Analysis of Key Drivers of Trading Performance

Bogdan Batrinca

A dissertation submitted in partial fulfilment
of the requirements for the degree of
Doctor of Philosophy
of
University College London

Department of Computer Science
University College London

2016

Declaration

I, Bogdan Batrinca, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

……………………………………………

Bogdan Batrinca

Abstract

This thesis is an applied study for understanding the key factors of trading volume, providing an in-depth investigation of liquidity demand and market impact. This research was conducted in collaboration with Deutsche Bank and presents a series of empirical studies, which examine several underlying factors affecting trading volume when executing orders algorithmically in the European equity markets, which ultimately translates into trading performance and liquidity modelling. This addresses several aspects: the size of the liquidity demand relative to the predicted or actual volumes traded in the market; the choice of execution strategy given the liquidity properties of the stock; and the timing of the trade, given the market circumstances and released or anticipated company news. All of these reflect the investment skill of the portfolio manager and the execution skill of the trader, and to some extent the quality of the execution algorithm being used.

The motivation is to investigate various factors that adversely affect the trading performance of algorithms, causing them to have excessive market impact or to under-participate when the market experiences periods of higher volatility. Although measuring the market fairness and efficiency is a crucial component for understanding the execution style and improving trading performance, the research into how to model and decompose trading performance requires further investigation. Trading volumes are a benchmark for determining an appropriate order size so as not to have excessive liquidity demand and therefore it is important to model well and accurately predict the volumes. The problem of sizing an order affects portfolio allocation and trade planning for multi-day trades, as well as intraday slices. To this end, four studies were conducted for time series analysis using machine learning methods based primarily on feature selection and regression. To achieve this, the thesis starts with a broad exploration of the trading volume drivers, followed by in-depth analyses of the effects of notable events, and concludes by proposing a volume prediction modelling framework.

The thesis consists of the following four studies:

1. Examining Drivers of Trading Volume. The first study is an in-sample volume analysis, which explores the market dynamics and identifies a series of drivers of trading volume based on the price-volume relation, lagged time series market data, the day-of-the-week effect, and a novel approach on the price-volume asymmetric relation.
2. European Trading Volumes on Cross-Market Holidays. The second study further extends the exploration of the drivers of trading volume and investigates the anecdotal evidence of lower trading volumes when other markets are not trading (i.e. 'the cross-market holiday effect'). The analysis considers the phenomenon in conjunction with the weekend effect, while indicating that the cross-market holidays are the real driver of the lower volumes on Mondays, and examines other aspects like lagged volumes, market capitalisation or multi-step ahead modelling.
3. Effect Expiry Day Effects on European Trading Volumes. This study examines the impact of sparse periodic events, such as stock index futures expiries and MSCI quarterly rebalances, on trading volume. The analysis explores anticipatory and subsequent effects of the index expiry and review dates. It investigates the main drivers of volume surges by discriminating between the Friday effect and the stock index futures expiries, and between the end-of-month effect and the MSCI quarterly reviews.

4. Developing a Volume Forecasting Model. The final study of this thesis incorporates the findings of the previous in-sample studies and provides an out-of-sample trading volume analysis, exploring the behaviour of time series variables in the context of volume prediction modelling, with seven statistical methods that are fit using the sliding and growing window approaches. The primary objective of the prediction model we propose in this final study is to achieve optimal accuracy in predicting the size of a trade given the market context, by proposing a dynamic model that switches between different models based on the temporal context. Finally, a stock-specific out-of-sample metamodel is constructed based on the recent performance of the initial stock-specific models that are independently fit.

The thesis presents the following contributions to science:

1. Detailed exploration of trading volume drivers. The main objective of this thesis is to investigate the causal factors of trading volume. Salient drivers of trading volume include the lagged time series, the price-volume relation and its asymmetry, and temporal factors, e.g. the day-of-the-week effect, holidays and other notable dates.
2. Focus on the volume dimension. The empirical studies in this thesis focus on the trading volumes, unlike the majority of the reviewed literature, which investigates the relation between calendar effects (e.g. day-of-the-week, cross-market holiday etc.) and stock returns; few studies consider the relation between these effects and trading volumes.
3. Comprehensive pan-European stock universe. To the best of our knowledge, this research is conducted on the largest European data set in the relevant literature, covering the daily market data for 2,353 stocks from 21 European countries since 1st January 2000. Most of the relevant literature either employs small data sets or focuses mostly on the US equities market.
4. Accurate trading calendar data. Due to the unavailability of a high-precision trading calendar, we constructed a consolidated and normalised calendar with a comprehensive breadth and depth of events influencing equities across a wide range of European countries. This robust calendar covers the most liquid European exchanges' trading calendar in 21 countries and the US trading calendar (since the US is the largest financial market and its trading holidays might influence the European liquidity), the stock index futures expiries for seven liquid European indices, and the MSCI quarterly review effective dates, along with the historical evidence of leavers and joiners for each analysed index, since 2000.
5. Established statistical methods applied to a new application domain. Advanced variable selection and machine learning methods have not been typically employed in the analysis of calendar effects and trading volumes; we choose a rather different approach and apply statistics and machine learning to this application domain.
6. Further insights into calendar effects. The field of behavioural finance and its literature on calendar effects contains mixed results that are often inconclusive. This thesis sheds light on a few calendar effects, e.g. the day-of-the-week, end-of-month, holiday and expiry day effects, and examines the extent of their impact on the trading volume.
7. Research validation. The advice and expert validation of a leading investment bank confirm the industry demand and necessity of this research. Deutsche Bank drove the analyses conducted in this thesis, addressing real-world problems, such as the liquidity extraction model, the multi-day trade planning using multi-step ahead forecasting, or the effect quantification of special events on trading volume.

Bogdan Batrinca, Analysis of Key Drivers of Trading Performance

Supervisors: Prof. Philip Treleaven and Dr. Christian Hesse

Acknowledgements

Pursuing a PhD is a challenging and difficult task. I would like to overstate my gratitude to my PhD supervisors, Prof. Philip Treleaven and Dr. Christian Hesse. They made this research possible with their friendly support, patience and technical expertise.

I would like to express my gratitude to UCL for providing the high performance computing resources to conduct this research. I would like to acknowledge Thomson Reuters and Andrew Fletcher for their technical help during the data acquisition phase and for appointing me as the Thomson Reuters technical liaison at UCL.

I would like to dedicate this thesis to my family, who showed love and support throughout my life. They have persuaded me to follow a PhD degree and have guided me continuously towards a solid education.

Contents

List of Figures

List of Tables

This page intentionally left blank.

1. Introduction

The objective of this chapter is to present an overview of this thesis by discussing the motivation behind the research problem, the objectives, experiments and contributions of this study and the structure of this thesis. The chapter starts by briefly introducing background information on algorithmic trading and trading performance while showing the importance of understanding the factors of trading volume, which leads to trading performance. The chapter then outlines the objectives, experiments, and contributions of this work and concludes with the thesis structure.

This thesis investigates the liquidity demand and assesses the degree of participation in the market: over-participation leading to market impact and under-participation leading to opportunity cost and price uncertainty. The trading volume is important for accurate predictions and good modelling; it is a benchmark for determining the optimal order size in order to avoid excessive liquidity demand. Order sizing influences the portfolio allocation and trade planning for both multi-day and intraday trades, although this thesis focuses on daily data. This liquidity modelling study explores the trading volume drivers such as the market conditions (e.g. price levels or fragmentation) or noticeable calendar events, and then proposes a volume prediction framework.

The majority of this study was conducted in collaboration with Deutsche Bank under the supervision of Dr. Christian Hesse, facilitating expert guidance and validation, and confirming the market demand for this study.

1.1 Motivation

Algorithmic trading (AT) is the business area consisting of trading systems that use both simple and advanced mathematical models for making optimal transaction decisions while executing client orders. The AT systems decide the timing, pricing and sizes of the buy or sell orders depending on the algorithm's parameters (commonly supported parameters include start/end times, duration, must-be-filled, execution style, limit price, maximum/minimum/child volume limit, and auction flags) and client constraints, aiming to cause the least amount of impact on a stock's price and to achieve the smallest execution cost. Therefore, large blocks of shares are usually executed, i.e. bought or sold, by dividing them into smaller parts; then, the execution algorithm decides when to execute these smaller blocks. Splitting an order with a huge amount of shares into smaller amounts that are incrementally executed on the market allows traders to disguise their activity and participate in the market over a prolonged period of time, which typically ranges from a few minutes to several days (or even weeks). The order lifetime is dictated by the chosen aggressiveness (i.e. passive, aggressive or neutral trading), order type, size, limit price, start/end time etc.

AT mainly focuses on providing the best electronic execution to client orders with respect to a certain benchmark. The order execution does not stop at the point when the trade is placed; instead, it stops when an order is completely filled. The execution algorithms, which are called 'algos', are classified by the specific benchmark that is selected to measure the algorithm's performance. Due to the high volume of shares handled (buy/sell) every day, the AT clients, who are typically large institutional investors and fund managers, can measure the trading performance against certain industry-standard benchmarks, such as volume-weighted average price (VWAP) or indices like FTSE 100, DJIA and S&P 500. AT clients have alpha generation, including price prediction, as their primary responsibility, while AT providers attempt to achieve the lowest trading cost as measured against certain benchmarks (e.g. VWAP). In order to minimise the discrepancy between VWAP and the price obtained during execution, the provider needs to predict the volumes and execute the majority of the client order in times of larger volumes. Hence, volume prediction, or liquidity prediction, plays a more critical role for some algos compared to price prediction.

The factors affecting the algorithm/strategy choice include investor requirements (e.g. benchmark, risk aversion, and trading goals), order-specific properties (e.g. size), and asset-specific properties (e.g. liquidity, volatility, and price trends); each of these factors can affect the overall cost (Johnson, 2010).

Because of the recent high volume of alleged trader misbehaviour, the financial system's executives are accelerating their efforts to replace humans with computers and become technology leaders (Schäfer & Strauss, 2014). Electronic trading has witnessed continuously increasing volumes in the past years. Greenwich Associates' research shows that 74% of global FX trading is now executed electronically (Greenwich Associates, 2014). Besides the increasing misbehaviour headlines, the stricter regulations and the highly competitive financial services industry recommend algorithmic trading as a core business unit, with continuously more sophisticated technology. However, the electronic trading systems' algorithms need to be kept up to date by humans, which still leaves scope to game the system; therefore, the accelerated move to computer platforms does not completely diminish the risk of human wrongdoing (Schäfer & Strauss, 2014).

The increasing volume of electronic trading implicitly generates more data and statistics, which need to be analysed, decomposed and data mined in order to make sense of the huge amounts of generated raw data and provide valuable insights. The algorithmic trading systems' logs of parameters, along with the corresponding context and market structure, provide all of the necessary data for decomposing the eventual performance and determine its key drivers.

The principal motivation of this thesis is identifying the key drivers of trading volume, which ultimately affects the trading performance. This study analyses the underlying causal elements that affect trading performance when systematically executing orders in the stock markets. It tackles aspects such as: the size of the liquidity demand relative to the predicted or actual volumes traded in the market; the choice of execution strategy given the liquidity properties of the stock; and the timing of the trade, given the market conditions and released or anticipated company news. All of these provide insights into the investment skill of the portfolio manager, the execution skill of the trader, and the quality of the execution algorithm used.

Trading performance analysis separates the underlying market and trading data into various constituent causal factors, out of which only a small subset has a salient effect on the resulting trading performance. Analysing the trading performance is based on the results of trading history in order to perform transaction cost analysis (TCA); this measures the traders/brokers and the AT systems against various benchmarks, such as arrival price, VWAP or participation percentage benchmarks. Transaction costs can be classified as: explicit costs, which are easily identifiable and measurable (e.g. commissions, fees, and taxes); and implicit costs, which are normally associated with the trading process and are harder to identify and measure (e.g. market impact, price trending, timing risk, opportunity cost, spread cost, and delay cost). Trading performance analysis reconstructs the market conditions at the time of a particular trade, taking into account the market liquidity, news and context at that point in time and charting the fill prices and volume against the market. This analysis also compares traders with their peers at the same desk, firm or industry in the most relevant categories, including country, sector, market capitalisation, volatility, momentum, size or average daily volume (ADV).

Order profiling allows portfolio managers, brokers and traders to see what has worked well in the past, either for them or for others at their desk, firm or entire industry (Bloomberg Finance, 2010). The system helps answer questions such as “What is the best algo for this order given this time of day and this order size?”, “What is the opportunity cost of not participating there?”, “Whose fault is it when having a bad trading performance, is it the algo, the trader or the portfolio manager?”, “Is a broker/trader skilful or lucky?”, “Does a certain trader’s VWAP algo perform better than another’s?” etc. Performance analysis is important for post-trade comparison with regard to the results of brokers, traders and algorithms. The most widely used tool for performance analysis is benchmark comparison, although this is not perfect and there are better alternative metrics to easily compare performance across assets and time, such as Robert Kissell’s relative performance measure (RPM) (Johnson, 2010). Measuring market impact typically consists of quantifying the markets efficiency (e.g. transaction cost and price discovery) and fairness (e.g. insider trading and market manipulation). Despite the fact that it is mandatory to accurately measure the market impact in order to understand the execution style and improve the trading performance, there is limited research on modelling and decomposing trading performance. Market impact has been covered in many empirical studies and they unanimously found a concave function of volume (i.e. the derivative of market impact is a decreasing function), approximately increasing as the square root of order size

(Moro, et al., 2009), however, the functional form of market impact varies from study to study and represents an unexplained puzzle in the finance research. The market impact cost cannot be directly gauged because both the occurrence and non-occurrence of a transaction cannot be observed and measured. Therefore, practitioners need to estimate the price of an asset in the absence of its transaction and compare this approximate price with the actual transaction price (Torre & Ferrari, 1998).

The ability to accurately anticipate trading volume is crucial in order to model the market impact of an order, since the market impact cost of a trade is a function of volume. The practical importance of market impact originates from the fact that it can reduce profits and it can turn a profitable strategy into a losing strategy by adversely moving the prices. Practitioners also take into account market impact because it increases with the trading size and therefore it limits the size of funds. Therefore, market impact plays an important role in determining the size distribution of funds as a fund could become too large (Moro, et al., 2009).

Furthermore, planning a multi-day trade is important for practitioners and we propose a couple of multi-step ahead analyses. These address the practical problem where traders and portfolio managers aim to size the allocation of their multi-day trades based on the available liquidity in order to minimise the market impact. This motivated us to model the investigated effects (e.g. the cross-market holiday effect or the expiry day effect) and evaluate how well these analyses perform the further we go into the future (i.e. by increasing the step size of the multi-step ahead forecasts).

The goal of evaluating any aspect of a trading plan is to minimise risk, create consistency and generate greater profitability. Identifying the factors driving the profitability and determining how good a trading plan is can bring plenty of benefits to the financial system's business management. The actor of this thesis' performance analysis can be an algo or a human trader, i.e. a person being engaged in the transfer of financial assets either for themselves or on behalf of a client.

Having posed the main problem this research is looking at, further work can be done to extend the proposed volume-based (trading) performance analysis framework within the algorithmic trading application domain by taking into account other drivers of performance.

The ongoing disputes between the efficient market hypothesis (EMH) supporters (e.g. Eugene Fama) and contesters (e.g. Richard Thaler and Daniel Kahneman) with regard to behavioural concepts (e.g. calendar, holiday and intraday effects) calls for further investigation in order to validate these theories.

Moreover, the availability and recent development of powerful computational statistics and machine learning techniques, along with the computationally powerful programming environments, allow this study to bring together concepts from different areas and apply them to conduct performance analysis in an application domain that has not been investigated enough academically.

