	110	91	108	98	1095	232	17	8.1142	10.07
	%	11.05	10.50	11.42	9.77	9.95	10.14	10.05	8.31	9.86	8.95		21.19	1.55	(0.5227)	
Molde	#	118	105	106	104	124	112	110	118	105	93	1095	230	13	6.5434	10.06
	%	10.78	9.59	9.68	9.50	11.32	10.23	10.05	10.78	9.59	8.49		21.00	1.19	(0.6845)	
Tr.heim	#	118	105	106	104	124	112	110	118	105	93	1095	230	13	6.5434	10.06
	%	10.78	9.59	9.68	9.50	11.32	10.23	10.05	10.78	9.59	8.49		21.00	1.19	(0.6845)	
Tromsø	#	121	106	98	106	118	117	108	103	114	104	1095	238	14	4.6804	10.04
	%	11.05	9.68	8.95	9.68	10.78	10.68	9.86	9.41	10.41	9.50		21.74	1.28	(0.8612)	
EE	#	115	118	124	94	120	114	102	102	104	102	1095	229	10	8.0594	10.07
	%	10.50	10.78	11.32	8.58	10.96	10.41	9.32	9.32	9.50	9.32		20.91	0.91	(0.5282)	
LV	#	108	110	107	105	126	115	113	104	103	104	1095	223	13	4.0776	10.04
	%	9.86	10.05	9.77	9.59	11.51	10.50	10.32	9.50	9.41	9.50		20.37	1.19	(0.9062)	
LT	#	104	117	107	97	118	124	109	111	102	106	1095	228	9	5.5023	10.05
	%	9.50	10.68	9.77	8.86	10.78	11.32	9.95	10.14	9.32	9.68		20.82	0.82	(0.7885)	
AT	#	20	13	17	21	25	16	15	14	18	23	182	36	0	7.7802	10.43
	%	10.99	7.14	9.34	11.54	13.74	8.79	8.24	7.69	9.89	12.64		19.78	0.00	(0.5564)	

Panel A: Distribution of the last digit or last two digits of the price.															Panel B: Clustering tests and HHI		
Last Digit		0	1	2	3	4	5	6	7	8	9	Total	0 & 5	00	X ²	HHI (%)	
BE	#	16	16	11	22	32	18	18	16	9	24	182	34	2	21.4066	11.18	
	%	8.79	8.79	6.04	12.09	17.58	9.89	9.89	8.79	4.95	13.19		18.68	1.10 (0.0110**)			
DE-LU	#	22	16	23	17	17	12	21	18	17	19	182	34	1	5.1429	10.28	
	%	12.09	8.79	12.64	9.34	9.34	6.59	11.54	9.89	9.34	10.44		18.68	0.55 (0.8217)			
FR	#	23	11	15	25	20	15	11	18	17	27	182	38	3	15.1429	10.83	
	%	12.64	6.04	8.24	13.74	10.99	8.24	6.04	9.89	9.34	14.84		20.88	1.65 (0.0871*)			
NL	#	16	17	15	23	21	23	21	12	14	20	182	39	0	7.5604	10.42	
	%	8.79	9.34	8.24	12.64	11.54	12.64	11.54	6.59	7.69	10.99		21.43	0.00 (0.5790)			
All	#	1955	1825	1892	1742	1933	1955	1778	1793	1785	1772	18,430	3910	227	33.2306	10.02	
	%	10.61	9.90	10.27	9.45	10.49	10.61	9.65	9.73	9.69	9.61		21.22	1.23 (0.0001****)			

Table 4.2 Panel A presents the frequency table regarding Nord Pool daily prices decimal values, for all the available bidding areas. The symbol # represents the absolute frequency, while % represents the relative frequency.

Names of the bidding areas are displayed by their initials, where: SYS stands for the system price; SE1, SE2, SE3 and SE4 refer to Sweden; FI refers to Finland; DK1 and DK2 refer to Denmark; Oda, Kr.sund, Bergen, Molde, Trondheim and Tromsø refer to cities in Norway; EJE stands for Estonia; LV refers to Latvia; LT refers to Lithuania; AT refers to Austria; BE refers to Belgium; DE-LU refers to Germany; FR refers to France; NL refers to Netherlands.

Panel B of Table 4.2 includes the tests price clustering tests applied for each bidding area. HHI stands for Hirschman-Herfindahl Index, while X² is the Chi-squared goodness-of-fit test, with the respective p-value displayed below in brackets.

*** represents values statistically significant for a level of confidence of 99%.

** represents values statistically significant for a level of confidence of 95%.

* represents values statistically significant for a level of confidence of 90%.

From the analysis of Table 4.2 Panel A, evidence of price clustering at both hundredth decimal place and whole numbers is less supportive to the existence of price clustering than the one obtained for hourly prices, when observing each bidding area individually. Areas such as *DE-LU* and *FR* still present a relatively higher frequency of prices ending in *XX.X0*, but not as significant as before. For the entire sample, the most frequent number in the last digit of daily prices are 0 and 5, followed by the numbers 4 and 2.

Such conclusions are supported when we observe the price clustering tests presented in Panel B from Table 4.2. Daily prices show once again less support for the existence of price clustering at the hundredth decimal place than the one observed at hourly data, as for a 90% confidence level only the bidding areas of *BE* and *FR* present statistically significant evidence that allows to reject the null hypothesis of uniform price distribution. Nevertheless, tests applied to the entire group of bidding areas' prices reveal that there is statistical evidence that suggests an abnormal concentration of prices, sufficient to reject the null hypothesis of a uniform price distribution for a 99% confidence level.

The temporal analysis of the cumulative number of occurrences for each hundredth number, in the daily prices for the entire sample, is displayed in the chart of Figure 4.2.

Figure 4.2 - Cumulative number of occurrences of each number, for Nord Pool day-ahead daily prices of the entire sample

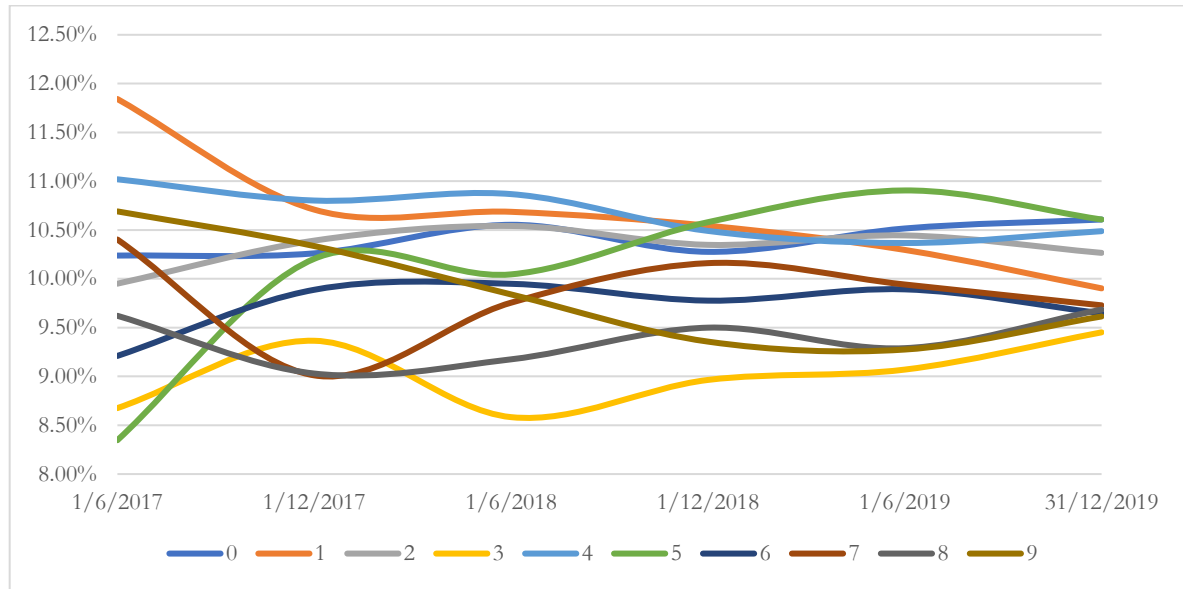


Figure 4.2 displays the cumulative number of occurrences for each number in the hundredth place of Nord Pool's day-ahead daily prices, for the entire sample, except the system price, from January 1st 2017 to December 31st of 2019. First semester was omitted from Figure 4.2 to eliminate the initial dispersion of the sample, although cumulative price occurrences of this period are also included.

Note: Data for the bidding areas of *AT*, *BE*, *DE-LU*, *FR* and *NL* is only available starting from July 3rd, 2019.

Comparing Figure 4.2 with the results obtained for hourly prices, several conclusions can be taken. Primarily, the number 0 does not show the same strength in daily prices throughout the sample, despite ending the period with 10.61% of the total occurrences. On the other hand, the number 5 already appears more frequently in the hundredth place, observing at the end a relative frequency close to the one of numbers ending in 0, as it is also seen in Table 4.2.

c) Weekly Prices

Analysing Nord Pool day-ahead market weekly prices, Table 4.3 Panel A displays the absolute and relative frequencies for the number of occurrences each number is present in the hundredth decimal place. Similarly to the previous two studies, the difference between whole digits versus cents is also assessed by observing the number of occurrences of prices with the last two digits being 0. Panel B presents the Herfindahl-Hirschman Index (HHI) and Chi-squared goodness-of-fit tests, where the p-value of the latter statistic is presented under brackets below the

respective value. Apart from the system price (SYS) and bidding areas' prices, clustering in the total sample is also analysed in the last row denominated as "All".

Table 4.3 - Frequency distribution for the hundredth decimal number and whole digits of Nord Pool day-ahead market weekly prices

Panel A: Distribution of the last digit or last two digits of the price.															Panel B: Clustering tests and HHI		
Last Digit		0	1	2	3	4	5	6	7	8	9	Total	0 & 5	00	X ²	HHI (%)	
SYS	#	13	17	19	22	12	15	11	17	15	15	156	28	0	6.3077	10.40	
	%	8.33	10.90	12.18	14.10	7.69	9.62	7.05	10.90	9.62	9.62		17.95	0.00	(0.7088)		
SE1	#	19	14	16	22	13	19	11	13	13	16	156	38	2	6.9487	10.45	
	%	12.18	8.97	10.26	14.10	8.33	12.18	7.05	8.33	8.33	10.26		24.36	1.28	(0.6425)		
SE2	#	19	14	16	22	13	19	11	13	13	16	156	38	2	6.9487	10.45	
	%	12.18	8.97	10.26	14.10	8.33	12.18	7.05	8.33	8.33	10.26		24.36	1.28	(0.6425)		
SE3	#	15	14	14	23	18	16	12	12	14	18	156	31	1	6.4359	10.41	
	%	9.62	8.97	8.97	14.74	11.54	10.26	7.69	7.69	8.97	11.54		19.87	0.64	(0.6956)		
SE4	#	13	16	20	14	11	14	17	20	17	14	156	27	1	5.0256	10.32	
	%	8.33	10.26	12.82	8.97	7.05	8.97	10.90	12.82	10.90	8.97		17.31	0.64	(0.8321)		
FI	#	16	16	18	18	12	7	13	17	21	18	156	23	2	9.1282	10.59	
	%	10.26	10.26	11.54	11.54	7.69	4.49	8.33	10.90	13.46	11.54		14.74	1.28	(0.4255)		
DK1	#	16	16	12	12	19	12	15	17	23	14	156	28	1	7.0769	10.45	
	%	10.26	10.26	7.69	7.69	12.18	7.69	9.62	10.90	14.74	8.97		17.95	0.64	(0.6291)		
DK2	#	19	9	16	18	20	16	20	12	9	17	156	35	3	10.1538	10.65	
	%	12.18	5.77	10.26	11.54	12.82	10.26	12.82	7.69	5.77	10.90		22.44	1.92	(0.3382)		
Oslo	#	17	20	22	17	12	14	15	12	13	14	156	31	3	6.5641	10.42	
	%	10.90	12.82	14.10	10.90	7.69	8.97	9.62	7.69	8.33	8.97		19.87	1.92	(0.6824)		

Panel B: Clustering tests and HHI

	X²	HHI (%)
SYS	6.3077	10.40
SE1	(0.7088)	
SE2	6.9487	10.45
SE3	(0.6425)	
SE4	6.9487	10.45
FI	(0.6425)	
DK1	5.0256	10.32
DK2	(0.8321)	
Oslo	9.1282	10.59
	(0.4255)	
	7.0769	10.45
	(0.6291)	
	10.1538	10.65
	(0.3382)	
	6.5641	10.42
	(0.6824)	

Panel A: Distribution of the last digit or last two digits of the price.															Panel B: Clustering tests and HHI	
Last Digit		0	1	2	3	4	5	6	7	8	9	Total	0 & 5	00	X ²	HHI (%)
Kr.sand	#	18	21	22	17	17	9	15	6	17	14	156	27	3	14.1282	10.91
	%	11.54	13.46	14.10	10.90	10.90	5.77	9.62	3.85	10.90	8.97		17.31	1.92	(0.1178)	
Bergen	#	12	23	22	20	11	12	16	9	20	11	156	24	1	15.7949	11.01
	%	7.69	14.74	14.10	12.82	7.05	7.69	10.26	5.77	12.82	7.05		15.38	0.64	(0.0713)	
Molde	#	24	21	14	12	20	15	13	12	16	9	156	39	1	12.7179	10.82
	%	15.38	13.46	8.97	7.69	12.82	9.62	8.33	7.69	10.26	5.77		25.00	0.64	(0.1758)	
Tr.heim	#	24	21	14	12	20	15	13	12	16	9	156	39	1	12.7179	10.82
	%	15.38	13.46	8.97	7.69	12.82	9.62	8.33	7.69	10.26	5.77		25.00	0.64	(0.1758)	
Tromsø	#	19	10	14	17	20	17	13	12	19	15	156	36	0	6.4359	10.41
	%	12.18	6.41	8.97	10.90	12.82	10.90	8.33	7.69	12.18	9.62		23.08	0.00	(0.6956)	
EE	#	14	17	21	21	17	10	12	15	11	18	156	24	2	8.7436	10.56
	%	8.97	10.90	13.46	13.46	10.90	6.41	7.69	9.62	7.05	11.54		15.38	1.28	(0.4613)	
LV	#	22	7	12	17	17	15	19	19	16	12	156	37	1	10.7949	10.69
	%	14.10	4.49	7.69	10.90	10.90	9.62	12.18	12.18	10.26	7.69		23.72	0.64	(0.2900)	
LT	#	19	14	11	16	11	15	22	17	18	13	156	34	2	7.2051	10.46
	%	12.18	8.97	7.05	10.26	7.05	9.62	14.10	10.90	11.54	8.33		21.79	1.28	(0.6158)	
AT	#	4	2	2	4	2	1	3	4	2	1	25	5	0	n/a	12.00
	%	16.00	8.00	8.00	16.00	8.00	4.00	12.00	16.00	8.00	4.00		20.00	0.00	n/a	

Panel A: Distribution of the last digit or last two digits of the price.															Panel B: Clustering tests and HHI	
Last Digit		0	1	2	3	4	5	6	7	8	9	Total	0 & 5	00	X ²	HHI (%)
BE	#	2	3	3	3	2	1	2	3	3	3	25	3	1	n/a	10.72
	%	8.00	12.00	12.00	12.00	8.00	4.00	8.00	12.00	12.00	12.00		12.00	4.00	n/a	
DE-LU	#	0	2	4	2	3	1	3	4	2	4	25	1	0	n/a	12.64
	%	0.00	8.00	16.00	8.00	12.00	4.00	12.00	16.00	8.00	16.00		4.00	0.00	n/a	
FR	#	2	4	4	2	3	1	2	1	4	2	25	3	0	n/a	12.00
	%	8.00	16.00	16.00	8.00	12.00	4.00	8.00	4.00	16.00	8.00		12.00	0.00	n/a	
NL	#	1	4	1	0	3	2	2	3	3	6	25	3	0	n/a	14.24
	%	4.00	16.00	4.00	0.00	12.00	8.00	8.00	12.00	12.00	24.00		12.00	0.00	n/a	
All	#	308	285	297	311	276	246	260	250	285	259	2777	554	27	17.8037	10.06
	%	11.09	10.26	10.69	11.20	9.94	8.86	9.36	9.00	10.26	9.33		19.95	0.97	(0.0375***)	

Table 4.3 Panel A presents the frequency table regarding Nord Pool weekly prices decimal values, for all the available bidding areas. The symbol # represents the absolute frequency, while % represents the relative frequency.