1.2 Research Objectives

The main objective of this research is to investigate the key factors that adversely affect the trading volume, and hence the performance of algorithms, causing them to produce excessive market impact or to under-participate when the market runs away. To achieve this, a series of experiments has been carried out in order to analyse the trading volumes and decompose the data by using current variable selection techniques in a controlled environment.

In order to explore the key drivers of behaviour in financial systems, an in-sample analysis of trading volumes is carried out to understand the factors driving volumes and illustrate what the variables are when the trading system is a tool (i.e. extraneous variables). The study compares various statistical methods for variable selection/complexity reduction to address the applied problem of this thesis.

The trading performance decomposition aims to also explain the trader's skills and the portfolio manager's skills within the sequential order execution process, as illustrated in Figure 1.1.

Figure 1.1. The chain of order placing.

The study further explores the market liquidity and its causal elements by analysing the temporal effects and their influence on trading volume, also looking to validate popular theories in the behavioural finance research literature. The experiment aims to determine the appropriate level of participation in the market and aims to build a volume prediction model based on the factors identified in the in-sample analysis.

1.3 Research Experiments

The thesis starts from a broad exploration of trading volume drivers in the first experiment, drills further down in the subsequent experiments, and concludes with an out-of-sample analysis based on the empirical evidence in the previous in-sample studies. It is split into four empirical studies:

1. Examining Drivers of Trading Volume. The first experiment aims to better understand the market dynamics and investigates the trading volume causal factors. The study considers various disjoint theories from the literature (e.g. the correlation between trading volume and price changes, the asymmetry of this relation, calendar effects etc.) and proposes a regression model which combines a broad feature set, ranging from autoregressive trading volumes and day-of-the-week indicator variables, to asymmetric price metrics (i.e. overnight and intraday returns). This is in contrast to the current research literature, which does not take into account the broad context when looking at a particular effect (e.g. the day-

of-the-week effect, where most of the models analyse the effect on its own and are essentially trained using an intercept and the day-of-the-week feature only). Unlike this approach, we analyse these effects using the variables' conditional expectation with respect to the other volume determinants. The analysis further explores and validates behavioural finance theories with regard to trading volume, namely the day-of-the-week effect. This in-sample time series analysis is important to determine the optimal order sizes and applies machine learning techniques such as feature selection with cross-validation and linear regression. The next two studies focus on special events, which are potentially influencing the trading volume.

2. European Trading Volumes on Cross-Market Holidays. This study investigates the effect of bank holidays on the trading volume in venues across European equity markets. We constructed an accurate trading calendar for the European countries and the US and, to the best of our knowledge, we conduct the first pan-European study investigating the cross-market holiday effect. Moreover, very little attention has been paid to the volume dimension when analysing calendar effects and the authors have traditionally focused on returns only. Our study aims to complement the existing literature by providing insights into the relation between the effects of these rare notable events and the trading volumes. This study investigates whether the higher volumes on cross-market holidays are significantly different from the Monday effect and explores a differentiated effect depending on market capitalisation.
3. Expiry Day Effects on European Trading Volumes. This analysis explores the trading volume seasonal variations associated with periodic events, such as the stock index futures expiries and MSCI quarterly reviews. The trading calendar includes pan-European futures and MSCI rebalances. We investigate the volume patterns around the index expiry and review dates, examining potential anticipatory and subsequent effects. We describe the phenomena of lower trading volumes on futures expiries in connection with the Friday effect and lower volumes on MSCI quarterly reviews in conjunction with the end-of-month effect in order to determine the actual driver of higher trading volumes.
4. Developing a Volume Forecasting Model. This study integrates the findings of the previous in-sample experiments for data exploration, and proposes an out-of-sample analysis of trading volume based on a dynamic model, which switches depending on the current context and market conditions. It forecasts the expected volume to mitigate market impact, while aiming to improve the error stability. This study trains a number of models using seven statistical methods, each being trained using two approaches (i.e. rolling window and growing window), and switches from one model definition to another in order to predict the trading volume given the contextual training data (e.g. market holidays, futures expiries, MSCI reviews, day-of-the-week etc.). The primary focus of this study is to achieve optimal accuracy by fine-tuning the models. We conclude by proposing a stock-specific out-of-sample metamodel, which chooses the prediction model based on the performance ranking of the initial stock-specific models within a given time window.

1.4 Scientific Contributions

This research contributes to the existing literature in a number of ways:

1. Detailed exploration of trading volume drivers. This thesis explores the causal factors of trading volume and, unlike the traditional literature that focuses on prices and returns, this study contributes to the existing finance literature by providing insights into the volume dimension of market data. Primary drivers of trading volume include the lagged time series, the asymmetric price-volume relation, and the temporal dimension, ranging from the day-of-the-week effect to holidays and other notable dates.
2. Focus on the volume dimension. A notable contribution is that the empirical studies in this thesis concentrate on the volume dimension unlike the traditional literature, which focuses on the relation between exogenous variables (e.g. calendar effects) and price returns only.
3. Comprehensive pan-European stock universe. To the best of our knowledge, we are the first authors to employ a huge pan-European data set for trading volume prediction. The data universe includes daily market data for 2,353 stocks trading in 21 European countries, since 1st January 2000, covering a timeframe of over 15 years of daily market data. Very few studies have analysed the European markets; most of the literature either investigates the US markets or it includes very small samples of European stocks, which are not representative of the overall European equity market.
4. Accurate trading calendar data. We constructed a thorough and robust calendar data set to be used in conjunction with the market data. The motivation to conduct this ample data retrieval and normalisation process was dictated by the unavailability of a high-precision trading calendar. The trading calendar includes a comprehensive breadth and depth of events that affect the pan-European equities, comprising the trading calendar for the most liquid exchanges in 21 European countries along with the US trading calendar, the European stock index futures expiries, and the MSCI quarterly review effective dates, along with historical evidence of index additions and eliminations, since the beginning of 2000.
5. Established statistical methods applied to a new application domain. The studies in this thesis are using advanced machine learning and statistical analysis methods in order to develop a framework for understanding the trading volume, which in turn affects the trading performance. These methods have been existing for a while, but have not been applied to this application domain before.
6. Further insights into calendar effects. The study investigates the liquidity demand and market impact, and sheds light on causal factors of trading volume, including some inconclusive calendar effects. This research corroborates previously disputed studies in the behavioural finance literature and provides further insights into the day-of-the-week effect, the cross-market holiday effect and the price-volume relation.
7. Research validation. The research in this thesis was carried out in collaboration with Deutsche Bank. The supervision and advice of a top-tier investment bank helped drive and validate this research, while addressing real-world problems that practitioners are currently facing, e.g. multi-day trade planning.

1.5 Thesis Structure

The structure of this thesis is organised as follows:

- Chapter 2 – Background and Literature Review. The relevant literature and the key concepts in the areas of this research are reviewed, in order to introduce the reader to the problems and environment of this thesis. The chapter starts by giving the reader a clear view of the financial context and concepts. It particularly presents the broad area of financial markets and algorithmic trading in order to place this research in a well-defined context and to introduce this thesis' industry-oriented application domain, along with a review of trading performance, performance benchmarks, market structure and other related financial research and knowledge. The chapter ends with a methodological review of machine learning and statistical analysis methods, surveying the data analysis problems in this thesis and the solutions that recent approaches offer.
- Chapter 3 – Examining Drivers of Trading Volume. This chapter provides an in-sample time series analysis of volume and looks into the application of variable selection techniques in order to empirically find the key drivers of trading volumes, ranging from lagged volumes, price measures, and price-volume asymmetry, to the days of the week. This data exploration study is conducted to understand the important factors influencing the trading volume by destroying the temporal dimension (i.e. treating the data as being stationary and time-agnostic).
- Chapter 4 – European Trading Volumes on Cross-Market Holidays. This second study explores the temporal frequent factors (i.e. cross-market holidays) affecting trading volumes. The analysis provides statistical significance tests to distinguish the cross-market holiday effect from the Monday effect. Multi-step ahead forecasting and potentially different effect magnitudes based on market capitalisation are also examined.
- Chapter 5 – Expiry Day Effects on European Trading Volumes. The third study extends the previous chapter's analysis and continues the exploration of temporal factors affecting trading volumes by examining sparse and periodic events, namely the futures expiries for seven European stock indices and the MSCI rebalances. The higher volumes on futures expiries and MSCI rebalances are contrasted with the Friday effect and the end-of-month effect, respectively.
- Chapter 6 – Developing a Volume Forecasting Model. This study provides an out-of-sample extension of the previous in-sample volume experiments, and proposes a machine learning framework for predicting the trading volume based on a dynamic model. The volume prediction model focuses on accurate prediction by improving the stability of error changes. It is trained either on a growing window basis, where more observations are available, or on a sliding window basis, where only the more recent data is used to train the model, using seven different statistical methods. The prediction model switches from one model definition to another depending on the temporal market context and ultimately proposes a stock-specific out-of-sample metamodel, which assesses the performance of the initial stock-

specific models across the most recent time series in order to choose the optimal prediction model.

- Chapter 7 – Conclusion and Future Work. The final chapter provides an overall conclusion of this research with a summary of the key findings of this work, and what can be learnt from the results of its models and experiments. The thesis ends with our recommendations for future work to be done in this area.

This page intentionally left blank.

2. Background and Literature Review

This chapter presents background information on a number of key concepts in the areas that this research focuses on. In the first part of this chapter, which is concerned with the application domain, particularly with trading volume and its impact on trading performance, the broad areas of algorithmic trading and financial markets are reviewed to place this work in context. The discussion centres on the investment and execution processes, which are followed by an illustration of market structure, as it is viewed from two conflicting perspectives. The chapter ends with a survey of machine learning and statistical analysis techniques for modelling trading volumes and trading performance.

The thesis consists of four experiments: the first three studies investigate the trading volume in-sample, while the last one proposes an out-of-sample volume prediction model. Every experiment gradually provides an in-depth analysis that is based on the previous findings of this thesis, so that the storyline flows logically. Due to the shared knowledge base required by the experiments of this thesis, the background literature review chapter is structured by research topic in order to avoid repetition, instead of following a structure similar to the thesis experiments.

Moreover, the background literature section within each experiment builds on the top of the literature surveyed in this broad background literature review. This section is a high-level presentation of the relevant literature, while each chapter particularly focuses on some of the aspects presented in the Background and Literature Review chapter and therefore the literature that is directly relevant to each experiment is reviewed in the subsequent studies.

The first part of the background chapter exhibits the narrative thread of this research by centring our research in the well-established and predominantly disjoint literature. The application domain section starts by introducing the broad context of this research (i.e. the investment process and decisions, and the execution/implementation process), where the role of volume in correctly sizing an order for minimal market impact is emphasised. The storyline is built in the next order:

- The execution process and its aim to minimise the impact cost by measuring the slippage to arrival. The impact is minimised by optimally sizing the order compared to what is usually trading. If the size is too small, then the trader is underperforming and incurs opportunity cost by missing the opportunity. Conversely, if the trader overparticipates, there is a risk of adversely pushing the price. In any case, volume prediction is crucial to correctly size the orders.

- Algorithmic trading (and trading in general) is believed to be business as usual, assuming that markets are efficient. However, according to the behavioural finance concepts, the market switches from one state to another (e.g. there are days when market dynamics are different), causing disturbance to equilibrium. The behavioural finance field is briefly introduced and a synopsis of the calendar effects is then exhibited to outline the vast literature investigating the various calendar effects, whose impact is predominantly analysed on price returns. There is not much emphasis on the volume dimension.
- Due to the lack of analysis of calendar effects explicitly impacting the trading volume, the next section reviews the relation between price and volume, where there are many theories on how to represent the price change (i.e. price change per se vs. absolute price change) and the asymmetry of the volume-price change relation. Eventually, the calendar effects on price returns and the relation between volume and price allow us to infer a direct link between calendar effects and trading volume.
- Once the trading volume is heavily documented, it can be used in the context of trading performance. In this section, we outline the liquidity and market impact (in relation to volume), and review the hardly accessible literature on trading performance and performance benchmarking.

The second part of this chapter outlines the fundamental methodology and techniques that act as a basis for the more advanced analyses performed in this research. It presents an overview of supervised learning, followed by an explanation of the analysis types (e.g. in-sample and out-of-sample analysis types; time series and cross-sectional data etc.) that are going to be used in this thesis, along with a methodological description of the randomisation tests.

2.1 Application Domain

2.1.1 Importance of Trading Volume

One of the main concerns of this thesis is the trading volume, i.e. the number of shares changing hands during a given period of time; a common time frame is per day, which is denoted as daily trading volume. The trading volume measures the level of trader participation (or market activity) and is a strong indicator of liquidity (i.e. the extent to which a stock's shares can be bought and sold); low volume indicates that the given stock is bought or sold infrequently, making it difficult to buy or sell the stock relatively quickly, while high trading volume means that the stock is highly liquid and can be easily traded. The trading volume affects trading strategies, e.g. an investor who wants to sell a large number of shares for a stock with low volume has to break down the sell order over a longer period in order to trade slowly with minimum market impact. The market impact is closely related to market liquidity and represents the effect of a participant, when buying or selling shares, i.e. the extent to which the price goes upward when buying or goes downward when selling. If both the price and the volume for a stock are rising, then it means there is higher demand in the market for the stock, as its direction is correlated with participation. The trading volume contains no information about the number of participants; instead, it is simply the number of shares traded and, consequently, a high volume change can be caused by a few participants.

The price movement can be an indicator of trading activity over a period. However, the underlying trading volume is a better indicator for demand trends and investor interest. Moreover, the volume is used to interpret the relevance of a market move, i.e. the higher the volume during a price move, the more significant the move is.

An important indicator related to volume is the turnover (or turnover ratio), which is the ratio of a stock's trading volume and its float (i.e. the number of shares available to trade publicly). Consequently, a stock with a high daily volume and a high number of shares outstanding has a high turnover, whereas a stock with a low daily volume and a high number of shares outstanding has a low turnover. While the trading volume is a liquidity indicator, the turnover ratio is also considered to indicate a stock's stability and volatility; high turnover means that there are few shares outstanding and therefore a sudden change in demand can result in a significant impact on the stock value, whereas a stock with low turnover has a large number of available shares and its price would not be affected considerably by a suddenly high demand.

2.1.2 Volume-Price Relation

The relation between trading volume and prices is important in order to better understand the financial market structure. Price changes indicate the market response to new information, while the trading volume measures the level of disagreement of the information among investors (Beaver, 1968).

Previous literature broadly presents two views on the relation between volume and price dynamics. Despite acknowledging the fact that traders have different interpretations of the public signals (i.e. priors and likelihoods), causing disagreement, which then results in trading activity, the first view considers that the trades caused by disagreement are idiosyncratic and cancel each other having no effect on stock prices; therefore, trading volume could be isolated from stock price analyses (Hong & Stein, 2007). In view of this theory, some papers discuss volume prediction models while disregarding any connection with prices; for instance, Kandel and Pearson (1995) argue that the empirical evidence on the relation between trading volume and stock returns around public announcements is inconsistent. The opposite viewpoint claims that trading volume is associated with investor sentiment; volume increases considerably when prices tend to be too high compared to the fundamentals, and can possibly lead to speculative bubbles and even broader cross-sectional samples (Hong & Stein, 2007). Harris and Raviv (1993) argue that price changes and volume are positively correlated and volume is *possibly* autocorrelated. Dichotomous evidence is found on the directionality of this relation: Assogbavi and Osagie (2006) find that stock price changes lead the trading volume in the emerging markets, while Morgan (1976), Harris (1987), and Kemal and Starks (1998) argue the opposite, i.e. that there is a positive association between volume and returns, but it is the volume leading returns, noting that the differences in institutions and information flows are strong enough to affect the stock valuation process.

The literature investigating the price-volume relation has found that volume is positively correlated with two price indicators: the magnitude (or absolute value) of the price change, i.e. $|\Delta p|$, and the 'price change per se', i.e. Δp, (Karpoff, 1987), where price change can refer to the log-price difference or to the price percentage change. These two correlations refer to two popular Wall Street adages: 1) "volume is relatively heavy in bull markets and light in bear markets", i.e. volume is correlated with the price change per se; 2) "it takes volume to make prices move", i.e. volume is correlated with the absolute value of the price change (Assogbavi & Osagie, 2006). Karpoff indicates two potential causes of simultaneous high volumes and large price changes: the sequential information arrival model and the mixture of distributions hypothesis; the volume being higher for a price increase than for a corresponding price decrease can be explained by the higher costs associated with short positions as opposed to long positions.

2.1.2.1 Volume and Absolute Price Change

The positive correlation between trading volume and absolute price changes was early investigated by Clark (1973), who found a curvilinear relation between volume and the square of the price, Westerfield (1977), Tauchen and Pitts (1983), who argue that one can determine the parameters of the joint distribution of volume and price change by maximum likelihood and then identify the conditional probability of the squared price of a stock given its volume, and Rutledge (1979), who found a volume-price correlation in the commodity futures markets. Epps and Epps (1976) support Clark's hypothesis and bring empirical evidence for the existence of a dependence relation between volume and the change in log-prices, using the $\Delta \log p_i$ as an approximation for $\Delta \frac{p_i}{p_{i-1}}$. Ying's empirical results (1966) on the S&P 500 composite index daily price data and NYSE daily volume data found significant correlation between small volume and price decrease, large volume and price increase (both results implying a relation between volume and Δp), and large volume increase and large price increase/decrease (implying a correlation between volume and $|\Delta p|$). Ying's study was conducted very early and has been criticised mainly because of the incomparable price and volume data sets. The results of Crouch (1970) support Ying's latter hypothesis, finding a relation between a stock's volume and the absolute value of its price change. An explanation for the relation between volume and absolute price change roots from Fama's study (1965) on the mixture of distributions hypothesis (MDH), which suggests that daily price changes are sampled from distributions with different variances and consequently daily price changes have a leptokurtic distribution (i.e. with excess kurtosis) compared to the normal distribution. Investigating the commodity futures contract in China, Chen et al. (2005) found that there is no correlation between trading volume and the price change per se; however, they found a significant correlation between volume and the absolute price change and performed a Granger causality test to examine the lead-lag relation between trading volume and the absolute return, finding that the absolute price change is leading volume.

2.1.2.2 Volume and Raw Price Change

Epps (1977) examined the relation between the trading volume and the raw price change, or 'price change per se', as coined in the background literature (i.e. the raw value and not the absolute value of the price change), for corporate bonds and found that the volume/price change ratio is larger for positive (i.e. upticks) than for negative price changes (i.e. downticks); Epps' empirical findings were supported by Hanna (1978), who replicated Epps' study. Rogalski's empirical work (1978) provides evidence for a significant cross-correlation between volume and the price change per se at lag zero, and explores the relation between the current price change and lagged volume, finding no dependence. Furthermore, Rogalski suggested that using an appropriate autoregressive-moving-average (ARMA) model to predict the trading volume from past volume observations would be feasible, but such a model would not provide any relevant information on the price change; only predictions of the variance of price change could be inferred from an ARMA volume model. A popular heuristic in technical analysis suggests that price is preceded by volume, e.g. if volume starts decreasing in an uptrend, then this might be evidence that the upward trend is ending soon. A bidirectional non-linear Granger causality relation between returns and volume was found in the developed markets (Hiemstra & Jones, 1994) and in the emerging markets (Assogbavi, et al., 2007).

2.1.2.3 The Uncorrelatedness of Volume and Price Change

The hypothesis that there is a positive correlation between volume and price change was rejected by some empirical studies, which found little correlation between the two series using weekly data, possibly because prices are believed to follow a random walk model in the short run (Morgenstern & Granger, 1963), although another study has demonstrated that this hypothesis is incorrect (Crouch, 1970). Morgenstern and Granger's study was further extended on daily and individual transaction data and, again, no relation between volume and prices (or the absolute value of the price difference) was found. However, this time, it revealed a modest correlation between volume and the difference between the daily high and low (Godfrey, et al., 1964), i.e. the intraday range. The mixture of findings and the divergent conclusions of this research area call for further work to provide insights on the relation between volume and price change.

2.1.2.4 Drivers of Trading Volume and Price Covariance

Hong and Stein (2007) investigated the covariance between price changes and trading volume changes, and suggested that the two series exhibit a highly significant correlation. They argue that the underlying factors, either at the level of market structure or individual cognition, that create trader disagreement on prices (and hence trading volumes) are gradual information flow, limited attention, and heterogeneous priors.

Gradual information flow assumes that some investors receive certain pieces of information before others, probably due to transmission channels or investor specialisation. The investors who received the information normally revise their valuations of the given stock, reflecting the positive (or negative) impact by an upward (or a downward) revision, while the other investors who have not seen the particular piece of information keep their valuation constant. Consequently, the disagreement between the two groups of investors increases. A representative example is the case of EntreMed. The New York Times featured a front-page article reporting on cancer breakthrough and EntreMed, the company holding the licensing rights for the product (Huberman & Regev, 2001). However, the same news story was reported more than 5 months earlier in Nature and New York Times itself, although less conspicuously. The investors' enthusiasm spilled over to other biotech stocks, the NASDAQ Biotech Index (excluding EntreMed and comprising 123 companies) increased by 7.5% and seven firms in the index experienced a return greater than 25% and a trading volume 50 times larger than the average. This story clearly differentiates the two types of investors. The first group reacts on the initial news story that came out in Nature and updates EntreMed's valuation, while the second and larger group of investors gets informed from the front page of New York Times. Therefore, the information flows such that specialists get informed before generalists, causing elevated trading volume around the news release time, followed by a gradual response of prices to the news story. Cohen and Frazzini (2008), and Menzly and Ozbas (2010) find that prices do not incorporate public news about economically linked firms.

The idea of limited attention is similar to that of gradual information flow and suggests that some cognitively overloaded investors pay attention to a small subset of public information, meaning that the price and trading volume will respond more prominently to a high-profile news release, as documented by Hirshleifer and Teoh (2003), and Klibanoff et al. (1998). A well-observed temporal pattern is the investor inattention to Friday earnings announcements (DellaVigna & Pollet, 2009), when there is lower immediate response, higher delayed response and 8% lower trading volume.

An important mechanism that can generate investor disagreement consists of heterogeneous priors, as investors have different economic models (Kandel & Pearson, 1995) that they use to interpret new information, hence increasing their disagreement on the fundamental value of stocks. Hong and Stein (2007) provide evidence of heterogeneous priors by analysing the trading volume around public earnings announcements; they report an elevated turnover around the announcement release time, which then remains constant for a week. This observation contradicts the rational-expectations model with common priors, where the investor disagreement should be decreased, rather than increased, by public earnings announcements.

2.1.3 Financial Markets

Financial markets broadly refer to any marketplace where buyers and sellers trade assets (e.g. equities/stocks, bonds, commodities, currencies or derivatives). This thesis concentrates on the stock market, but there is a variety of market types: capital markets (stock markets and bond markets), commodity markets, money markets, derivatives markets, foreign exchange markets, and commodity markets. These markets are typically characterised by pricing transparency, basic trading regulation, costs and fees. The main roles of financial markets include capital raising (capital markets), risk transferring (derivatives markets), liquidity transferring (money markets), international trading (currency markets), and price discovery (i.e. the process of determining an asset's price by the buyers and sellers interacting in the marketplace). However, the most important activity is the intermediation for raising capital. Financial markets match those looking to raise capital (i.e. borrowers, such as individuals, companies, corporations, or governments) to those who have and lend it (i.e. lenders, such as companies and pools of individuals) in return to some compensation (e.g. interest or dividends). The market dynamics consist of price signals caused by the change in the market's demand and supply.

2.1.3.1 Market Structure

In order to investigate the relation between market dynamics and some exogenous variables, we need to first review some important aspects of the market structure:

- Changes in the regulatory landscape, e.g. the attempts to impose a minimum tick size or the implementation of market-wide circuit breakers, which, unlike the stock-specific Limit Up-Limit Down mechanism that halts an individual stock trading, halts the market-wide trading if the S&P 500 index falls by 7%, 13%, or 20% in a trading day.
- Electronic trading strategies. Nowadays high frequency trading (HFT) is a dominant component of the market structure and calls for careful consideration of technology and infrastructure, with emphasis on execution speed.
- Fragmentation. The higher competition provides investors lower fees, tight bid/ask spreads, and innovative trading platforms and services. However, this can lead to trading fragmentation when many trading venues compete against each other to obtain continuous order flow for a stock (U.S. Securities and Exchange Commission, 2013).
- Dark pools and off-exchange trading. Dark pools are Alternative Trading Systems (ATS) that trade in venues other than the traditional exchanges by hiding the bids and offers, and therefore providing anonymity. Consequently, they decrease the transaction costs by reducing the information leakage and the signalling risks for investors. Dark pools are an appropriate execution tool for large orders and stocks with wide spreads or low liquidity (BlackRock, 2014).

2.1.3.2 Algorithmic Trading

The trading performance attribution mainly examines algorithmic trading machines. This section comprises an outline of the algorithmic trading business, its increasing usage, and the steps of the AT process. This section concludes with the importance of liquidity prediction (and hence volume prediction) in the context of algorithmic trading. As of 2012, approximately 78% of the trading volume of the US equity market was performed via algorithmic trading (NeverLossTrading, 2014), which involves sophisticated algorithms (partially or fully) automating the trading process, i.e. algorithms performing trading on behalf of people on the execution side. HFT is specifically characterised by extremely short position-holding periods (Treleaven, et al., 2013). The chase for very low trading speeds has driven challenging and innovative research on high-performance computing, such as graphics processing unit (GPU) computing and field-programmable gate array (FPGA) programming (Che, et al., 2008). Algorithmic trading attempts to spot market anomalies and capitalise on statistical patterns, optimal order execution, or rivals' strategies exploitation, with profits coming from client commissions, cost savings, or proprietary trading (Nutti, et al., 2011).

A shared centralised order book allows brokers to execute orders by listing the buy and sell orders sorted by price (in descending order for buy position and ascending order for sell position) and quantity for a given security. The centralised order book attempts to match the orders on the top of the book (i.e. highest buy price with lowest sell price); exchanges usually prioritise orders by price and arrival time (on a first-in-first-out basis).

Algorithmic Trading Cycle

Although not directly relevant to the thesis, it is crucial to understand the trading cycle and understand where a trading volume analysis can improve the trading performance. The algorithmic trading process (Treleaven, et al., 2013) typically consists of the following cycle:

- Data access and data cleansing. Real-time and historical financial, economic, social and news data driving the algorithmic trading decision-making is acquired and cleansed (i.e. correcting/removing erroneous data points).
- Pre-trade analysis. Financial data and news are analysed to evaluate the properties of the assets and spot potential trading opportunities. This analysis is performed with various techniques, e.g. fundamental analysis (investigating the macro-economic and company-specific variables affecting an asset's valuation, using ratios, such as price-to-earnings ratio, or fundamental properties, such as earnings yield), technical analysis (identifying and understanding historical changes in prices and volumes), or quantitative analysis (exploiting and predicting financial/economic/news data patterns using computational concepts ranging from statistics to physics and machine learning, while focusing on an asset price's stochastic nature). The pre-trade analysis can be further divided into three main computational components (Narang, 2009):
    - Alpha model. This component predicts the future behaviour of the selected asset.
    - Risk model. The level of exposure and risk associated with the given asset and portfolio is assessed.
    - Transaction cost model. The expected costs of trading the asset are calculated, including commissions, slippage and market impact.
- Trading signal generation. This stage gives buy and sell recommendations (i.e. what and when to trade) and consists of the portfolio construction model, which decides the optimal portfolio of assets (i.e. the selected securities and their quantities) that should be owned in the next time horizon in order to maximise profits, minimise trading costs, and limit risks.
- Trade execution. At this stage, the orders for the recommended portfolio's assets are executed. The trade execution can be performed either as agency/broker execution (on behalf of a client), or as principal/proprietary trading (on the institution's own account).
- Post-trade analysis. The results of the trading activity are evaluated to measure the trade performance, e.g. slippage (i.e. the difference between a trade's expected price and the actual price it was executed at), profit & loss statement (P&L), algorithmic performance measurement, cost measurement etc.

Liquidity Prediction

This research analyses and predicts the trading volume; it is important to understand the crucial role volume is playing in the algorithmic trading environment. Although the main goal is to simply get the best price, volume prediction, i.e. liquidity prediction, plays a more important role in algorithmic trading than price prediction. The clients of AT businesses are fund managers whose main responsibility consists of price prediction. However, they require help in order to minimise their trading costs, which explains the availability of a wide range of execution algorithms. The trading costs can be classified as explicit (e.g. commissions and fees) or implicit (e.g. spreads, market impact, price trends, timing risk, and opportunity cost) (Johnson, 2010).

Minimising trading costs leads to finding liquidity. Despite the simplicity of its definition, i.e. the extent to which a security can be traded (bought or sold) in the market at a given price, liquidity is a complex factor in the non-stationary environment of financial time series and it is a multidimensional quantity that is only partially observed. Therefore, it is a challenging task for clients to choose from the various algorithmic trading providers and evaluate the brokers' liquidity-seeking skills in order to guarantee the lowest trading costs for the clients' trading activity.

Clients of algorithmic trading providers typically use benchmarks to measure the trading performance since accessible liquidity cannot be measured or predicted reliably. The benchmarks are calculated after the order execution and good trading performance is considered to have the lowest deviation from the best price that is potentially available during the order lifetime. The best price is a placeholder for any benchmark; one of the most popular benchmarks is volume-weighted average price (VWAP), which motivates the broker to execute the trade at a price where most of the shares change hands during the order lifetime. Hence, the price is less relevant in this situation as long as the order is executed simultaneously with everybody else at the current market price. The main goal of algos is to be able to predict the total execution volume for the next time period. Provided that the algo predicts a higher trading volume for the next time bucket, then one should increase their trading rate or, otherwise, decrease it. Even if modern algos are more complex and sophisticated than this rudimentary example, volume prediction still plays a crucial role in algorithmic trading and a slightly better volume prediction rate can result in a significant decrease of the algo slippage against its benchmark. The implementation slippage (illustrated in Figure 2.1), or implementation shortfall, is the difference between the actual portfolio performance and its theoretical performance. The implementation slippage is a reflection of transaction costs and, generally, the larger the order, the larger the slippage.

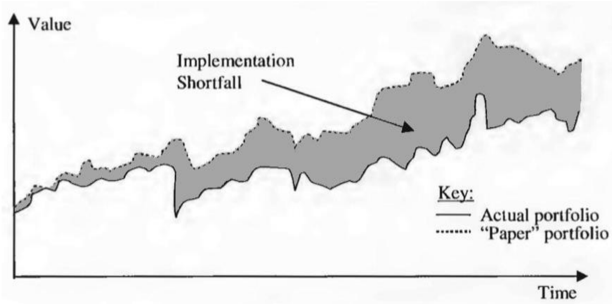

Figure 2.1. Implementation shortfall (Johnson, 2010).

The brokers and algorithmic trading providers are very restricted because orders are placed with a particular side, i.e. buy/sell, so that the algo cannot sell a client's shares when the original order was buy and vice versa; also, the orders can have a limit price, start/end time, minimum/maximum participation level, minimum/maximum trade size, ability/requirement to participate in auctions and routing instructions, among many other constraints. However, they can choose to trade more or less aggressively. Scenarios where additional functionalities are enabled to hedge an order, i.e. selling short another asset in order to temporarily offset the market exposure, are not considered in this context.