Names of the bidding areas are displayed by their initials, where: SYF stands for the system price; SE4, SE2, SE3 and SE4 refer to Sweden; FI refers to Finland; DK1 and DK2 refer to Denmark; Oda, Kr.sund, Bergen, Molde, Trondheim and Tromsø refer to cities in Norway; EE stands for Estonia; LV refers to Latvia; LT refers to Lithuania; AT refers to Austria; BE refers to Belgium; DE-LU refers to Germany; FR refers to France; NL refers to Netherlands.

Panel B of Table 4.3 includes the tests price clustering tests applied for each bidding area. HHI stands for Hirschman-Herfindahl Index, while X² is the Chi-squared goodness-of-fit test, with the respective p-value displayed below in brackets. Note that for the bidding areas of AT, BE, DE-LU, FR and NL the Chi-squared goodness-of-fit test could not be applied, as the number of available observations is not sufficient.

*** represents values statistically significant for a level of confidence of 99%.

** represents values statistically significant for a level of confidence of 95%.

* represents values statistically significant for a level of confidence of 90%.

The broader timeframe reduces the number of available observations, which can introduce additional limitations to the conclusions that can be taken. Thus, even though most bidding areas show a certain abnormal concentration of prices around specific numbers, as seen in Panel A from Table 4.3, no generalized inference is available.

In Table 4.3 Panel B, due to the small number of observations, the Chi-squared goodness-of-fit test could not be applied to the bidding areas of *AT*, *BE*, *DE-LU*, *FR* and *NL*. For the remaining bidding areas, results from the test support the previous conclusions, as only *Bergen* presents enough evidence to reject the null hypothesis for a 90% confidence level. The aggregate prices from all bidding areas still show sufficient statistical evidence to reject the null hypothesis of a uniform price distribution, this time for a confidence level of 95%.

As proceeded for the hourly and daily timeframes, a temporal analysis of the cumulative number of occurrences for each hundredth number, in the weekly prices for the entire sample, is displayed in the chart of Figure 4.3.

Figure 4.3 - Cumulative number of occurrences of each number, for Nord Pool day-ahead weekly prices of the entire sample

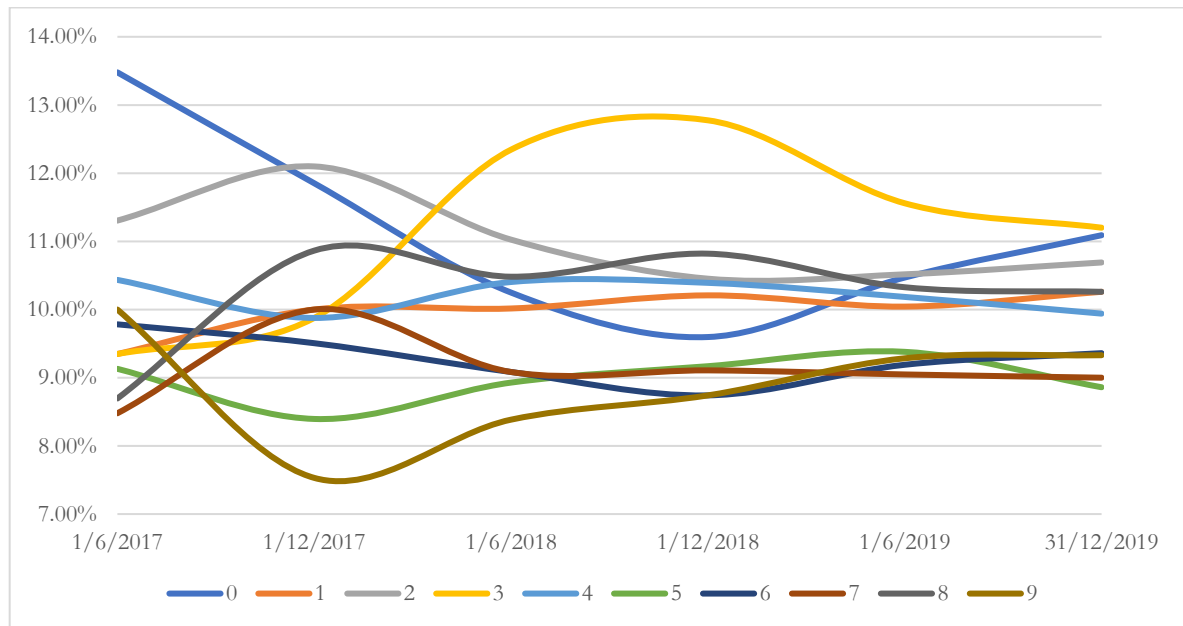


Figure 4.3 displays the cumulative number of occurrences for each number in the hundredth place of Nord Pool's day-ahead weekly prices, for the entire sample, except the system price, from January 1st 2017 to December 31st of 2019. First semester was omitted from Figure 4.2 to eliminate the initial dispersion of the sample, although cumulative price occurrences of this period are also included.

Note: Data for the bidding areas of *AT*, *BE*, *DE-LU*, *FR* and *NL* is only available starting from week 28 of the year 2019.

The lower sample size for weekly prices increases the dispersion of the relative frequencies obtained throughout the period of analysis. Prices ending in 0 and 3 are the most frequent within the sample, while those ending in XX.X5 show the lowest relative frequency, equal to 8.86% as observed in Table 4.3.

4.2. Test on Price Clustering Nature

Having obtained statistical support for the existence of price clustering in a subgroup of bidding areas, namely through the Chi-squared goodness-of-fit test, a subsequent study of the nature of such phenomenon becomes adequate in order to not only position this case within the existent literature, but also retrieve possible implications for the results obtained.

Taking into account the results obtained by the previous tests, the multivariate regressions presented in Equations 3.3 and 3.4 were applied to the entire sample for which data is available⁸, including the system price. For all the subsamples, the dependent variable *Clustering* – $E(Clustering)$ is defined as the difference between the observed and the expected relative frequency of price clustering for numbers ending in XX.X0 (expressed in percentage), as suggested by the tests performed in the previous section, and following the approach from Brooks et al. (2013) and Chung and Chiang (2006). In addition the formation of price clustering of whole numbers - i.e. numbers ending in XX.00 – was also studied, considering this time the percentage difference between the observed and expected relative frequencies for these as a dependent variable.

As previously mentioned, two multivariate regression models were applied with the difference consisting in the consideration of the log-transformed daily standard deviation as a proxy for volatility. Table 4.4 presents the results of the model without this measure, based in Equation 3.3, more consistent and comparable with similar studies. Panel A studies the formation of price clusters for numbers with 0 in the hundredth decimal place, while Panel B shows the results from the study of price clustering in whole numbers.

⁸ Volume data is not available for the bidding areas of *Oslo*, *Kr.sand*, *Bergen*, *Molde*, *Tr.heim* and *Tromsø*, so the multivariate regressions were not applied to these.