2.1.3.3 Investment Process

This research investigates the trading volume in order to derive a model to assess market impact (and hence the trading performance). Figure 2.2 exhibits a simplified view of the investment process and covers both the traditional broker trading and the algorithmic execution. This process is of great importance to this thesis since the market impact modelling applies both to traders executing orders and investment managers allocating portfolios, and it is vital to understand the transaction costs (and therefore the market impact costs) spanning the entire investment process, from the initial decision of a buy/sell order to the order creation and execution. The transaction costs are inevitable, but they can be minimised by modelling the volume distribution in order to lower the market impact.

The investment process starts on the buy-side with a potential buy or sell idea, whose impact on the investment portfolio is assessed by the portfolio manager, who determines the target positions using optimisation and risk analysis techniques. Then, an internal trader identifies the optimal way of trading these orders, by estimating the potential transaction costs and considering the historical broker performance. This is followed by the orders being routed for execution and they are usually sent to a sell-side institution, which ultimately sends them out to the market. This final process can be performed either by a salesperson (or trader) or by an algorithmic trading, or direct market access (DMA), process. Eventually, the orders are executed and the buy-side receives the trade reports (Johnson, 2010).

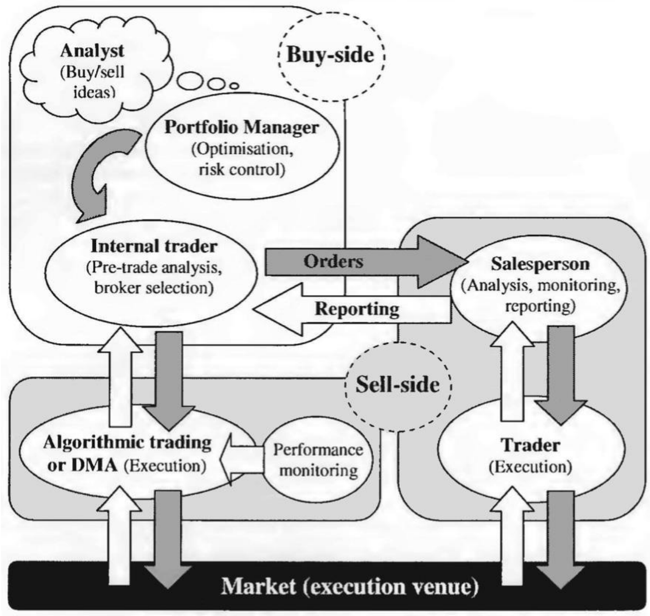

Figure 2.2. The investment process (Johnson, 2010).

We differentiate between two approaches of investment process and decisions: fundamental investing and systematic/quantitative investing; the thesis will focus on the latter.

Fundamental Investment Strategies

Fundamental investing strategies, or macro-strategies, apply theories in order to invest on the long-term (i.e. passive investment).

'Buy-and-hold' is a passive investment strategy where the investor does a financial accounting evaluation of various stocks, and buys and holds them for a long time horizon no matter what the short-term price movements are. There is a rule of thumb implying that investing in the long run increases the rate of return. The opposite of buy-and-hold is day trading, where investors buy low and sell high, with potentially higher profits obtained on greater volatility.

Other fundamental investing strategies include value investing (i.e. buying a quality stock at a discounted price, while knowing its intrinsic value), qualitative measures (i.e. evaluating stocks depending on management, corporate governance, company, industry, competition etc.), growth investing (i.e. investing in companies with future growth prospects with insignificant emphasis on the current price value), and income investing (i.e. focusing on companies that provide a steady stream of income).

Systematic Investment Strategies

Systematic (or quantitative) investment strategies are looking at the trends in the market and the patterns in prices, exploiting profitable conditions and seeking short-term profits (buy low, sell high). This investing technique typically requires sophisticated systems and technically skilled quants. Systematic investing examines large amounts of data in order to find repeating patterns of a phenomenon, correlations between two liquid assets (e.g. pairs trading), or price-movement patterns (e.g. trend following or mean reversion). Alpha is a risk-adjusted performance measure representing the return in excess of what would be predicted by an equilibrium model. Quant strategies are also known as alpha generators or alpha gens. Alpha strategies focus on generating alpha across a range of product segments.

The goal of quant investing is like any other investment strategy, i.e. adding value, alpha or excess returns. Quants (i.e. the quant strategy developers) create complex mathematical models in order to spot investment opportunities. In systematic investing, the model (and ultimately the computer) makes the buy and sell decisions, removing any human biases that are normally associated with a person that buys or sells investments.

Some examples of systematic investing strategies include: momentum investing (capitalising on the continuance of existing market trends), pair trading (matching a long position with a short position in two historically correlated stocks, i.e. the pair, belonging to the same sector), and mean reversion (assuming that a stock's price will have the tendency of moving to the average price over time).

2.1.3.4 Execution Process

Trade execution is a key concept for this research and our empirical work addresses the issue of measuring the market impact of a trading order. The trade execution/implementation process makes decisions on the trading strategy (e.g. a popular execution strategy is volume-weighted average price), the trading venue, order type (e.g. market orders, which are executed at the current market prices and provide immediate execution if the market is liquid enough, or limit orders, where a given security is bought at no more, or sold at no less, than a specified price limit, guaranteeing the execution price, although the order might not be fully executed), order size, degree of trader's anonymity etc., while being constrained on the transaction costs and trading duration, and aiming to minimise market impact and timing risk. These decisions, i.e. how to trade: where, what and when, are handled by order routing.

The objectives of algorithmic trading vary depending on the type of trading: broker algorithmic trading systems aim to optimise the execution strategy (e.g. minimise market impact cost or execution time) in order to minimise the cost of trading, while proprietary algorithmic trading systems aim to maximise risk-adjusted profits.

Many financial instruments can be traded in more than one market (i.e. multiple trading venues). Trading systems typically choose trading venues depending on liquidity, trading mechanism, degree of trader anonymity, different execution costs etc. However, the most important characteristic in choosing a trading venue is liquidity, because highly liquid markets imply immediate order execution and low transaction costs.

Trading and Market Impact

Market impact is the central focus of this thesis. It is defined as the price change caused by a specific trade or order, usually having an adverse effect and driving the price against the transaction itself: a buy order would increase the price, while a sell order would decrease the price. The rationale behind this stems from standard economic theory; an increase in demand should drive the prices up, and an increase in supply should drive them down. The market impact cost is a price-per-share amount and is calculated as the difference between the actual price and the hypothetical price if the order was not created (Johnson, 2010). The goal of our volume analysis experiments is to minimise market impact and it is crucial to further explain it. The investment process has three key components: alpha, risk, and costs. The former two components are widely studied by academics and practitioners, whereas the cost factor is not covered enough in academic research, despite the fact that the investment performance is mainly driven by transaction costs. These can be classified as direct costs (i.e. commissions and fees that are explicitly specified and therefore easy to measure, e.g. custody fees, taxes and infrastructure costs) and indirect costs (i.e. costs that are not explicitly specified and that are extremely difficult to measure, e.g. market impact cost and opportunity cost). The most important indirect cost for large trades and the most relevant to this thesis is the market impact of a trader's own actions, which is defined as the effect on market liquidity caused in a financial market by a participant buying/selling an asset. More specifically, investors can move the price against themselves, going upward if buying and downward if selling. The market impact costs are notoriously difficult to measure, but thorough trade management and execution can easily improve market impact costs (Almgren, et al., 2005).

Investors can leak their intention to buy or sell and, for example, the price can go up against themselves when buying large orders, causing market impact. Market impact is crucial for large investors (e.g. financial institutions) when making a trade; if it is large relative to the usual volume of the chosen asset, then traders need to strategically disguise their presence in the stock market (e.g. using hidden orders, where large orders are executed incrementally through a sequence of smaller orders) in order to minimise the transaction cost. For example, a trader might want to buy a large number of shares in an asset in order to profit on a future price increase of that asset. The investment strategy is to buy the shares at the lowest price possible; however, the trader's buy orders would drive the price up (against the trader). Therefore, if the trader splits this large order and executes the order incrementally, a part of the order could be bought at a low price, allowing the trader to minimise the overall cost, until the impact was fully reflected by the market and the price is pushed up (Moro, et al., 2009). Investors typically minimise the market impact by intermediating the transactions through a broker, who charges a commission in order to take responsibility for the trade risk, or by gradually splitting the order and risk leaking information.

The market impact can be analysed by looking at the slippage (also known as implementation shortfall), which is defined as the difference between the final execution price (including commissions, taxes etc.) and the decision price, which can be either the arrival price (preferred by the brokers/trade executors) or the closing price of the day's trading (preferred by the fund managers/decision-makers). The market impact cost is a type of transaction cost that reflects the cost incurred by an index/security trader. Put simply, the market impact cost represents the additional value a trader must pay over the initial price due to the cost incurred by the transaction itself (which moved the asset price).

There are two types of market impact costs: temporary and permanent (Fabozzi, et al., 2010). The temporary market impact cost is a transitory cost consisting of the additional liquidity concession needed for the market maker to take the order. The permanent market impact cost is a consequence of the information content of a trade causing the market participants to adjust their views, resulting in persistent price changes. For example, a sell order might suggest that a security could be overvalued. The total market impact cost is the sum of these two costs.

Traders attempt to decrease the temporary market impact by extending the order's trading horizon. For example, a trader could execute a larger order by splitting it into smaller portions over a longer time period, while ensuring that each portion represents a small percentage of the average trading volume for a given time interval. This method is feasible for orders that are less urgent and it might incur opportunity costs, delay costs and price movement risk.

The goals of agency/broker trading are the minimisation of implementation shortfall and finding liquidity. Their strategies are provided to institutional investors in order to minimise slippage from certain benchmarks (e.g. VWAP, TWAP, or implementation shortfall).

Kyle (1985) introduced a common market impact statistical measure called Kyle's lambda (λ). Kyle assumed a normal distribution for the random variables and derived a linear market impact function. The market dynamics can be perceived as a contest between an insider, who has unique information regarding the fair price of an asset, noise traders, who trade randomly in the absence of actual knowledge, and market makers, who set the prices based on the trading flow. Kyle's lambda gauges the market impact (and price fluctuations) of a trade as a consequence of volume flow. His model is driven by asymmetric information where the order flow moves prices due to uninformed traders' anticipation of trading against informed traders. Kyle's lambda can be approximated using the regression in Equation (2.1), where ΔP_t represents an asset's price change, V_t is the signed trading volume and u_t are the residuals (John, et al., 2015). A positive V_t generates a buy signal ($\Delta P_t > 0$), while a negative V_t generates a sell signal ($\Delta P_t < 0$):

$$\Delta P_t = \alpha + \lambda V_t + u_t. \tag{2.1}$$

Kyle's lambda equation can be further reduced for very short periods as:

$$\lambda = \frac{|\Delta P_t|}{V_t}. \tag{2.2}$$

Kyle (1985) also concluded that modelling price innovation based on trading volume is consistent with modelling price innovation based on new information.

The market impact is modelled in various forms. For example, MSCI Barra (Torre & Ferrari, 1998) define it as a function of trade volume; traded asset parameters, ranging from elasticity (the response of order flow to price signals) and volatility (the variability of an asset's price) to intensity (how often an asset trades) and shape (the distribution of trade sizes); market tone (the price of liquidity), and the investor skill (the impact of an investor's trading process on the cost). Market impact increases with several different factors in industry based on the company defining the model, but the common denominator across these market impact models is the relative order size, i.e. the order size divided by the expected daily volume (Deutsche Bank, 2008):

- Spread, volatility, and relative order size (Bloomberg);
- Variance, relative order size, and trading rate (JP Morgan);
- Relative order size, volatility, trading rate, and spread (Deutsche Bank).

The micro-structure theory investigates how specific trading mechanisms influence the price formation process. In stock markets, price formation is mainly based on the principle of demand and supply; a security's price is the value a potential buyer is willing to pay for that asset.

Multilateral Trading Facilities

Multilateral trading facilities (MTFs) are trading systems offering investors an alternative venue to trading on traditional exchanges, by providing a non-exchange financial trading venue. This term was introduced in a European financial law, i.e. the Markets in Financial Instruments Directive (MiFiD), which defined three trading venue classes: regulated markets, which are run by a market operator; multilateral trading facilities; and systematic internalisers, who do not comply with the former two classes and can decide to share their quotes only with their retail clients, only with their professional clients, or both (European Parliament, Council of the European Union, 2006). MTFs are the European equivalent of the United States' alternative trading systems. They are characterised by the exchange of securities between multiple parties, with less restrictions on the traded financial instruments. Popular MTFs include Turquoise and BATS Chi-X Europe; the latter is derived from BATS Europe's acquisition of Chi-X Europe in 2011 and is no longer considered an MTF per se, as it received the Recognised Investment Exchange status from the Financial Services Conduct Authority (FCA) in 2013. The MTFs' trading volume has now reached considerable levels. In order to get a stock's consolidated trading volume (i.e. the total trading volume across primary exchanges and MTFs), we add up the trading volume from the principal European MTFs: BATS, CHI-X, and Turquoise.

Trading Challenges and Order Routing

Large orders are normally broken into smaller orders and submitted to the market over a period of time in order to minimise the trade's market impact. Smaller orders are more likely to be executed than larger orders, however the delayed execution of breaking a large order into smaller sequential orders can potentially bring additional risks to the trader, e.g. adverse price movements or opportunity cost (i.e. the profits the trader could have incurred by executing a larger order). Therefore, an optimal trading schedule can be translated to achieving the desired balance between price impact and opportunity cost (Nutti, et al., 2011). Alternatively, a large order can be executed in alternative markets (e.g. dark pools and MTFs/alternative trading systems), where the order book is not publicly listed.

Smart order routing (SOR) is a system that automatically routes dually listed stock orders to both the primary and the secondary exchanges in order to minimise market impact. It was promoted as a technological answer to market structure fragmentation, being able to find hidden liquidity without moving the market dynamics. Its aim is to provide best execution for clients, potentially obtaining increased liquidity access and better prices (compared to the price that could have been accessed on one exchange only). The first SOR solutions in Europe were introduced in late 2007 around MiFID's effective date, when Chi-X was starting to compete with the primary exchanges. SOR was improved and, consequently, CHI-X's volumes soared and other MTFs launched in 2008. SOR looks for best order execution by considering price (i.e. holistic view of level 1 and level 2 order books in order to select the venues with the optimal price/volume values), cost and probability of execution and constantly responding in real-time to every market micro-structure change and every trade. Once an order is out, SOR compares the historical data assumptions to the real-time feedback and, in case it identifies an opportunity at a different venue than the originally selected trading venue, it could then decide to move part of the order from the initial location to a more convenient location. Large trading firms have many venues to choose from for executing an order; selecting the best venue depends on the market conditions, business relationships and trading parameters.

Illustrating the concepts of market impact, market fragmentation, market liquidity and the market dynamics sets an appropriate context for examining the potential impact of special events (e.g. futures expiry dates and MSCI index rebalances) on the market dynamics and volumes.

2.1.4 Behavioural Finance

Behavioural finance is a relatively new field of finance that proposes explanations for the systematic irrational financial decisions of market participants, by combining behavioural and cognitive psychology with conventional economics and finance. A large portion of the behavioural finance literature provides extensive studies looking at the link between anomalies (and mainly calendar effects) and prices. Another part of newer research of behavioural finance focuses on sentiment analysis applied to newswires, microblogs and search engine databases.

In the subsequent chapters, we review the empirical studies analysing the impact of calendar effects on stock prices and the relation between prices and volumes in order to infer a direct connection between calendar effects and trading volume, since there is not enough evidence on the calendar effects on trading volume. However, this section broadly introduces the field of behavioural finance (e.g. efficient market hypothesis and common biases and anomalies), which will be extended in the following chapters when focusing on specific calendar effects. Algorithmic trading relies on business as usual on the presumption that markets are efficient. Behavioural finance though acknowledges that the financial markets switch from one state to another, causing disturbance to equilibrium. Eugene Fama developed the Efficient Market Hypothesis (EMH) in the 1960s and defined an efficient capital market as a market that is efficient in processing information. In an efficient market, prices 'fully reflect' the available information at that time (Fama, 1969).

EMH implies that prices jump to the right level to reflect new information immediately in an efficient market and nothing happens as a result of non-news, such as an index fund manager buying stock that was added to an index or a fund manager selling stocks after recent withdrawals. For example, even if additions to stock market indices do not affect a stock's fundamentals, when a stock is included in S&P 500, its beta with respect to other S&P 500 stocks increases, while its beta with respect to non-S&P 500 stocks decreases (Barberis, et al., 2005). Although the market price is the outcome of supply and demand, there is a large amount of inexplicable price volatility. Finance traditionalists would explain this by risk factors, while behaviourists would point the finger at market anomalies (i.e. stock return patterns violating the market efficiency).

In the light of new evidence of market inefficiency, the relatively new field of behavioural finance was born. It explains why people make irrational financial decisions, by applying theories of behavioural and cognitive psychology along with the theories of traditional economics and finance. Notable contributors to this field include Daniel Kahneman, Amos Tversky, Robert Schiller and Richard Thaler. The former two are considered the fathers of behavioural finance. Behavioural finance is the study of how investors make decisions and the effect of these decisions on stock prices and broad market movements. It explores the irrational behaviour of investors, which can potentially jeopardise the market efficiency; for example, Barber, Odean and Zhu (2009) found a correlation in trading behaviour of almost 75% in their sample data.

De Bondt and Thaler (De Bondt & Thaler, 1994) illuminate the psychology of decision-making and reiterate the most important anomalous behavioural concepts within investment theory and corporate finance. The most common cognitive biases that behavioural finance researchers have been analysing include: small sample bias, framing effect, extrapolation, anchoring, conservatism, overconfidence effect, non-Bayesian forecasting, optimism bias, limited attention, recency illusion, confirmation bias, status quo bias, loss aversion, mental accounting, familiarity bias etc. These come into play into the investment process twice: firstly when estimating, and secondly when making an investment decision, as illustrated in Figure 2.3 (Goldman Sachs Global Investment Research, 2013).

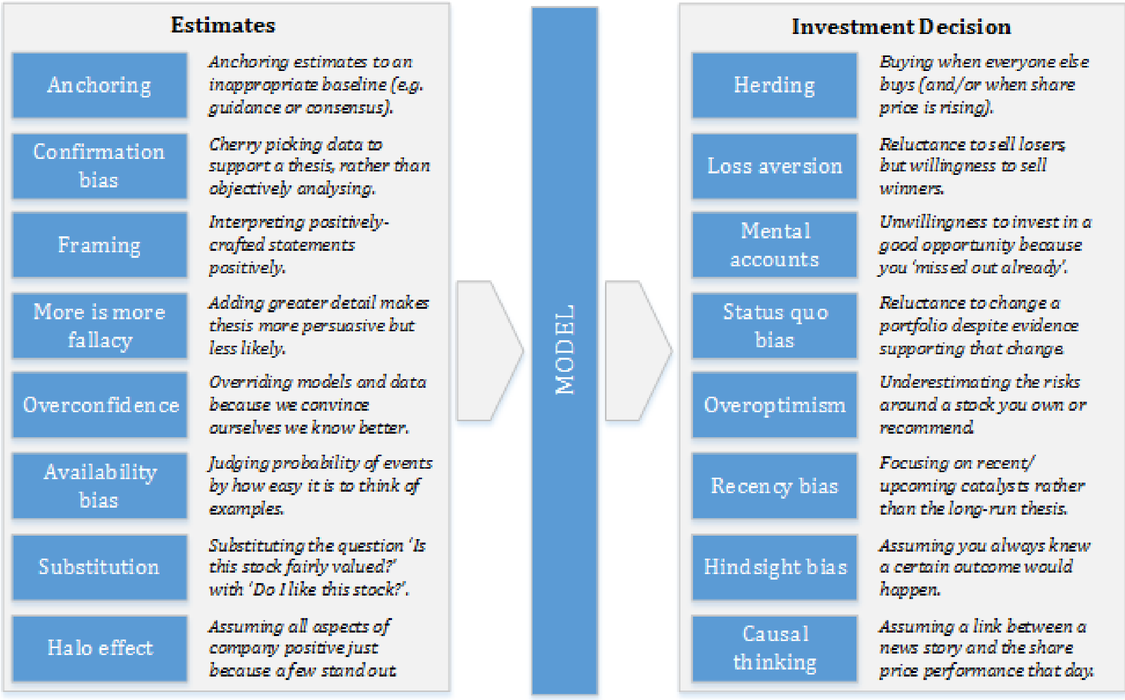

Figure 2.3. Biases in the investment process (Source: Goldman Sachs Global Investment Research).

The analysis of these biases stems from the cognitive nature of human mind, which comprises two systems. The first system is automatic and almost effortless, and judges according to familiarity. On the other hand, the second system works methodically, requiring effort and focus. These two systems collaborate continually, however their interaction is not always smooth (Kahneman, 2012).

Anomalies are evidence of market inefficiency, which can occur either only once or frequently, following a cyclic pattern. Popular anomalies include the calendar effect, medium-term momentum, post-earnings announcement drift, value effect, long-term reversal, size effect etc. This wide range of variables, which have the power to predict stock returns without a clear relation to risk, have been replicated in a large pool of samples and are now deemed as established facts.

It should be noted that behavioural finance consists of a high collection of empirical facts and isolated one-off models, calling for academic consensus and cohesiveness.

2.1.5 Trading Performance

Trading performance in algorithmic trading systems is a grey area and there is not much material published on this topic. We outline a few ways of measuring trading performance and reiterate the importance of liquidity in achieving optimal trading performance (i.e. minimal market impact). The impact is measured by the extent of slippage to arrival. This can be minimised by optimally sizing the trading order related to the volume being traded. If it is too small, the trader underperforms and misses out opportunities to make profits; conversely, if the trader over-participates, the trader can move the price against themselves. This stresses once again the need for accurate prediction of trading volume, which would result in better order sizing and order timing.

Trading performance is typically measured by the trade's risk-adjusted returns. The best-known performance ratio was developed in 1966 by William Sharpe (1966) and later became an industry standard known as the Sharpe ratio. It is calculated by subtracting the risk-free rate from a portfolio's rate of return and dividing this difference by the standard deviation of the portfolio returns. Other metrics for assessing the performance of trading models, while incorporating the risk exposure, include maximum drawdown (i.e. the largest decline from equity high to a succeeding equity low), maximum run-up (i.e. the largest increase from an equity low to a subsequent equity high), risk-adjusted rate of return (RAR), reward-to-risk ratio (RRR) etc.

Before an algorithmic trading strategy is implemented in a live environment, it must be back-tested. Back-testing determines the trading strategy performance for past periods as if the trades were executed at those times' market conditions. This testing approach attempts to find vulnerabilities in real-world conditions from the past, while simulating the whole environment; however, no future performance indicators can be predicted using this technique. Forward testing (or walk forward testing) simulates real markets in order to identify the optimal parameter values for a trading strategy (Pardo, 2008).

A trading system needs to be constantly evaluated not only during the testing phase, but also during live trading in order to increase profitability, minimise risks and assess the efficiency of the trading plan. Equity curves are a way of determining the correlation between live trading results and testing results, by displaying the profits and losses over a time period. Even if adequate tests have been performed for in-sample, out-of-sample and forward performance testing, there might still be unexpected events in the live market that can impact performance. These include trader errors (or pilot errors, i.e. when a trader diverts from the trading plan or makes an error while entering an order), technical problems (e.g. Internet/exchange connection problems), unique trading conditions (e.g. salient unexpected news causing spikes in the market activity), or fills (i.e. the expectation that the algo fills hypothetical orders, whereas the live market environment can cause trades not to be filled at all due to slippage, or the difference between the expected price and the actual trading) (Folger, 2013).

2.1.5.1 Market Liquidity

One of the fundamental objectives of stock markets is to provide liquidity, i.e. the ability to quickly convert shares into cash at minimum transaction costs. Liquid markets are characterised by intense trading activity. Measuring liquidity is problematic as there is no specific formula for this; instead, various liquidity ratios based on different asset types are employed to calculate liquidity (e.g. current ratio, quick ratio, cash ratio, and cash conversion cycle). Liquidity measures can be broadly classified as trade-based measures (e.g. trading value, trading volume, frequency/number of trades, and turnover ratio) and order-based measures (e.g. bid-ask spread and order depth). An important shortcoming of trade-based measures is the inability to predict the capacity to trade immediately and the associated transaction cost due to its ex post nature (i.e. they indicate trading values from the past). Aitken and Comerton-Forde (2003) provide evidence that the order-based measures are better proxies for market liquidity.

Market quality is assessed by both efficiency, i.e. transaction costs (spreads, price impact, latency, fees) and price discovery (information share and common factor share of an execution channel, resiliency, price volatility), and integrity, i.e. market manipulation (ramping, churning, squeezing, rumours, disclosures), information leakage (insider trading, trading ahead of price-sensitive announcements), and broker/client conflict (front running, payment for order flow) (Harris, et al., 2011).

Closing auctions are single-price Dutch auctions that maximise the amount of tradable stock by matching buy and sell orders. They enhance the accuracy of the price discovery process by consolidating the order flow at a single point in time. Closing call auctions concentrate liquidity at this time and reduce the chances of manipulating the closing price, which is the most commonly quoted price and is used for portfolio valuations and for evaluating fund performance. Aitken et al. (2005) concluded that closing call auctions are consolidating the liquidity, while having no economically significant effect on the cost of trading (i.e. the bid-ask spreads at the end of the trading day).

2.1.5.2 Performance Benchmarks

Michael Aitken's Market Quality Dashboard (Hall, 2013) predicts the potential effect of a market design change, by measuring the contribution to market fairness and efficiency using academic measures including transaction cost and price discovery for market efficiency, and insider trading and market manipulation for market fairness. It compares the values of the metrics before and after implementing a market change.

2.2 Method Development

This section provides the foundation of the statistical methodology employed in this research. We outline the analysis problems and the solutions that traditional and more recent approaches offer. This methodological survey starts by outlining the distinction between time series and cross-sectional data (i.e. continuous distribution analysis) and follows with a review of supervised learning methods, focusing on a variety of regression techniques. The section ends with a contrast between the in-sample and out-of-sample analyses and an outline of the randomisation test methodology.

2.2.1 Dynamic (Time Series) vs. Static Data

This research makes intensive use of time series data and therefore it is important to emphasise the difference between the ordered (i.e. dynamic) and unordered (i.e. static) representation of data.

Time series data (also called longitudinal data) are a representative example of dynamic data (or the 'sequential data' class); they consist of a series of data points that occur at successive points in time and at regular intervals (e.g. every day for the daily market data). Dynamic data points cannot be considered independent and identically distributed. In our financial forecasting context, we would like to predict the next values in a time series based on the previous values, but we expect that recent observations are more relevant to our prediction compared to more historical observations (or randomly sampled observations). Therefore, temporal ordering of data is important in the analyses that are carried out in this research, where we need to consider the state evolution dynamics by capturing the temporal structure.

Conversely, static (or cross-sectional) data consist of one-dimensional variables, which are collected by observing several subjects either at the same point in time, or by disregarding the temporal dimension. In such situations, one can assume that each point was generated independently and identically, and that there is no natural (or temporal) ordering associated with these data points (Roweis & Ghahramani, 1999). Since data ordering is random, such an analysis would disregard any temporal dependence and permuting the ordering of the data points would not result in any information loss.

This distinction is important in order to understand the methodology by which the models are cross-validated in the following studies. We first compute the feature matrix (i.e. observations of inputs and outputs) and then partition the data for cross-validation. This process can be either random or stratified, depending on the input data characteristics. Some models exhibit sparse data and therefore the data needs to be cross-validated using stratified partitioning in order to make sure that each fold contains roughly the same proportions of variables.

2.2.2 Machine Learning

Machine learning can be classified into three broad categories:

- Supervised learning: the algorithm builds a prediction model based on the training data set, which includes the input data and the response values (i.e. labelled training data). Predictions are based on a training sample of previously solved cases, where a response variable is provided for a given set of predictors (i.e. the supervisor provides the correct target variables or the degree of error for each observation). There are two classes of algorithms: classification and regression. Regression algorithms represent the central focus of this thesis and we will describe the basic concepts in the following section.
- Unsupervised learning: the algorithm draws inferences from the training data, where the input data does not have any labelled responses associated. There are no correct target variables provided for the input data and the main goal is to infer the properties of the probability density of the input variables (from the data itself) without a training sample. Another contrasting aspect is the success measuring. This is easily accomplished in supervised learning (e.g. cross-validation), whereas there is no direct measure of success in unsupervised learning and therefore it is hard to assess the inferences drawn from unsupervised learning algorithms. Decorrelation, decomposition, dimension/complexity reduction and clustering are some problems that unsupervised learning techniques deal with. Common unsupervised learning algorithms include k-means clustering, Gaussian

mixture models, principal component analysis (PCA), factor analysis, independent component analysis (ICA), and hidden Markov models (HMM).

- Reinforcement learning: the learning agent interacts with a dynamic environment and its algorithms are provided with a reward function, which indicates when it is performing well, and when it is performing poorly. The algorithm is never presented with the correct input-output pairs.

2.2.2.1 Multiple Linear Regression

Ordinary least squares (OLS), or linear least squares (LLS), is a statistical method that estimates the features' unknown parameters in a linear regression model by minimising the sum of the squared vertical distances between the linear approximation of the response values and the observed response values. Unlike the multivariate linear regression, where the model predicts multiple correlated dependent variables, multiple linear regression predicts a single scalar variable. In Equation (2.3), the model predicts the scalar output Y, using an intercept (or bias) β_0, and p column vectors of input variables X_i:

$$f(X) = \hat{Y} = \beta_0 + \sum_{i=1}^{p} X_i \beta_i. \tag{2.3}$$

For convenience, the intercept is included as a constant variable 1 in the input matrix X and therefore its coefficient β_0 will be included in the coefficients vector β, which will now have $p + 1$ elements.

Gauss-Markov Theorem

The Gauss-Markov Theorem states that the coefficients' best linear unbiased estimator (BLUE) is given by the OLS estimator if and only if the OLS errors:

- Have expectation zero;
- Are uncorrelated;
- Have equal variances.

The best estimator is the one that gives the lowest variance of the estimate. The regression errors do not need to be normally distributed, nor do they need to be independent and identically distributed; instead, they only need to be uncorrelated with mean zero and homoscedastic, and with finite variance.