Table 4.4 - Multivariate Regression applied to Nord Pool day-ahead hourly prices, without $\log(\text{Volatility})$ as an explanatory variable

Bidding Area	Price	$\log(\text{Buy Volume})^9$	$\log(\text{Sell Volume})^{10}$	Adjusted R ²	F-stat
<i>Expected Signal</i>	+	-	-		(p-value)
<i>Panel A: Dependent variable as the difference between the observed and expected frequency (in percentage) of hourly prices ending in XX.X0</i>					
<i>SYS</i>	-0.007 (0.000 ^{***})	-0.282 (0.000 ^{***})	n/a ¹¹	0.034	153.900 (0.000 ^{***})
<i>SE1</i>	-0.009 (0.000 ^{***})	-2.824 (0.000 ^{***})	0.381 (0.000 ^{***})	0.100	326.189 (0.000 ^{***})
<i>SE2</i>	-0.011 (0.000 ^{***})	-1.244 (0.000 ^{***})	0.347 (0.000 ^{***})	0.073	231.278 (0.000 ^{***})
<i>SE3</i>	-0.011 (0.000 ^{***})	-0.121 (0.150)	-1.566 (0.000 ^{***})	0.074	233.046 (0.000 ^{***})
<i>SE4</i>	-0.009 (0.000 ^{***})	0.120 (0.007 ^{***})	-0.528 (0.000 ^{***})	0.059	183.774 (0.000 ^{***})
<i>FI</i>	0.008 (0.000 ^{***})	-1.765 (0.000 ^{***})	-0.249 (0.000 ^{***})	0.113	373.508 (0.000 ^{***})
<i>DK1</i>	-0.022 (0.000 ^{***})	0.731 (0.000 ^{***})	-0.549 (0.000 ^{***})	0.061	191.469 (0.000 ^{***})
<i>DK2</i>	-0.001 (0.581)	-0.280 (0.000 ^{***})	-0.556 (0.000 ^{***})	0.118	392.691 (0.000 ^{***})
<i>EE</i>	0.004 (0.000 ^{***})	-0.365 (0.000 ^{***})	-0.155 (0.000 ^{***})	0.026	79.416 (0.000 ^{***})
<i>LV</i>	0.005 (0.000 ^{***})	-1.027 (0.000 ^{***})	0.126 (0.000 ^{***})	0.023	70.572 (0.000 ^{***})
<i>LT</i>	-0.008 (0.000 ^{***})	0.162 (0.007 ^{***})	-0.662 (0.000 ^{***})	0.046	141.935 (0.000 ^{***})
<i>AT</i>	-0.006 (0.015 ^{**})	0.001 (0.000 ^{***})	0.003 (0.000 ^{***})	0.048	51.074 (0.000 ^{***})
<i>BE</i>	0.008 (0.005 ^{***})	-0.001 (0.003 ^{***})	-0.001 (0.184)	0.003	4.186 (0.006 ^{***})

⁹ The bidding areas of *AT*, *BE*, *DE-LU*, *FR* and *NL* present observations for which *Buy Volume* is zero, so this variable is not log-transformed for these cases in particular.

¹⁰ The bidding areas of *AT*, *BE*, *DE-LU*, *FR* and *NL* present observations for which *Sell Volume* is zero, so this variable is not log-transformed for these cases in particular.

¹¹ For the system price *SYS*, volume is not segregated by Nord Pool in buy and sell. Instead, it is used the total turnover, represented in Table 4.4 under the column $\log(\text{Buy Volume})$ as an exception.

DE-LU	-0.010 (0.000 ^{***})	0.001 (0.000 ^{***})	0.001 (0.000 ^{***})	0.089	100.817 (0.000 ^{***})
FR	0.006 (0.009 ^{***})	0.002 (0.000 ^{***})	0.003 (0.000 ^{***})	0.157	191.284 (0.000 ^{***})
NL	0.014 (0.029 ^{**})	0.004 (0.000 ^{***})	0.000 (0.573)	0.134	159.005 (0.000 ^{***})

Panel B: Dependent variable as the difference between the observed and expected frequency (in percentage) of hourly prices ending in XX.00

<i>SYS</i>	-0.014 (0.000 ^{***})	-0.057 (0.000 ^{***})	n/a ¹²	0.329	2147.386 (0.000 ^{***})
<i>SE1</i>	-0.000 (0.243)	0.613 (0.000 ^{***})	-0.078 (0.000 ^{***})	0.064	199.105 (0.000 ^{***})
<i>SE2</i>	0.000 (0.599)	0.213 (0.000 ^{***})	0.064 (0.000 ^{***})	0.045	137.983 (0.000 ^{***})
<i>SE3</i>	0.002 (0.000 ^{***})	0.158 (0.000 ^{***})	0.486 (0.000 ^{***})	0.170	597.001 (0.000 ^{***})
<i>SE4</i>	0.001 (0.000 ^{***})	0.194 (0.000 ^{***})	0.135 (0.000 ^{***})	0.151	519.205 (0.000 ^{***})
<i>FI</i>	-0.000 (0.000 ^{***})	0.013 (0.577)	0.167 (0.000 ^{***})	0.038	115.413 (0.000 ^{***})
<i>DK1</i>	-0.010 (0.000 ^{***})	0.328 (0.000 ^{***})	-0.079 (0.000 ^{***})	0.067	211.893 (0.000 ^{***})
<i>DK2</i>	-0.001 (0.000 ^{***})	0.025 (0.129)	-0.062 (0.000 ^{***})	0.018	54.958 (0.000 ^{***})
<i>EE</i>	-0.002 (0.000 ^{***})	0.105 (0.000 ^{***})	0.055 (0.000 ^{***})	0.049	152.958 (0.000 ^{***})
<i>LV</i>	-0.002 (0.000 ^{***})	-0.092 (0.000 ^{***})	0.075 (0.000 ^{***})	0.025	76.877 (0.000 ^{***})
<i>LT</i>	-0.001 (0.000 ^{***})	0.106 (0.000 ^{***})	-0.012 (0.060 [*])	0.016	49.216 (0.000 ^{***})
<i>AT</i>	-0.002 (0.120)	0.001 (0.000 ^{***})	0.002 (0.000 ^{***})	0.048	52.480 (0.000 ^{***})
<i>BE</i>	-0.019 (0.000 ^{***})	0.007 (0.000 ^{***})	0.005 (0.000 ^{***})	0.212	277.075 (0.000 ^{***})
<i>DE-LU</i>	-0.002 (0.138)	0.000 (0.892)	0.000 (0.000 ^{***})	0.026	27.927 (0.000 ^{***})
<i>FR</i>	0.010 (0.000 ^{***})	0.001 (0.000 ^{***})	0.002 (0.000 ^{***})	0.282	403.458 (0.000 ^{***})
<i>NL</i>	0.002 (0.325)	0.002 (0.000 ^{***})	-0.001 (0.091 [*])	0.180	225.630 (0.000 ^{***})

¹² Similarly to Panel A of Table 4.4, for the system price *SYS* volume is not segregated by Nord Pool in buy and sell. Instead, it is used the total turnover, represented in Table 4.4 Panel B under the column $\log(\text{Buy Volume})$.

Table 4.4 presents the results obtained by the regression defined in Equation 3.3, without the explanatory variable $\log(Volatility)$. Both results for price clustering in the hundredth number being 0 and of whole digits are displayed in Panel A and Panel B, respectively. *Price* is defined in euros, while $\log(Buy Volume)$ and $\log(Sell Volume)$ consist in the natural log-transformation of the respective values, originally expressed in MW/h.

Names of the bidding areas are displayed by their initials, where: *SYS* stands for the system price; *SE1*, *SE2*, *SE3* and *SE4* refer to Sweden; *FI* refers to Finland; *DK1* and *DK2* refers to Denmark; *EE* stands for Estonia; *LV* refers to Latvia; *LT* refers to Lithuania; *AT* refers to Austria; *BE* refers to Belgium; *DE-LU* refers to Germany; *FR* refers to France; *NL* refers to Netherlands.

The bidding areas of *Oslo*, *Kr.sand*, *Bergen*, *Molde*, *Tr.beim* and *Tromsø* were not included in the multivariate analysis, since Nord Pool calculates the volume for these Norwegian regions through a different segregation.

Each column presents the estimated coefficient for the respective explanatory variable, followed by the correspondent p-value under brackets. The coefficient of determination R^2 and the global F-statistic with the respective p-value are displayed on the two columns in the right.

*** represents estimates that present individual statistical significance for a confidence level of 99%.

** represents estimates that present individual statistical significance for a confidence level of 95%.

* represents estimates that present individual statistical significance for a confidence level of 90%.

From a broad analysis, models do not show in general a significant explanatory capacity as seen by the relatively low coefficient of determination *Adjusted R²*. Nevertheless, such results are acceptable taking into consideration the data from similar works on price clustering, namely Brooks et al. (2013).

Regarding the variable *Price*, several bidding areas suggest a negative correlation between this and the formation of price clusters, and the overall magnitude and direction of this relation is not supportive to Ball et al. (1985). The results from Panel A and Panel B are consistent overall, with the respective coefficients showing often the same signal for both types of price clusters. Thus, regarding price, evidence found is globally less supportive to the existent literature, namely when compared to the results from Brooks et al. (2013).