2.2.2.2 Autoregressive Moving Average

The autoregressive models in this thesis are based on the traditional autoregressive-moving-average (ARMA) model and therefore it is important to understand how ARMA models work. They are parsimonious models comprising two polynomials, one for auto-regression and one for moving average, which describe a weakly stationary stochastic process. The model has two hyperparameters and is usually referred to as ARMA(p, q), where p is the order of the autoregressive model and q is the order of the moving average model. As shown in Equation (2.4), the ARMA model contains both the AR(p) and MA(q) models, where X_t is the given time series, β_0 is the constant/intercept, β_i are the AR parameters, β'_i are the MA parameters, and ε_t is the random variable of white noise error terms:

$$X_t = \beta_0 + \sum_{i=1}^{p} \beta_i X_{t-i} + \sum_{i=1}^{q} \beta'_i \varepsilon_{t-i} + \varepsilon_t. \tag{2.4}$$

A peculiar version of ARMA is the autoregressive moving-average model with exogenous input (ARMAX), which includes a linear combination of the last b terms of the time series d_t, which is known and external to the autoregressive variable. Equation (2.5) shows the addition of the exogenous time series into the ARMA model:

$$X_t = \beta_0 + \sum_{i=1}^{p} \beta_i X_{t-i} + \sum_{i=1}^{q} \beta'_i \varepsilon_{t-i} + \sum_{i=1}^{r} \beta''_i d_{t-i} + \varepsilon_t. \tag{2.5}$$

We employ a slightly modified autoregressive model in the first study, such that we employ the moving average of the time series data X_t (in order to smooth it) and not the white noise error terms ε_t (as it is traditionally used in the ARMA model). The autoregressive models are calculated by using the observed lagged time series data, consisting of lagged time series and lagged smoothed time series instead. For better clarity, we are not labelling this type of models as 'ARMA' and, instead, we describe it as a prediction model for trading volume based on time-lagged observations, both raw (i.e. autoregressive past observations) and smoothed (i.e. moving average of the last observations). The lagged smoothed time series part acts as a low-pass filter effect in the data and represents our rationale for using a different variable (i.e. the time series data and not the error terms) for the moving average model.

2.2.2.3 Goodness of Fit

The models' goodness of fit can be compared with root mean squared error (RMSE), or root-mean-square deviation, which estimates the standard deviations of the error distribution by measuring the differences between an estimator's predicted values and the actual values. The individual errors or differences are called "residuals" when computed in-sample, and "prediction errors" when computed out-of-sample. The RMSE of an estimator/model $\hat{\theta}$ is defined in Equation (2.6) as the square root of the mean square error (between the estimated value $\hat{\theta}$ and the real value θ):

$$\varepsilon_{\mathrm{RMSE}}(\hat{\theta}) = \sqrt{E\left((\hat{\theta} - \theta)^2\right)}. \tag{2.6}$$

For n predictions in a time series, the model RMSE of the predicted values $\hat{y}_t$ of the regression's dependent variable y_t at time t is computed as the square root of the mean of the deviations' squares, as in the following equation:

$$\varepsilon_{\mathrm{RMSE}}(\hat{y}) = \sqrt{\frac{1}{n}\sum_{t=1}^{n} (\hat{y}_t - y_t)^2}. \tag{2.7}$$

The RMSE is computed slightly different for regression analysis. Here it refers to the unbiased estimate of the error variance; the denominator is $n-p$ instead of n, representing the degrees of freedom, where n is the sample size and p is the number of the estimated model parameters (i.e. p regressors). The denominator becomes $n-p-1$ if an intercept is used when estimating the model.

Another indicator of goodness of fit for statistical linear models is the coefficient of determination, or R-squared (denoted R^2), which indicates the proportion of total variation of outcomes explained by the model, ranging from 0 to 1. If the linear regression includes an intercept, then R-squared is calculated as the square of the sample correlation coefficient between the real values and their predicted values. R-squared is the square of the coefficient of multiple correlation in case more explanatory variables are included. R-squared, whose formula is shown in Equation (2.10), is the ratio of the total sum of squares, exhibited in Equation (2.8), and the residual sum of squares (RSS), shown in Equation (2.9), subtracted from 1:

$$S_{\text{total}} = \sum_i \left(y_i - \frac{1}{n}\sum_{i=1}^{n} y_i \right)^2 \tag{2.8}$$

$$S_{\text{residuals}} = \sum_i (y_i - \hat{y}_i)^2 \tag{2.9}$$

$$R^2 \equiv 1 - \frac{S_{\text{residuals}}}{S_{\text{total}}}. \tag{2.10}$$

The closer the value of R-squared is to 1, the better the regression model fits the data in comparison to the average (i.e. the mean of the observed data). However, R-squared is not ideal to evaluate time series prediction models because it mainly reflects how well a model fits past values, but it does not indicate how well it predicts future values. In theory, one could obtain a perfect fit for a model by adding an increasing number of regressors, most probably leading to poor prediction performance when fitting the model outside the sample.

2.2.2.4 Variable Selection

Variable selection, also known as feature selection, attribute selection or variable subset selection, is a machine learning and statistical technique that selects a subset of relevant features in order to construct a model. It assumes that the input data has many redundant (i.e. features that do not provide any more information than the currently selected features) and/or irrelevant features (i.e. features that do not provide any relevant information regardless of the context). The main benefits of variable selection include faster data training and cost-effective predictors, improved model generalisation (by reducing the model overfitting), and improved prediction performance. Unnecessary features add noise to estimating the target variable and degrees of freedom are wasted; therefore, the reduced model (i.e. with the lower number of features) is the best model.

Guyon and Elisseeff (2003) outline a few useful heuristics that need to be taken into account when performing variable selection:

- Variables that are perfectly correlated are completely redundant since adding them brings no extra information.
- Variables that are anti-correlated or very highly correlated might provide additional (and complementary) information when they are all added to the model.
- A variable that is totally useless alone might provide significant performance improvement when considered together with others.

The features must be pre-processed before performing variable selection: identify and potentially exclude outliers and transform variables that seem appropriate. For instance, price log-ratios are used instead of the raw price values and log-volumes are used to smooth/normalise the volume data in this research.

The performance of the feature subsets is typically evaluated using an objective function, which is based on the trade-off between two competing criteria: 1) the (maximum) goodness-of-fit; 2) the (minimum) number of variables/degrees of freedom.

The simplest and most rudimentary method tests all the possible subsets of features and returns the subset that minimises the error rate. If there are n potential predictors, then there is a total number of 2^n possible models. This method performs an exhaustive search of the space and therefore it is computationally consuming and intractable for large numbers of variables. Criterion-based procedures fit all the models (although special best subsets algorithms do not fit all 2^n models) and output the best model according to some criterion, such as the Akaike Information Criterion, the Bayes Information Criterion, Adjusted R^2, Predicted Residual Sum of Squares (PRESS), Mallow's Cp Statistic etc. Another simple method is sequential search, which replaces each variable at a time with all the remaining variables in order to assess whether the model has been improved.

Stepwise Regression

Stepwise regression is a regression model using greedy search algorithms, whose predictor variables are selected automatically, using statistical techniques such as sequences of F-tests/t-tests, adjusted R-square, Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC) etc. There are three methods of performing stepwise regression:

1. Forward selection starts from an empty model with no variables and tests whether the addition of each new variable improves the model, repeatedly adding the variable that brings the highest improvement at each step, until no variable can improve the current model anymore;
2. Backward elimination is the opposite of forward selection and starts with a full model comprising all the explanatory variables and tests whether the elimination of each variable improves the model, repeatedly removing the variable whose elimination brings the highest improvement at each step, until no variable's elimination can improve the current model anymore;

3. Bidirectional elimination is a combination of the first two methods, testing for variables to be included or excluded at each step. This starts with an empty model and adds a variable with the largest F-statistic at each step. After adding a variable at each iteration, the algorithm tries to eliminate any variable that is not significant (i.e. backward elimination)

Stepwise regression commences from a starting model, which could be a constant model (i.e. only a constant/intercept term), a linear model (i.e. an intercept and linear terms for the predictors), an interaction model (i.e. the intercept, linear terms and all products of pairs of distinct predictor variables) etc., and then it sequentially adds and removes predictors. It performs forward selection to add new variables having *p*-values that are less than a predefined value for the improvement measure for adding a term, and backward elimination to eliminate features from the model whose *p*-value is greater than a given improvement measure for removing a term. One drawback of stepwise procedures is that the optimal model could be missed due to the 'one-at-a-time' approach of including/excluding variables.

MATLAB's sequential feature selection function starts from an empty feature set and creates possible feature subsets by selecting each of the features that are not yet included in the feature set. It tests each feature subset by performing 10-fold cross-validation.

More specific variable selection techniques include genetic algorithms, particle swarm optimisation, ant colony optimisation, least absolute shrinkage and selection operator (LASSO), and elastic net (which combines the penalty terms of LASSO and ridge regression).

The LASSO is a regression method that minimises the RSS (like OLS) and, similarly to ridge regression, it constrains the sum of the absolute values of the coefficients to be less than a given constant. LASSO performs variable selection because the coefficients shrinkage process is designed to shrink some variables' coefficients to exactly zero and eliminate them.

Part of this research employs variable selection on models that include dummy variables (e.g. day-of-the-week: Monday-Friday). If a categorical predictor has *n* possible values, we do not encode the categorical variable into *n-1* dummy variables. Instead, we use *n* dummy variables because the ultimate purpose is to perform feature selection on them and see which ones are relevant. This is consistent with Hastie et al. (2011), who represent an *n*-level qualitative variable by a vector of *n* binary variables (or bits), where one and only one of them is 'on' at a time.

2.2.2.5 Shrinkage Methods

Regularisation is used for model selection by introducing additional information, i.e. penalty for complexity (e.g. restrictions for smoothness or bounds on the norm of the solutions), in order to prevent model overfitting. The most common regularisation techniques are the L_1 and L_2 regularisers, which are additional information introduced to regression models that minimise a loss function $E(X, Y)$ by minimising $E(X, Y) + \alpha\|w\|$ instead, where w is the model's vector of weights, α is a free parameter whose value is determined empirically (e.g. by cross-validation), and ||·|| is either the L_1 norm or the squared L_2 norm. The resulting linear regression models that are fit by applying this method are called ridge regression or LASSO. Other popular regularisation techniques for model selection are AIC, BIC, and minimum description length (MDL). A very simple example of regularisation is the least squares method.

The L_q Space

The L_q space represents a finite-dimensional vector space. The most important regularisation methods are ridge regression (L_2 regularisation) and the lasso (L_1 regularisation). The lasso applies an L_1 penalty in order to recover sparse solutions and perform feature selection, whereas ridge regression applies an L_2 penalty to shrink coefficients toward zero.

The contours of the constant value for five values of q are exhibited in Figure 2.4. The constraint region for ridge regression is a disk, while the lasso has a diamond region. Since lasso's constraint region has corners, if the solution occurs at a corner, then a coefficient β_j is set to zero. When $q > 2$, the diamond becomes a rhomboid, having many corners, flat edges and faces. Hence, there can be many more parameters estimated to be zero (Hastie, et al., 2011). These constraint regions are sections of 3-dimensional L_q balls, e.g. $q = \infty$ corresponds to a cube and $q = 2$ to a sphere (this is the common Euclidean distance). The volumes of the L_q balls decrease with q.

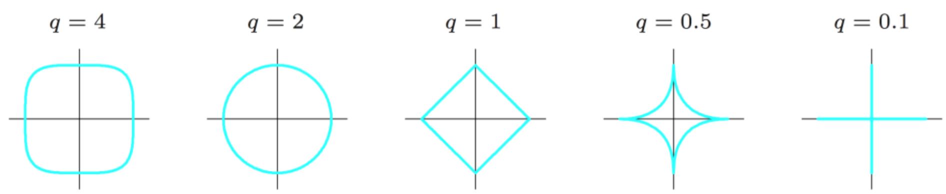

Figure 2.4: Contours of the L_q family (Hastie, et al., 2011).

Ridge Regression

Another simple form is ridge regression (or Tikhonov regularisation, or linear regularisation), which is a trade-off between better fitting the data and reducing the norm of the solution. The standard approach is OLS, which aims to minimise the residuals $||\beta X - Y||^2$ for the regression problem $\beta X = Y$, where ||·|| is the Euclidean norm. Ridge regression includes a regularisation term in the OLS minimisation, $||\beta X - Y||^2 + ||\Gamma X||^2$, where Γ is an empirically chosen matrix, called Tikhonov matrix. Usually, the Tikhonov matrix is chosen to be the identity matrix, i.e. $\Gamma = I$, aiming for solutions with small norms.

Lasso, Elastic Net and LAR

The lasso (Tibshirani, 1996) is another shrinkage method, similar to ridge regression. However, it differs from ridge regression in a few ways: the L_2 ridge penalty $\sum_1^p \beta_j^2$ is substituted by the L_1 lasso penalty $\sum_{j=1}^p |\beta_j|$; lasso performs a sort of continuous variable selection, setting some coefficients exactly to zero, unlike ridge regression, where some coefficients are valued close to zero, but not exactly zero.

Lasso treats and estimates β_0 in the same way as ridge regression and therefore the predictors are standardised and the model is fit without an intercept. It is defined as:

$$\beta_{\text{lasso}} = \operatorname{argmin}\left\{\frac{1}{2}\sum_{i=1}^{N}\left(y_i - \beta_0 - \sum_{j=1}^{p} x_{ij}\beta_j\right)^2 + \lambda\sum_{j=1}^{p}|\beta_j|\right\}. \tag{2.11}$$

Another method regarded as a compromise between ridge regression and lasso is the elastic net penalty (Zou & Hastie, 2005), which is defined in Equation (2.12). Elastic net selects variables like the lasso, and shrinks together the coefficients of correlated predictors like the ridge regression:

$$f_{\text{penalty}} = \lambda\sum_{j=1}^{p}\left(\alpha\beta_j^2 + (1-\alpha)|\beta_j|\right). \tag{2.12}$$

Another relatively new method is the least angle regression (LAR), which was introduced by Efron et al. (2004) and can be regarded as a 'democratic' version of forward stepwise regression (Hastie, et al., 2011).

2.2.2.6 Other Common Supervised Methods

Support Vector Machines

Support vector machines (SVMs) construct a hyperplane (or a set of hyperplanes) in a high-dimensional space, which is used for both classification and regression analyses. SVMs are a representative algorithm of the class of kernel methods (the name of this class is linked to using kernel functions, which allow the operation in a high-dimensional space without calculating the data coordinates in that space, but rather by simply computing the inner products between the images of all pairs of data points in the feature space). A hyperplane with the largest distance to the nearest training data point of any class provides the optimal separation. One of the advantages of SVMs is that despite the nonlinear optimisation of the SVM training, the solution of the optimisation problem is quite straightforward because the objective function is convex (Bishop, 2007).

In the context of regression, in order to obtain sparse solutions using SVMs, the quadratic error function is replaced by an ϵ-insensitive error function (Vapnik, 1999). This function gives zero error if the absolute difference between the prediction and the target variable is less than ϵ (where $\epsilon > 0$). The optimisation problem can be also expressed by introducing slack variables (ξ_n and $\hat{\xi}_n$), which are defined for each data point x_n. For a prediction $\hat{y}_n$ with a correct target value y_n, ξ_n corresponds to a point where $y_n > \hat{y}_n + \epsilon$, and $\hat{\xi}_n$ corresponds to a point where $y_n < \hat{y}_n - \epsilon$. This allows the creation of the ϵ-insensitive 'tube' and the target points (y_n) lie inside the tube if $\hat{y}_n - \epsilon \leq y_n \leq \hat{y}_n + \epsilon$.

k-Nearest Neighbours

The k-nearest neighbours algorithms (kNN) is a non-parametric method for classification and regression. For a given query point, the input consists of the k closest training points (in terms of the feature space distance). In kNN classification, the output is a class membership, which is determined by the majority vote among its neighbours. In kNN regression, the output is the average of the values of the k nearest neighbours. A special case of this method is the 'nearest neighbour' when $k = 1$. The kNN method is memory-based and does not require a model to be fit. Since the entire training data needs to be stored, this method is computationally expensive when the data set is large.