Relatively to the liquidity, measured using volume as a proxy, including both $\log(Buy Volume)$ and $\log(Sell Volume)$, signs of the coefficients do not show a clear pattern, as for the same bidding area these are often contrary to each other. This result also comes as not supportive to Ball et al. (1985) price resolution hypothesis. Conclusions for the system price are similar, as for both prices ending in XX.X0 and XX.00 signals of the coefficients are contrary to the expected. In addition, conclusions remain contradictory both when analysing Panel A and Panel B results.

Table 4.5 introduces the proxy for volatility as displayed in Equation 3.4, intended to consider for the market uncertainty and its influence in the formation of price clusters at hourly prices from Nord Pool day-ahead market.

Table 4.5 - Multivariate Regression applied to Nord Pool day-ahead hourly prices, incorporating $\log(Volatility)$ as an explanatory variable

Bidding Area	Price	$\log(Volatility)$	$\log(Buy Volume)^{13}$	$\log(Sell Volume)^{14}$	Adjusted R ²	F-stat
<i>Expected Signal</i>	+	+	-	-		(p-value)
<i>Panel A: Dependent variable as the difference between the observed and expected frequency (in percentage) of hourly prices ending in XX.X0</i>						
<i>SYS</i>	-0.006 (0.000 ^{***})	0.042 (0.000 ^{***})	-0.317 (0.000 ^{***})	n/a ¹⁵	0.035	107.654 (0.000 ^{***})
<i>SE1</i>	-0.009 (0.000 ^{***})	-0.007 (0.693)	-2.815 (0.000 ^{***})	0.380 (0.000 ^{***})	0.100	244.657 (0.000 ^{***})
<i>SE2</i>	-0.012 (0.000 ^{***})	-0.074 (0.000 ^{***})	-1.235 (0.000 ^{***})	0.370 (0.000 ^{***})	0.075	178.223 (0.000 ^{***})
<i>SE3</i>	-0.013 (0.000 ^{***})	-0.105 (0.000 ^{***})	-0.017 (0.842)	-1.544 (0.000 ^{***})	0.079	188.383 (0.000 ^{***})
<i>SE4</i>	-0.009 (0.000 ^{***})	-0.139 (0.000 ^{***})	0.153 (0.001 ^{***})	-0.591 (0.000 ^{***})	0.068	161.537 (0.000 ^{***})
<i>FI</i>	0.009 (0.000 ^{***})	-0.048 (0.001 ^{***})	-1.764 (0.000 ^{***})	-0.322 (0.000 ^{***})	0.114	283.179 (0.000 ^{***})
<i>DK1</i>	-0.023 (0.000 ^{***})	-0.048 (0.101)	0.745 (0.000 ^{***})	-0.544 (0.000 ^{***})	0.061	144.302 (0.000 ^{***})
<i>DK2</i>	-0.001 (0.221)	-0.225 (0.000 ^{***})	-0.210 (0.000 ^{***})	-0.551 (0.000 ^{***})	0.132	335.385 (0.000 ^{***})
<i>EE</i>	0.005 (0.000 ^{***})	-0.020 (0.128)	-0.384 (0.000 ^{***})	-0.159 (0.000 ^{***})	0.026	60.151 (0.000 ^{***})
<i>LV</i>	0.004 (0.000 ^{***})	0.063 (0.000 ^{***})	-0.981 (0.000 ^{***})	0.137 (0.000 ^{***})	0.025	58.050 (0.000 ^{***})
<i>LT</i>	-0.009 (0.000 ^{***})	0.072 (0.000 ^{***})	0.164 (0.006 ^{***})	-0.606 (0.000 ^{***})	0.048	110.908 (0.000 ^{***})
<i>AT</i>	-0.004 (0.163)	-0.401 (0.000 ^{***})	0.001 (0.000 ^{***})	0.003 (0.000 ^{***})	0.053	43.650 (0.000 ^{***})
<i>BE</i>	0.008 (0.005 ^{***})	-0.004 (0.970)	-0.001 (0.004 ^{***})	-0.001 (0.187)	0.003	3.139 (0.014 ^{**})

¹³ The bidding areas of *AT*, *BE*, *DE-LU*, *FR* and *NL* present observations for which *Buy Volume* is zero, so this variable is not log-transformed for these cases in particular.

¹⁴ The bidding areas of *AT*, *BE*, *DE-LU*, *FR* and *NL* present observations for which *Sell Volume* is zero, so this variable is not log-transformed for these cases in particular.

¹⁵ For the system price *SYS*, volume is not segregated by Nord Pool in buy and sell. Instead, it is used the total turnover, represented in Table 4.5 under the column $\log(Buy Volume)$ as an exception.

<i>DE-LU</i>	-0.009 (0.000 ^{***})	0.158 (0.033 ^{**})	0.001 (0.000 ^{***})	0.001 (0.000 ^{***})	0.090	76.832 (0.000 ^{***})
<i>FR</i>	0.007 (0.002 ^{***})	-0.261 (0.004 ^{***})	0.002 (0.000 ^{***})	0.003 (0.000 ^{***})	0.159	145.861 (0.000 ^{***})
<i>NL</i>	0.012 (0.070 [*])	0.268 (0.233)	0.004 (0.000 ^{***})	0.001 (0.483)	0.134	119.627 (0.000 ^{***})

Panel B: Dependent variable as the difference between the observed and expected frequency (in percentage) of hourly prices ending in XX.00

<i>SYS</i>	-0.015 (0.000 ^{***})	-0.023 (0.000 ^{***})	-0.038 (0.014 ^{**})	n/a^{16}	0.332	1452.029 (0.000 ^{***})
<i>SE1</i>	0.001 (0.121)	0.010 (0.027 ^{**})	0.600 (0.000 ^{***})	-0.076 (0.000 ^{***})	0.064	150.610 (0.000 ^{***})
<i>SE2</i>	0.004 (0.244)	0.017 (0.000 ^{***})	0.210 (0.000 ^{***})	0.059 (0.000 ^{***})	0.046	107.661 (0.000 ^{***})
<i>SE3</i>	0.001 (0.000 ^{***})	-0.040 (0.000 ^{***})	0.197 (0.000 ^{***})	0.494 (0.000 ^{***})	0.183	490.811 (0.000 ^{***})
<i>SE4</i>	0.001 (0.000 ^{***})	-0.044 (0.000 ^{***})	0.204 (0.000 ^{***})	0.115 (0.000 ^{***})	0.169	445.794 (0.000 ^{***})
<i>FI</i>	0.000 (0.708)	-0.042 (0.000 ^{***})	0.014 (0.540)	0.104 (0.000 ^{***})	0.060	140.670 (0.000 ^{***})
<i>DK1</i>	-0.000 (0.803)	0.001 (0.000 ^{***})	0.003 (0.000 ^{***})	-0.001 (0.000 ^{***})	0.071	169.587 (0.000 ^{***})
<i>DK2</i>	-0.001 (0.002 ^{***})	0.026 (0.000 ^{***})	0.016 (0.314)	-0.063 (0.000 ^{***})	0.020	46.573 (0.000 ^{***})
<i>EE</i>	-0.001 (0.000 ^{***})	-0.050 (0.000 ^{***})	0.057 (0.000 ^{***})	0.045 (0.000 ^{***})	0.077	184.291 (0.000 ^{***})
<i>LV</i>	-0.001 (0.000 ^{***})	-0.044 (0.000 ^{***})	-0.124 (0.000 ^{***})	0.068 (0.000 ^{***})	0.046	106.855 (0.000 ^{***})
<i>LT</i>	-0.001 (0.000 ^{***})	-0.056 (0.000 ^{***})	0.105 (0.000 ^{***})	-0.056 (0.000 ^{***})	0.046	106.662 (0.000 ^{***})
<i>AT</i>	-0.002 (0.082 [*])	0.043 (0.314)	0.001 (0.000 ^{***})	0.002 (0.000 ^{***})	0.048	39.613 (0.000 ^{***})
<i>BE</i>	-0.019 (0.000 ^{***})	0.364 (0.000 ^{***})	0.007 (0.000 ^{***})	0.005 (0.000 ^{***})	0.219	216.606 (0.000 ^{***})
<i>DE-LU</i>	-0.001 (0.273)	0.100 (0.006 ^{***})	0.000 (0.896)	0.000 (0.000 ^{***})	0.028	22.860 (0.000 ^{***})
<i>FR</i>	0.008 (0.000 ^{***})	0.291 (0.000 ^{***})	0.001 (0.000 ^{***})	0.002 (0.000 ^{***})	0.289	313.056 (0.000 ^{***})
<i>NL</i>	0.002 (0.532)	0.109 (0.197)	0.002 (0.000 ^{***})	-0.000 (0.127)	0.180	169.676 (0.000 ^{***})

¹⁶ Similarly to Panel A of Table 4.5, for the system price *SYS* volume is not segregated by Nord Pool in buy and sell. Instead, it is used the total turnover, represented in Table 4.5 Panel B under the column $\log(\text{Buy Volume})$.

Table 4.5 presents the results obtained by the regression defined in Equation 3.4, with the explanatory variable $\log(Volatility)$ included. Both results for price clustering in the hundredth number being 0 and of whole digits are displayed in Panel A and Panel B, respectively. *Price* is defined in euros, while $\log(Buy\ Volume)$ and $\log(Sell\ Volume)$ consist in the natural log-transformation of the respective values, originally expressed in MW/h. $\log(Volatility)$ expresses the natural log-transformation of the daily standard deviation.

Names of the bidding areas are displayed by their initials, where: *SYS* stands for the system price; *SE1*, *SE2*, *SE3* and *SE4* refer to Sweden; *FI* refers to Finland; *DK1* and *DK2* refers to Denmark; *EE* stands for Estonia; *LV* refers to Latvia; *LT* refers to Lithuania; *AT* refers to Austria; *BE* refers to Belgium; *DE-LU* refers to Germany; *FR* refers to France; *NL* refers to Netherlands.

The bidding areas of *Oslo*, *Kr.sand*, *Bergen*, *Molde*, *Tr.heim* and *Tromsø* were not included in the multivariate analysis, since Nord Pool calculates the volume for these Norwegian regions through a different segregation.

Each column presents the estimated coefficient for the respective explanatory variable, followed by the correspondent p-value under brackets. The coefficient of determination R^2 and the global F-statistic with the respective p-value are displayed on the two columns in the right.

*** represents estimates that present individual statistical significance for a confidence level of 99%.

** represents estimates that present individual statistical significance for a confidence level of 95%.

* represents estimates that present individual statistical significance for a confidence level of 90%.

Conclusions taken for Table 4.4 can be extrapolated to Table 4.5, as models still remain with a relatively low coefficient of determination. But once again, similar levels of overall regression quality are obtained by comparable studies, namely Brooks et al. (2013). In addition, both price and volume variables still do not suggest a direct correlation with the formation of price clusters at both hundredth decimal places and whole digits.

Considering both tables, and taking into consideration the theoretical hypothesis existent in literature, the differences between the expected and the observed signals for the coefficient's estimates show little to no support in explaining why price clusters are formed in Nord Pool's hourly prices. Recalling Ball et al. (1985) price resolution hypothesis, higher levels of price would theoretically lead agents to rely on coarser levels of price resolution, making them choose rounded values. In the same study, the level of liquidity is pointed to have a contrarian effect to this phenomenon, as higher levels of volume proportionate more probabilities for agents to trade at their desired quote. Regarding to price, the coefficient estimate shows a negative sign on a significant number of bidding areas, when analysing both frequency of prices ending in a 0 at the hundredth decimal place or when analysing price clustering in whole digits. When focusing on the difference between the observed and expected frequency of prices ending in 0 at the hundredth decimal place, estimates for the volume coefficients support in some cases the theoretical hypothesis that propose a negative sign. Nevertheless, the behaviour from the variables $\log(Buy\ Volume)$ and $\log(Sell\ Volume)$ show discrepancies between each other, as

contrary signs for these coefficients often appear within the same bidding area. In the Panel B of Table 4.5, the conclusions regarding the impact of volume are still applicable, but this time are even less significant in terms of magnitude.

Conclusions when volatility is introduced into the analysis remain mixed and not supportive to the theoretical hypothesis, as seen in Table 4.5. Retrieving once again Ball et al. (1985) price resolution hypothesis, higher levels of volatility are expected to be caused mostly by an increase of investors' uncertainty, everything else held constant, which can lead them to use a lower level of price resolution and less accurate values. Expecting a positive sign for the coefficient estimator, results in general do not support clearly this theoretical hypothesis, both for Panel A and B of Table 4.5 since even though one bidding area might present a positive sign for this coefficient in one test, the contrary is often verified on the other panel.

5. Conclusion

The energy markets are essential for many stakeholders, going from the buyers and sellers of energy to the household or company that uses it as the final consumer. Investors depend also on this utility market for several purposes, using for example derivative instruments to either diversify their portfolio, achieve exposure to this asset, or hedge their position against certain outcomes. Such high importance raises the need for a deep understanding of how the energy markets such as Nord Pool functions, in order to eventually adopt public policies and take regulatory measures to promote the maximum social and environmental welfare. One way of doing so passes by the comparison of markets such as Nord Pool, a major energy market in Europe, with the characteristics of other financial markets already deeply studied such as the stock and the currency markets. Thereby, the study presented assumes an important contribute to the financial anomalies' academic literature, at the moment of writing still scarce regarding the study of price clustering in utilities markets. Besides, the relevance and impact that Nord Pool has in a wide array of stakeholders make the study of its efficiency even more important, a purpose that the present study also aims to fulfil.

Nord Pool prices present in several bidding areas evidence that suggests the existence of price clusters, namely in the hourly timeframe. Curcio and Goodhart (1991) attraction hypothesis suggests that each ending figure as a different “gravitational” strength within human minds – with numbers ending in 0 being the strongest - that makes them more compelling by facilitating the required mental calculations. Evidence found supports to a certain degree such expectation, as hourly prices show a relatively higher number of frequencies for both numbers ending in 0 at the hundredth decimal place and whole prices ending in XX.00. The presence of number 5, the second “strongest” number according to Curcio and Goodhart (1991), shows already no direct support from the results obtained.

When analysing the nature of the phenomenon through a multivariate regression for bidding areas for which the sample suggested that the formation of price clusters was statistically significant, namely through the Chi-squared goodness-of-fit test, evidence was not clearly supportive to the theoretical hypothesis. In particular, Ball et al. (1985) price resolution hypothesis, pioneer in the literature and the foundation of Harris (1991) negotiation hypothesis,

suggested a positive impact of the price level and volatility in the formation of price clusters and a negative impact of liquidity. Analysing the results from the regression models developed, clustering in hourly prices does not seem to be formed due to these explanations, meaning that other dimensions are into taking place that distinguishes Nord Pool from other financial markets.

While studying the formation of price clustering in Nord Pool day-ahead market, it is essential to understand how price is first formed. Despite functioning as an auction, where each party – the buyers and the sellers of energy – set their respective bid and ask quotes (set in euros per MW/h) to trade in the following day, two important dynamics must be highlighted. On the first moment, after the quotes are settled, all these are converted into euros at the prevailing exchange rate, thus introducing a distortion factor that, as Nord Pool itself states, introduces sometimes significant gaps between the price set by each party and the price that effectively takes place. Secondly, the algorithm used to define the final prices for each bidding area as well as the system price (Appendix C. PCR EUPHEMIA) must take into account several factors such as the volume offered and demanded for each area, as well as other constraints specific to the electricity market. This introduces a non-human agent in the middle of the price settlement process, that in practice is not subject to the same biases as we should expect if this task was performed by a human intermediary.

Both these dimensions reveal that there is a long path between the quotes set by each counterparty and the last settlement price that comes out publicly. The introduction of bid and ask quotes to the multivariate regression models could potentially change the outcome of the results, but Nord Pool itself stated that such data was not publicly available. Brooks et al. (2013) did so to the Australian water market, finding that buy and sell bids from traders show similarities to the ones from other financial markets more developed, namely regarding the existence of price clustering.