2.2.3 In-Sample and Out-of-Sample Analyses

The first three experiments are in-sample analyses that use the entire data set in order to find structure in the data. The final study extends the findings of the in-sample analyses in order to go out-of-sample and predict the target variables on previously unseen data.

2.2.3.1 In-Sample Analysis

The in-sample analysis investigates all the observations in order to best fit a model for generating a target variable (e.g. trading volume). This analysis finds structure in the data; for instance, by performing stepwise regression, the model outputs the coefficients for the most significant features. This research evaluates the results of the in-sample analysis by cross-validation, which is described in the next section.

2.2.3.2 Cross-Validation

Cross-validation is a standard model validation technique for evaluating the extent to which the results of a model generalise on independent data in order to avoid overfitting. The cross-validation starts by partitioning the data in two subsets, one for the training set, which is used to learn the model, and one for the testing set, which is used to evaluate the trained model's predictive accuracy. It performs the model analysis on the training set and validates it on the testing set. Multiple rounds of cross-validation with different training and testing sets are performed and their results are averaged in order to reduce variability. A particularly important technique is k-fold cross-validation (the value of k is commonly 10), which partitions the data in k equally-sized data subsets, containing k-1 subsets for the training set and a single subset for the testing set; the process is repeated k times (i.e. the cross-validation folds) such that each subset is used for the testing set once and only once, and eventually the results of the k rounds are averaged for the final model estimation. The obvious advantage of this method is the use of every observation for both training and testing the model. 10-fold cross-validation is used across the analyses of this thesis. The stepwise objective function (or loss function) of the cross-validation of the studies included in this thesis computes the mean squared error (MSE) and averages it over the k rounds.

2.2.3.3 Out-of-Sample Analysis

The out-of-sample analysis starts by taking the in-sample analysis model and applies it on unseen data in order to confirm whether the model works in real-life conditions. The main benefit of out-of-sample forecasting is its trustworthiness, unlike the in-sample evidence where the models can suffer of small sample bias and can be sensitive to outliers. Due to its nature, out-of-sample is designed to provide improved performance on new observations and therefore its robustness is significantly higher. Moreover, out-of-sample analyses are of predictive nature and can be employed in real-time forecasting systems to provide estimates ex ante, whereas in-sample analyses are based solely on past data, finding data patterns ex post.

Given an out-of-sample setting, the data set is split up into two disjoint subsets: a training set for inferring the in-sample model and a testing set for applying the in-sample model on previously unseen data (i.e. data that were not used in training the model). This research employs two methods for assessing the performance of the out-of-sample analysis (i.e. the prediction accuracy):

- Sliding/moving/rolling window. This technique loops over the entire data set and, at each step, it discards the oldest/first training observation and adds the next data point to the training set; the testing data set is the first observation following the training set. This technique refreshes the model over time and trains it on the most recent data, instead of relying on a static training set from a further point in the past.
- Growing window. Unlike the sliding window technique, this method does not discard the oldest/first training observations at each iteration; instead, the training window keeps growing by a new data point at each iteration. This model assumes that historical data are relevant.

Figure 2.5 illustrates two loops/iterations over the sliding and growing window techniques, with a starting training set size of 7 observations and a fixed test set size of 1 observation.

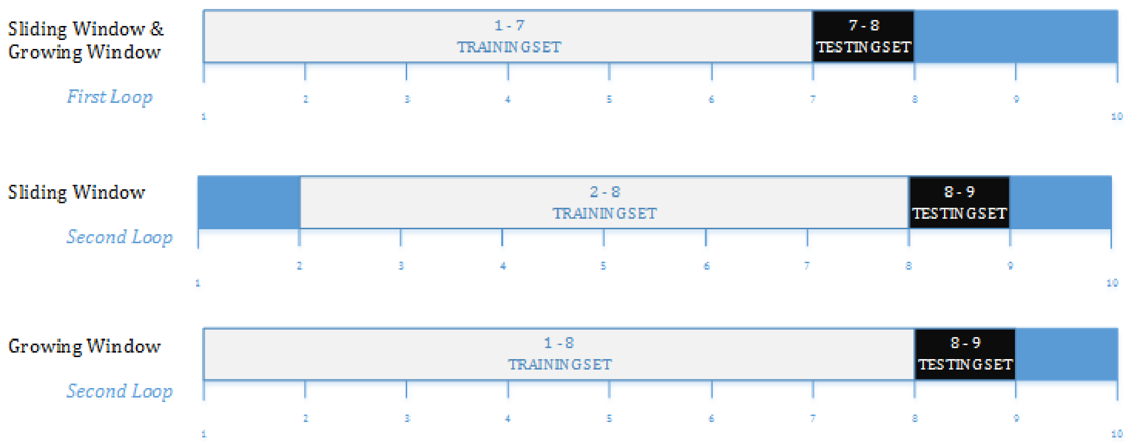

Figure 2.5: Out-of-sample testing techniques: sliding window vs. growing window.

2.2.4 Multiple Comparisons Problem

The multiple comparisons problem (or multiplicity/multiple testing problem) occurs when a set of statistical inferences are considered simultaneously or when a subset of parameters is selected based on the observed variables. When statistical tests are used repeatedly, the occurrences of type I errors increases. Therefore, the goal of multiple comparisons corrections is to decrease the number of false positives.

It is important to understand and compare two main controlling procedures of type I errors. The first method is the familywise error rate (FWER), which is the probability of making at least one false discovery (i.e. type I error) throughout all the hypotheses while employing multiple statistical tests. The 'family' term refers to a set of inferences. The second method is the false discovery rate (FDR), where the expected proportion of incorrectly rejected null hypotheses (i.e. false discoveries) is controlled. The main difference between these two controlling procedures is that FWER controlling procedures provide a more stringent control of false discoveries compared to the FDR procedures, because the FWER methods control the probability of at least one false discovery.

There are other popular procedures (Hochberg & Tamhane, 1987), such as the Šidák correction, Turkey's procedure, the Holm's step-down procedure, Hochberg's step-up procedure etc.

The multiple comparisons problem can be corrected by recalculating the probabilities of a statistical test that was used multiple times. The most common FWER controlling procedure is the Bonferroni correction, which adjusts the familywise error rate. Let p_i be the p-value for rejecting a test H_i. Considering that there are m repeated tests, the Bonferroni procedure rejects test H_i if $p_i \leq \frac{\propto}{m}$.

However, a more powerful correction that adjusts the false discovery rate is the Benjamini-Hochberg procedure (Benjamini & Hochberg, 1995). The false discovery rate approaches recalculate the p-values and the resulting values are called q-values. The Benjamini-Hochberg procedure controls the FDR at level α. It is a step-up procedure, since the p-values of the m null hypotheses tested are sorted in increasing order, denoting the reordered p-values by the new rank: $P_{(1)}, \dots, P_{(m)}$. Then the procedure finds the largest k such that $P_{(k)} \leq \frac{k}{m}\alpha$ for a given α, which represents the chosen false discovery rate.

However, a more attractive approach is to use the permutation distribution (Hastie, et al., 2011) – this is our preferred method and is employed in the cross-market holiday study and in the expiry day study. This method makes no assumptions regarding the data distribution. This approach consists in computing a large number (usually 1,000 iterations are used in practice) of permutations of the sample labels and computes the test statistic for each permutation. There is more background information about the permutation test (or resampling) in a subsequent section.

2.2.5 Measures of Central Tendency and Dispersion

Two frequent types of descriptive statistics are used to describe the trading volume data, namely measures of central tendency and dispersion.

A measure of central tendency (or location) is the central or typical value of a probability distribution. We employ this measure to determine the tendency of the observed data to cluster around a central point. Although the arithmetic mean, the geometric mean and the mode are common measures of central tendency, we use the median since it is very robust in handling outliers. Other measures of central tendency that deal with outliers include the truncated mean, interquartile mean, trimmed mean, and Winsorised mean.

The geometric mean was also a candidate to aggregating the benchmark volumes. It consists of a set of numbers and uses their product in order to denote their central tendency. It is defined as the n^{th} root of the numbers' product, as shown in Equation (2.13), where the formula on the right-hand side shows the geometric mean expressed as the arithmetic mean of logarithms (sometimes called log-average). The latter representation is preferred in computational statistics since it overcomes overflow or underflow problems arising when many numbers are multiplied:

$$\left(\prod_{i=1}^{n} a_i\right)^{\frac{1}{n}} = \sqrt[n]{a_1 a_2 \dots a_n} = \exp\left(\frac{1}{n}\sum_{i=1}^{n} \ln a_i\right). \tag{2.13}$$

The other measure is for statistical dispersion (also called spread, variability or scatter) and indicates how stretched or squeezed a distribution is. Popular measures of dispersion include the sample standard deviation, range, interquartile range (IQR), mean absolute deviation, median absolute deviation etc. The latter two measures are abbreviated MAD, where one is about the mean and one is about the median. Absolute deviation is defined as the absolute difference between each data point and a central point (either the mean or the median). The mean absolute deviation and median absolute deviation, whose formulae for an input vector X are shown in Equations (2.14) and (2.15), are considered better than the standard deviation in representing real life data sets, where $\bar{X}$ is the mean of the data. The absolute deviations are robust estimators of dispersion which are unaffected by outliers:

$$\text{mean absolute deviation: } m_{\text{mean}} = \tfrac{1}{n}\textstyle\sum_{i=1}^{n}|X_i - \bar{X}| \tag{2.14}$$

$$\text{median absolute deviation: } m_{\text{median}} = \text{median}(|X_i - \text{median}(X)|). \tag{2.15}$$

IQR and the two MAD measures are very robust to outliers. As noted by Leys et al. (2013), the median absolute deviation is easier to implement and it is a more robust measure of dispersion than the standard deviation since it is less sensitive to outliers. Identifying outliers using the mean plus/minus three standard deviations involves a few issues; using the mean as the central tendency measure assumes data, including outliers, follow a Gaussian distribution, is severely impacted by outliers and does not cope well with small samples. However, the median is very insensitive to outliers and the median absolute deviation is immune to sample sizes, making it a more robust scale estimator than the interquartile range.

We used the median to measure the dispersion of the benchmark period trading volume while conducting the special date analyses (i.e. cross-market holidays, stock index futures expiries, and MSCI rebalances) and exploring the various benchmark volume aggregation possibilities, as it provided marginally better results than median absolute deviation.

2.2.6 Randomisation Tests

A common way of assessing the statistical significance of results is to employ the Student's t-test. However, this parametric test assumes that data points come from a normal distribution. Unfortunately, this is not the case of the trading volume data. Alternatively, non-parametric tests, such as the Wilcoxon signed-rank test, the sign test, or the Mann–Whitney U (also called Wilcoxon rank sum test), could be applied as they cope with unknown distributions.

Another alternative to the parametric Student's t-test is to employ a randomisation test (also called permutation test) and examine the empirical p-value. A randomisation test is a statistical significance test that performs random rearrangements (or permutations) of the data in order to validate the results based on a data sample. The permutations consist of exchanging labels on data points. The test calculates a very large number of possible values of the test statistic under rearrangements of the labels throughout the observed data points. Under the null hypothesis, the data permutations have no effect on the outcome, and the reshuffled data exhibit the same properties as the observed data (Bordino, et al., 2012). The empirical p-value is determined by the rank of the observed test statistic among the randomised test statistics. This value indicates the probability that the test statistic is at least as extreme as the observed value, given that the null hypothesis is true. For our chosen significance level of 0.05, if the original statistic is greater than 95% of the randomised statistics, then the null hypothesis H_0 is rejected with confidence $p < 0.05$, meaning that the probability of observing a value as extreme as the original test statistic is less than 5%. The aim of the randomisation test is to assess whether the observed difference between the sample of means of the observed data vector and a control data vector is large enough in order to reject the null hypothesis that these two vectors have the same probability distributions. The randomisation test methodology is based on the works of R. A. Fisher (1935) and E. J. G. Pitman (1937).

All of the in-sample studies in this thesis are based on randomisation tests, where we produce reshuffled data by defining a target data set (i.e. the observed data) and a paired control data set (i.e. artificial data conditioned on the target data). The randomisation of the data set occurs over 1,000 iterations, resulting in 1,000 reshuffled bootstrap vectors. Given two vectors consisting of target data (denoted by X) and control data (denoted by Y), the randomisation test starts from the real data and computes the (original) observed test statistic as $\mathrm{mean}(X) - \mathrm{mean}(Y)$. We further take the absolute value of this difference if the test is two-tailed (i.e. the aim is to show that the means of the two vectors are significantly different, but we are not particularly interested in the test sidedness). Under the null hypothesis, the randomisation has no effect on the observed statistic. We compute the 1,000 reshuffled test statistics using the same equation we used to define the original test statistic, by replacing the data labels between the target and control vectors. The test eventually compares the original test statistic (i.e. the difference in means between the original observed data) with the reshuffled statistics and computes the proportion of rearrangements when the original statistic is larger (for the two-tailed test, where the aim is to prove the alternative hypothesis that $X \neq Y$, and the right-tailed test, where the alternative hypothesis is $X > Y$) or smaller (in the case of the left-tailed test, i.e. when we want to prove that $X < Y$) than the reshuffled statistics. If this percentage is greater than the confidence level of 95%, then the null hypothesis is rejected, as the statistic lies in the upper tail. It is this percentage subtracted from 1 that constitutes the empirical p-value. If the p-value is less than the significance level of 0.05, then the observed data difference is statistically significant.

Some randomisation tests in this thesis are pairwise, meaning that the values/elements of the two vectors (i.e. the target vector and the control vector) are interchanged on a controlled pairwise basis. For each element, the test flips a coin in order to decide whether the values for the current index are shuffled. This is mainly because certain data points are peculiar to a time period or feature and we need to control the corresponding value to interchange with. In this case, each permutation is composed of pairs (x_i, y_i) for each time series observation at index i, which are used to randomly pair the target vector with the control vector.

For other randomisation tests, the values by which the target vector is potentially replaced with are not conditioned on the particular target element and therefore the order does not matter. In this case, we perform a random permutation of the values of the two vectors.

3. Examining Drivers of Trading Volume

This study presents an in-depth exploration of market dynamics and analyses potential drivers of trading volume. The study considers established facts from the literature, like calendar anomalies, the correlation between volume and price change, and this relation's asymmetry, while proposing a variety of time series models. The results identified some key volume predictors, such as the lagged time series volume data and historical price indicators (e.g. intraday range, intraday return, and overnight return). Moreover, the study provides empirical evidence for the price-volume relation asymmetry, finding an overall price asymmetry in over 70% of the analysed stocks, which is observed in the form of a moderate overnight asymmetry and a more salient intraday asymmetry. We conclude that volatility features, more recent data, and day-of-the-week features, with a notable negative effect on Mondays and Fridays, improve the volume prediction model.

3.1 Introduction

This study investigates the drivers affecting the trading volume with an in-sample analysis. We explore the interaction between truly exogenous determinants and trading volume. Several hypotheses are evaluated while looking at the previous literature, where various factors are discussed in isolation, and we propose a liquidity extraction model by placing these findings in a unified context.

Identifying the drivers of trading volume is crucial in order to anticipate and minimise market impact, by accurately sizing and executing orders. Achieving optimal order sizing relies on precise volume prediction, i.e. planning trades and deciding how much to trade given the current market context and the predicted volume levels. To better illustrate the importance of trading volume, some recent facts include the total turnover value, which was $63tn in 2011 (World Federation of Exchanges, 2012) and $49tn in 2012 (World Federation of Exchanges, 2013). The NYSE's turnover averaged more than 100% between 2004 and 2009, with 138% in 2008 (NYSE Euronext, 2016), meaning that the entire market value has changed hands once a year, although it has decreased to significantly lower levels during the following years, averaging 72% for the 2010 – 2015 period.