Overall, the magnitude of the results conclusions can differ according to the perspective adopted. On one hand, despite the price dynamics involved, there is still evidence for some bidding areas of price clustering at the hourly timeframe, which can influence how market participants set their quotes at the exchange. Nevertheless, from the perspective of an external

investor that gets most of his/her exposure via derivative instruments, no relevant implications are expected as the system price, the major reference price of the Nord Pool exchange, does not present significant statistical evidence that compromises the assumption taken by most traditional valuation models according to which prices of the underlying asset follow a uniform distribution.

The introduction of Nord Pool aimed from the beginning to promote a more efficient way of exchanging energy between producers and suppliers, and its continuous growth namely through Europe means that it is reaching more and more stakeholders. Thus, it is naturally expected to see more academics concerned about the practical implications of this fairly new system of transacting energy, and a closer attention to its characteristics and dynamics will continue to be essential for public policymakers, regulators and market participants to guarantee its well-functioning.

Glossary

TSO – Transmission System Operator

NEMO – Nominated Electricity Market Operator

SDAC – Single Day-ahead Coupling

PCR – Price Coupling of Regions

CET – Central European Time

MW – Megawatt

TW - Terawatt

MW/h – Megawatts per hour

TW/h – Terawatts per hour

EUR – Euros

NOK – Norwegian Krone

SEK – Swedish Krona

DKK – Danish Krone

Appendix

Appendix A – Day-Ahead Market Orders

- a. Single hourly orders: comprehends the largest percentage of trades. Each member specifies his buying or selling volume (in MW) for each hour and can choose between a price dependent or a price independent order. Price independent orders must be defined at least with two price steps, at the minimum price (EUR -500) and at the maximum price (EUR 3,000), and in this case the participant will sell/receive the volume requested for each hour at any price between these extremes. On the other hand, a price dependent order is more suited for participants who want to trade different volumes in each hour, depending on the price practiced. In this second option, the participant establishes then the amount of volume to buy/sell, segregated by each hour and price level (which must also be comprehended between the previous EUR -500 and EUR 3,000 levels). In order to build the supply and demand curves Nord Pool uses a linear interpolation that generates a linear relationship between price and volume, and for every price-volume combination a linear curve segment is created. Note that for prices between the ranges established by the participants the volume traded will result from this linear interpolation. For example, in a price dependent order, if the participant wants to acquire 100 MW at the price of EUR 10 and acquire 50 at the price of EUR 20, if the market price is equal to EUR 15 he will only acquire 75 MW. This does not happen for the price independent order, since the participant establishes a unique amount of volume for every price between the limits mentioned.
- b. Block orders: similar to single hourly orders, but in this case the participant specifies the volume and price only for a certain number of consecutive hours, being more suited for those who want to have access to power for intervals larger than one hour since it allows to minimise costs related to the interruptions that would otherwise occur. There are four main types of block orders in Nord Pool markets:
 - i. Regular block orders: have an “all-or-nothing” condition, being fully rejected or fully accepted, in which case the contract covers the entire period and volume defined. The acceptance or rejection of these orders depend on whether the price

is higher, lower or equal than the average day-ahead price for those hours, with a symmetrical criterion whether it is a sales block or purchase block.

- ii. Linked block orders: consists in the linking of two or more block orders, where one order acceptance depends on the acceptance of another. Note that all linked blocks need to belong to the same bidding portfolio, and one bidding area.
 - iii. Curtailable block orders: correspond to block orders that can be partially accepted by the participant, according to a pre-specified Minimum Acceptance Ratio (MAR). For example: a block with $MAR = 1$ (100%) corresponds to a regular “all-or-nothing” block; a block with $MAR = 0.5$ (50%) will be accepted if at least half of the specified volume is traded; a block with $MAR = 0$ (0%) will be accepted independently of the specified volume traded. Note that MAR should be unique for all hours, and curtailable blocks can also be linked.
 - iv. Profile block orders: this option allows participants to specify a different volume over the time span of the block, with the similar possibility to set a MAR and link several blocks. In order to accept this order the volume-weighted average price is calculated, and compared with the volume-weighted average day-ahead price for the periods for which the block is defined.
- c. Exclusive group: defined as a cluster of sell and/or buy blocks, from which only one block can be activated. These blocks in the exclusive group cannot be linked with other block orders. A Minimum Acceptance Ratio can be set, but this must be higher than 50%.
 - d. Flexi orders: correspond to a block order with maximum duration of 23 consecutive hours, comprehended in the same day. In this order, the starting hour of the trading block is not defined by the user but instead by an algorithm that aims to provide the best social welfare.

Appendix B - Intraday Market Orders

- a. Limit orders: similarly to other financial markets, a limit order allows the participant to set a level for which their actions will be triggered. In a buy limit order the participant sets a price and the order will take place if the price reaches that level or below. Symmetrically, sell limit orders are executed at the price set of values above.
- b. User-defined block orders: consists in a block order with one or several consecutive hourly products, with a “all-or-nothing” regime.
- c. Predefined block orders: correspond to block orders, specific to the United Kingdom region.
- d. Iceberg orders (IBOs): type of limit order suited for large volume transactions, allowing to divide the full size of the orders into smaller ones, that will only show in the market after the previous tranche is fully matched.
- e. Execution constraints: just like other exchanges, limit orders can have additional constraints.
 - i. Fill-or-Kill (FoK): corresponds to a limit order which is withdrawn from the market if the entire volume is not immediately fulfilled.
 - ii. Immediate-or-Cancel (IoC): similar to the previous, but in this case only the non-fulfilled portion of volume is withdrawn from the market.

Appendix C - PCR EUPHEMIA

Europe has been developing several efforts throughout recent history in order to develop and integrate their power markets. The Single Day-ahead Coupling (SDAC) surged as an initiative between the Nominated Electricity Market Operators (NEMOs) and Transmission System Operators (TSOs), with the objective to enable cross-border trading across Europe via auction-based processes for delivery of power for the following day. SDAC intervention can be divided into three main phases: process inputs from all involved NEMOs and TSOs (consisting of bids, offers and network capacities/constraints); match these inputs using a common algorithm; validate and send the outputs (matched trades, clearing prices and scheduled exchanges) back to NEMOs and TSOs.

For this process SDAC relies on the Price Coupling of Regions (PCR) solution, previously developed by a group of eight power exchanges (EPEX SPOT, GME, HEnEx, Nord Pool, OMIE, OPCOM, OTE and TGE). From this project, the common price coupling algorithm PCR EUPHEMIA (acronym for Pan-European Hybrid Electricity Market Integration Algorithm) was created and continues to be further developed. The complex nature of this algorithm seeks to fulfil several objectives, namely to generate a fair and transparent determination of day-ahead electricity prices, energy allocation and net positions across Europe, properly taking into account specific features of the various power markets and the electricity network constraints. In sum, with this project it is intended to maximize overall social welfare and increase the transparency of the process, consequently increasing the efficiency of the entire European power market.

The exhaustive description of the algorithm is published by the NEMO Committee (Committee, 2019a).

References

- Aitken, M., Brown, P., Buckland, C., Izan, H. Y., & Walter, T. (1996). Price clustering on the Australian Stock Exchange. *Pacific-Basin Finance Journal*, 4, 297-314. doi:10.1016/0927-538X(96)00016-9
- Almeida, C. M. N. d. (2018). *Barreiras psicológicas no mercado elétrico : estudo aplicado ao mercado Nord Pool*. (Master in Economics), Universidade de Aveiro,
- Alves Luís, M. (2018). *Price clustering in bank stocks during the global financial crisis*. Porto: [Edição do autor].
- ap Gwilym, O., & Alibo, E. (2003). Decreased price clustering in FTSE100 futures contracts following a transfer from floor to electronic trading. *Journal of Futures Markets*, 23(7), 647-659. doi:10.1002/fut.10080
- Aşçıoğlu, A., Comerton-Forde, C., & McInish, T. H. (2007). Price Clustering on the Tokyo Stock Exchange. *Financial Review*, 42(2), 289-301. doi:10.1111/j.1540-6288.2007.00172.x
- Ball, C., Torous, W., & Tschoegl, A. (1985). The Degree of Price Resolution: The Case of the Gold Market. *Journal of Futures Markets*, 5, 29-43. doi:10.1002/fut.3990050105
- Blau, B. M. (2019). Price Clustering and Investor Sentiment. *Journal of Behavioral Finance*, 20(1), 19-30. doi:10.1080/15427560.2018.1431887
- Booth, G., & Yuksel, A. (2006). Price resolution in an emerging market: Evidence from the Istanbul Stock Exchange. *European Journal of Finance*, 12, 137-152. doi:10.1080/13518470500146017
- Botterud, A., Kristiansen, T., & Ilic, M. D. (2010). The relationship between spot and futures prices in the Nord Pool electricity market. *Energy Economics*, 32(5), 967-978. doi:<https://doi.org/10.1016/j.eneco.2009.11.009>
- Brooks, R., Harris, E., & Joymungul, Y. (2013). Price clustering in Australian water markets. *Applied Economics*, 45(6), 677-685. doi:10.1080/00036846.2011.610747
- Brown, P., Chua, A., & Mitchell, J. (2002). The influence of cultural factors on price clustering: Evidence from Asia-Pacific stock markets. *Pacific-Basin Finance Journal*, 10(3), 307-332. doi:[https://doi.org/10.1016/S0927-538X\(02\)00049-5](https://doi.org/10.1016/S0927-538X(02)00049-5)
- Brown, P., & Mitchell, J. (2008). Culture and stock price clustering: Evidence from The Peoples' Republic of China. *Pacific-Basin Finance Journal*, 16(1), 95-120. doi:<https://doi.org/10.1016/j.pacfin.2007.04.005>
- Christie, W. G., Harris, J. H., & Schultz, P. H. (1994). Why did NASDAQ Market Makers Stop Avoiding Odd-Eighth Quotes? *The Journal of Finance*, 49(5), 1841-1860. doi:10.2307/2329273
- Christie, W. G., & Schultz, P. H. (1994). Why do NASDAQ Market Makers Avoid Odd-Eighth Quotes? *The Journal of Finance*, 49(5), 1813-1840. doi:10.2307/2329272

- Chung, H., & Chiang, S. (2006). Price clustering in E-mini and floor-traded index futures. *Journal of Futures Markets*, 26(3), 269-295. doi:10.1002/fut.20196
- Comerton-Forde, C., & Putniņš, T. J. (2013). Stock Price Manipulation: Prevalence and Determinants*. *Review of Finance*, 18(1), 23-66. doi:10.1093/rof/rfs040
- Committee, N. (2019a). EUPHEMIA Public Description. Retrieved from <https://www.nordpoolgroup.com/globalassets/download-center/pcr/euphemia-public-description.pdf>
- Committee, N. (2019b). Single Day-Ahead Coupling (SDAC). Retrieved from <http://www.nemo-committee.eu/sdac>
- Curcio, R., & Goodhart, C. (1991). The Clustering of Bid/Ask Prices and the Spread in the Foreign Exchange Market. *Financial Markets Group*(dp110).
- Fama, E. F. (1970). Efficient Capital Markets: A Review of Theory and Empirical Work. *The Journal of Finance*, 25(2), 383-417. doi:10.2307/2325486
- Geoffrey Booth, G., Kallunki, J.-P., Lin, J.-C., & Martikainen, T. (2000). Internalization and stock price clustering: Finnish evidence. *Journal of International Money and Finance*, 19(5), 737-751. doi:[https://doi.org/10.1016/S0261-5606\(00\)00022-X](https://doi.org/10.1016/S0261-5606(00)00022-X)
- Grossman, S. J., Miller, M. H., Cone, K. R., Fischel, D. R., & Ross, D. J. (1997). Clustering and Competition in Asset Markets. *The Journal of Law & Economics*, 40(1), 23-60. doi:10.1086/467365
- Harris, L. (1991). Stock Price Clustering and Discreteness. *The Review of Financial Studies*, 4(3), 389-415. doi:10.1093/rfs/4.3.389
- Haugom, E., Westgaard, S., Solibakke, P. B., & Lien, G. (2011). Realized volatility and the influence of market measures on predictability: Analysis of Nord Pool forward electricity data. *Energy Economics*, 33(6), 1206-1215. doi:<https://doi.org/10.1016/j.eneco.2011.01.013>
- Hellström, J., Lundgren, J., & Yu, H. (2012). Why do electricity prices jump? Empirical evidence from the Nordic electricity market. *Energy Economics*, 34(6), 1774-1781. doi:<https://doi.org/10.1016/j.eneco.2012.07.006>
- Hu, B., Jiang, C., McInish, T., & Zhou, H. G. (2017). Price clustering on the Shanghai Stock Exchange. *Applied Economics*, 49(28), 2766-2778. doi:10.1080/00036846.2016.1248284
- Huang, R. D., & Stoll, H. R. (2001). Tick Size, Bid-Ask Spreads, and Market Structure. *The Journal of Financial and Quantitative Analysis*, 36(4), 503-522. doi:10.2307/2676222
- Kristiansen, T. (2012). Forecasting Nord Pool day-ahead prices with an autoregressive model. *Energy Policy*, 49, 328-332. doi:<https://doi.org/10.1016/j.enpol.2012.06.028>
- Lobão, J., Pacheco, L., & Alves, L. (2019). Price Clustering in Bank Stocks During the Global Financial Crisis. *Scientific Annals of Economics and Business*, 66, 465-486. doi:10.2478/saeb-2019-0043
- Määtä, T., & Johansson, T. (2011). *The System Price of Electricity on Nord Pool : A Matter of Fundamental Factors?*, Umeå University,

- Mitchell, J. (2001). Clustering and psychological barriers: the importance of numbers. *Journal of Futures Markets*, 21(5), 395-428. doi:10.1002/fut.2
- Narayan, P. K., Narayan, S., Popp, S., & D'Rosario, M. (2011). Share price clustering in Mexico. *International Review of Financial Analysis*, 20(2), 113-119. doi:<https://doi.org/10.1016/j.irfa.2011.02.003>
- Narayan, P. K., & Smyth, R. (2013). Has political instability contributed to price clustering on Fiji's stock market? *Journal of Asian Economics*, 28, 125-130. doi:<https://doi.org/10.1016/j.asieco.2013.07.002>
- Nord Pool - Bidding Areas. (2019). Retrieved from <https://www.nordpoolgroup.com/the-power-market/Bidding-areas/>
- Osborne, M. F. M. (1962). Periodic Structure in the Brownian Motion of Stock Prices. *Operations Research*, 10(3), 345-379.
- Palmon, O., Smith, B., & Sopranzetti, B. (2004). Clustering in Real Estate Prices: Determinants and Consequences. *Journal of Real Estate Research*, 26, 115-136.
- Parkinson, M. (1980). The Extreme Value Method for Estimating the Variance of the Rate of Return. *The Journal of Business*, 53(1), 61-65.
- Pool, N. (2019). Nord Pool Website. Retrieved from www.nordpoolgroup.com
- Simonsen, I. (2005). Volatility of power markets. *Physica A: Statistical Mechanics and its Applications*, 355(1), 10-20. doi:<https://doi.org/10.1016/j.physa.2005.02.062>
- Sonnemans, J. (2006). Price clustering and natural resistance points in the Dutch stock market: A natural experiment. *European Economic Review*, 50(8), 1937-1950. doi:<https://doi.org/10.1016/j.eurocorev.2005.09.001>
- Sopranzetti, B. J., & Datar, V. (2002). Price clustering in foreign exchange spot markets. *Journal of Financial Markets*, 5(4), 411-417. doi:[https://doi.org/10.1016/S1386-4181\(01\)00032-5](https://doi.org/10.1016/S1386-4181(01)00032-5)
- Urquhart, A. (2017). Price clustering in Bitcoin. *Economics Letters*, 159, 145-148. doi:10.1016/j.econlet.2017.07.035
- Waterpool. (2019). Retrieved from <https://www.waterpool.org.au/>
- Weron, R., & Misiorek, A. (2008). Forecasting spot electricity prices: A comparison of parametric and semiparametric time series models. *International Journal of Forecasting*, 24(4), 744-763. doi:<https://doi.org/10.1016/j.ijforecast.2008.08.004>
- Weron, R., & Zator, M. (2014). Revisiting the relationship between spot and futures prices in the Nord Pool electricity market. *Energy Economics*, 44, 178-190. doi:<https://doi.org/10.1016/j.eneco.2014.03.007>