In order to better understand the factors affecting the trading volume, it is necessary to survey and combine apparently disjoint literature concepts. We start by reviewing the relevant areas of the behavioural finance literature. Here, a large amount of research has mainly investigated the calendar effects on price returns and there is very little emphasis on the calendar effects on trading volume. We particularly focus on the day-of-the-week effect, which, once investigated, can formulate several hypotheses to analyse other calendar effects (e.g. the effect of stock index futures expiries and cross-market holidays). We then connect the behavioural finance findings with evidence from the literature on the relation between price changes and volume (i.e. the price-volume lead-lag effect). Following this reverse path, we test the direct relation between calendar effects, represented in this study by the day-of-the-week effect, and volume.

Behavioural finance mainly consists of regression models built on a collection of indicator variables, implying a certain limitation with regard to its statistical significance. We propose a model based on lagged time series and lagged smoothed time series in order to explain observed volumes in terms of recent time series; this follows the behavioural finance paradigm and represents market dynamics on the run, while assuming stationarity and disregarding outliers. However, the financial data is a strong non-stationary and non-constant mean time series, due to the existence of notable event dates (e.g. MSCI rebalance dates, futures expiry dates, company earnings announcement dates etc.). This analysis aims to bridge the gap between behavioural finance and traditional finance and explores the feasibility of a potential special event effect (e.g. futures expiries or cross-market holidays) on trading volume by starting with an analysis of the day-of-the-week effect on trading volume. The financial markets are event-driven and their dynamics are permanently shifting. Therefore, it is important to predict the trading volume at different time horizons.

The main motivations of this study include: the insufficiency of literature looking at the calendar effects on trading volume (and not on returns), the inconclusive results of the price-volume relation and whether it is characterised by asymmetry, and the abundance of studies investigating certain volume determinants in complete isolation from other types of volume drivers.

Out of a total number of fifty-five surveyed articles, which are all cited in this experiment and investigate the price-volume relation and the day-of-the-week effect, only seven of them use data sets after 2000 and none of the cited papers employs market data after 2006. Moreover, only seven studies include a few European stocks or indices among their international data sets, and only two papers are based on European data sets exclusively. Given the lack of a broad European stock universe and post-2000 data sets, we employ an extensive pan-European stock universe consisting of 2,353 stocks, for which we use daily market data between 1st January 2000 and 10th May 2015, and we also test for structural breaks by comparing the results before and after the financial crisis of 2007-08.

The aim of this study is to define a unified volume prediction model, while exploring the endogenous variables in conjunction with exogenous variables and performing feature selection. We investigate a pan-European stock universe for a sample period of over 15 years in order to test the improvement of an autoregressive volume model, by sequentially adding features, such as volatility, more recent data, and day-of-the-week, and test additional hypotheses such as the existence of an asymmetric price-volume relation. The rest of the study is organised as follows: section 2 provides a literature review of the main research topics addressed by this study: volume dynamics, price actions and volume-price correlations, along with a survey of the relevant calendar effects; section 3 introduces the sample data set, while section 4 outlines the main models and the analytical approach; section 5 describes the methodology of the trading volume study, while gradually introducing the various variables we are examining in order to better predict the volume; section 6 exhibits the main results of the various volume prediction models and the types of price-volume asymmetry; section 7 provides a conclusion of this study, together with a discussion on the results and potential suggestions for other researchers to further extend this study.

3.2 Background Literature

The literature review starts by setting the context of this study, i.e. why volume prediction is important, followed by a review of studies on the types of the price-volume relation and its potential asymmetry. We then switch to the behavioural finance literature by outlining the main calendar effects and elaborate on the day-of-the-week effect.

3.2.1 Trading Volume Historical Dynamics

Trading volume is extraordinarily large across developed stock exchanges and many interesting patterns in prices and returns are closely related to the volume movement; volume is highly used in conjunction with price actions. For instance, the volume of high-priced 'glamour stocks' tends to be larger than the volume of low-priced value stocks, and a stock with higher trading volume tends to have lower future returns (Hong & Stein, 2007). Trading volume is a strong indicator of economic activity.

Auctions account for a high trading volume and there are three types of auctions: opening auctions, intraday auctions and closing auctions. A normal day starts with pre-trading auctions, or opening auctions, in order to set the price after the non-trading hours during the night, when news came out, and is followed by continuous trading. In Europe, this phase can be temporarily halted by volatility interruptions, which trigger a 2- or 5-minute auction, called intraday auction, in case the price is changing more than +/- 5%, in order to set the price correctly.

The literature has mostly examined the relation between trading volume and the following three variables: 1) bid-ask spread (i.e. negatively correlated); 2) price changes – the literature has predominantly found a positive correlation between volume and the absolute price change; 3) information – a volume increase means that the investors interpret the information either differently or identically by beginning with different priors; the market institutional design affects volume around the informational events (Karpoff , 1986).

3.2.2 Volume-Price Relation

The relation between trading volume and prices is important in order to better understand the financial market structure. Price changes indicate the market response to new information, while the trading volume measures the level of disagreement of the information among investors (Beaver, 1968).

There is wide evidence in the literature (Harris & Raviv, 1993) (Hong & Stein, 2007) for a positive correlation between trading volume and price dynamics. Volume has been found to be positively correlated either with the magnitude (i.e. absolute value) of the price change (Assogbavi & Osagie, 2006), i.e. $|\Delta p|$, or with the price change per se (i.e. the raw value of the price change), i.e. Δp (Karpoff, 1987), where price changes can be represented as log-price difference or percentage price change.

Moreover, Godfrey et al. (1964) reported a modest correlation between volume and the difference between the daily high and low, i.e. the intraday range, which will also be employed as a price metric in this study.

3.2.2.1 Volume-Price Relation Asymmetry

Given that there is a potential relation between trading volume and price change, we further investigate the price change representation, by taking into account the evidence on the asymmetric relation between trading volume and price changes, as previously shown in the finance literature. Having an in-depth understanding of the asymmetry in price indicators helps fit the models more accurately. More specifically, instead of having a single feature for the magnitude of a price indicator (e.g. intraday or overnight price returns), we can discriminate between positive and negative values, and represent the magnitude of each sign independently, i.e. one feature for the magnitude of positive values and another feature for the magnitude of negative values, which would result in two different coefficients when regressing. The following two equations summarise the symmetry and asymmetry of the volume/price relation:

$$\text{symmetry: } (v_t|\Delta p_t^+) = (v_t|\Delta p_t^-) \tag{3.1}$$

$$\text{asymmetry: } (v_t|\Delta p_t^+) > (v_t|\Delta p_t^-) \text{ or } (v_t|\Delta p_t^+) < (v_t|\Delta p_t^-). \tag{3.2}$$

Ying's early work (1966) was the first to draw attention to the asymmetry in the volume-price relation, finding evidence that price asymmetry exhibits greater volume when associated with a price increase than when associated with a price decrease. Furthermore, Jain and Joh (1988) used hourly volume and returns data from NYSE, concluding that the relation between volume and price returns is steeper for positive price changes. This was subsequently confirmed by other researchers, such as Epps (1975) (1977), Smirlock and Starks (1985) and Al-Deehani (2007), who also offered a possible interpretation for the price-volume relation asymmetry, namely the fact that short sellers respond faster to information stimulating price change than long investors, causing higher volume on price upticks. Conversely, Woord et al. (1985) and Moosa et al. (2003) found a reverse asymmetry, where the volume/price change ratio is smaller for upticks than for downticks in stock markets. The results of these studies imply that the absolute price changes must depend on whether the price change is positive or negative, which will be an important decision in the price indicator variables across this study. Other common explanations for the existence of asymmetry include the optimistic and pessimistic investors' disparity of opinion or the higher costs of short selling compared to the costs of taking long positions (Assogbavi, et al., 1995).

3.2.3 Calendar Effects

This research builds on the top of previous literature by merging disjoint findings on volume prediction, price-volume relation and the asymmetry of this relation. However, we complement these findings by introducing some exogenous variables from the behavioural finance literature that could potentially drive trading volume. Given the abundance of isolated papers analysing the calendar effects on price returns and the papers discussing the price-volume relation, there is an auspicious context for exploring a direct relation between calendar effects and trading volume. Moreover, in this particular study, the day-of-the-week effect is investigated in a volume prediction context, along with the endogenous variables based on time-lagged volume and price-related metrics; this is opposite to most of the behavioural finance articles on calendar effects, where the authors define dummy variables only for particular effects in complete isolation from the endogenous predictors. Discovering an explicit relation between trading volume and the days of the week would allow us to subsequently investigate the effect of futures expiries and cross-market holidays, since their abnormal returns might potentially explain the day-of-the-week effect and they could impact on the trading volume. The market is typically in a steady state with a relatively constant price formation process that drives the fairly expected price and volume metrics. However, when special events occur, such as the futures expiries or cross-market holidays, the market is in a different condition during these days and calls for a state-switching model.

3.2.3.1 Behavioural Finance and Calendar Effects

The behavioural finance literature introduced several anomalies (e.g. calendar effects) that affect prices. This contradicts the traditional paradigm that markets are efficient, and suggest that markets switch to different states that disturb the equilibrium.

In the 1960s, Eugene Fama introduced the efficient market hypothesis (EMH), defining an efficient market as one that efficiently processes information, i.e. prices fully reflect the publicly available information at a given time (Fama, 1969). This hypothesis is shared among the finance traditionalists and was the driver of an opposing view from behaviourists, who explored various stock return patterns that violate the market efficiency. The field of behavioural finance explains the decision-making process of investors and its consequences on the market movements.

Behaviourists analysed a huge amount of samples of past market data and identified evidence of market inefficiency in the form of anomalies, which can either occur once or follow a periodic pattern.

The most popular anomalies include the calendar effects, medium-term momentum, value effect, size effect, post-earnings announcement drift etc.

The calendar effects are market anomalies that involve a sudden change in the behaviour of stock markets at certain times of the year. These event-driven irregularities have been documented in a wide range of studies. Some of the most interesting calendar effects include the weekend effect (and more generally the day-of-the-week effect), the month-of-the-year effect, the January effect, the holiday effect (and more specifically the cross-market holiday effect), the expiry day effect, and the intra-month effect.

3.2.3.2 The Weekend (Day-of-the-Week) Effect

The weekend effect (or Monday effect) consists of a lower closing price on Monday than the closing price of the previous Friday. It is a particular instance of the broader day-of-the-week effect. The literature on calendar effects focuses on the connection between these effects and returns; extremely few articles investigate the impact of calendar effects on trading volume and hence it is important to first understand the findings on calendar effects and price returns, and then connect them with the insights on the price-volume relation, in order to infer a direct link between calendar effects and trading volume.

The weekend effect is intriguing because empirical results contradict the expectation to have higher returns on Monday, since its returns reflect three consecutive days. The average return for Monday is negative (French, 1980) (Gibbons & Hess, 1981) (Jaffe & Westerfield, 1985) (Pettengill, 2003), as Cross (1973) first indicated that Monday returns are significantly different from Friday returns. The weekend effect has been widely documented in the literature (Dubois & Louvet, 1996) (Harris, 1986) (Abraham & Ikenberry, 1994). Other authors found non-cyclical patterns for the day-of-the-week effect, which could be explained by other calendar effects: Rogalski (1984) found that the day-of-the-week returns are connected to the January, firm size and turn-of-the-year anomalies, while Wang et al. (1997) found that the Monday effect occurs mainly in the last two weeks of the month (i.e. the fourth and fifth weeks).

Contrarily, Steeley's research (2001) suggests that the weekend effect in the UK stock prices has disappeared after 1990, while Smirlock and Starks conclude that this weekend return is positive (1986). More confusingly, Brusa et al. (2000) confirmed the existence of a weekend effect for small firms, but reported the existence of a reverse weekend effect for medium- and large-sized firms, where Monday returns are positive and significantly greater than the average of the other four weekdays.

Berument and Kiymaz (2001) found a day-of-the-week effect in both returns (with highest returns on Wednesday and lowest returns on Monday) and volatility (with highest volatility on Friday and lowest on Wednesday); later, they discovered the maximum and minimum days are different across international markets, with highest volatility occurring on Thursdays in the UK (Berument & Kiymaz, 2003). As for the relation between the day-of-the-week and the trading volume, Lakonishok and Maberly (1990) found a relative increase in the individuals' trading activity on Mondays.

Potential justifications of the strong Monday effect include general measurement-error explanations (Keim & Stambaugh, 1984), the delay between trading and stocks settlement, and in clearing checks (Lakonishok & Levi, 1982), the individual investors' trading pattern (i.e. selling pressure) on Monday (Lakonishok & Maberly, 1990), and, partially, the institutional behaviour (Flannery & Protopadakis, 1988) (Sias & Starks, 1995).

Most of the literature on calendar effects consists of an ample collection of studies conducted on isolated one-off models applied to certain past samples of market data. Because the calendar effects are highly data-driven and the inter-dependence of economic variables is ambiguous, the calendar effects have been investigated ex post and usually the stock universe of the data samples used in the studies is too narrow in order to draw a generalised conclusion. Besides the small stock universe, most of the studies are conducted on older sample periods. However, the market structure keeps changing and what happened in the 1970s might not be valid anymore. This motivates this study to consider structural breaks around the financial crisis of 2007-08, and we fit the models for 2000-2007 and 2008-2015 in order to explore potential structural breaks.

3.3 Data Set

The analysis is conducted on a pan-European stock universe comprising 2,353 stocks (7,197,065 daily observations) with price and volume market data for the time period between 1st January 2000 and 10th May 2015. The midpoint of our data set coincides with the financial crisis of 2007-08, whose peak consisted of the collapse of Lehman Brothers on 15th September 2008. Therefore, we investigate a potential structural break in the market dynamics before and after the crisis, by splitting the data set into two subsets: the pre-crisis subset (1st January 2000 – 31st December 2007) and the post-crisis subset (1st January 2008 – 10th May 2015).

3.3.1 Data Acquisition

The analysis market data for the extensive pan-European stock universe consists of the constituents of the most important European indices, along with a comprehensive index from Thomson Reuters. The indices' constituent list is compiled as of 10th May 2015 and does not contain historical evidence of index additions and eliminations throughout the entire duration of the experiment; this list is a representative stock sample for the European stock market. The final stock universe consists of the list of unique constituents of the indices included in Table 3.1, along with their RICs (Reuters Identification Codes).

Table 3.1
The European indices whose constituents were part of the study data sample.

RIC	Index Name	RIC	Index Name
.STOXX	STOXX Europe 600 EUR Price Index	.PSI20	Euronext Lisbon PSI 20 Index
.FTSE	FTSE 100 Index	.OMXS30	OMX Stockholm 30 Index
.FTMC	FTSE Mid 250 Index	.OBX	Oslo Stock Exchange Equity Index
.FTLC	FTSE 350 Index	.OMXHPI	OMX Helsinki_PI
.FTSC	FTSE Small Cap Index	.BFX	BEL 20 Index
.FTAS	FTSE All Share Index	.OMXC20	OMX Copenhagen 20 Index
.GDAXI	Deutsche Boerse DAX Index	.ATG	Athex General Composite Share Price Index
.MDAXI	MDAX Performance Index	.ISEQ	ISEQ Overall Price Index
.SDAXI	SDAX Share Index	.JTOPI	Johannesburg Stock Exchange Top 40 Tradeable Index
.FCHI	CAC 40 Index	.ATX	Austrian Traded Index
.CN20	CAC Next20 Index	.FTMIB	FTSE MIB Index
.CACMD	CAC Mid 60 Index	.MSPE	MSCI International Pan Euro Price Index
.CACS	CAC Small Index	.MCX	MICEX Composite Index
.SSMI	Swiss Market Index	.WIG20	Warsaw SE WIG-20 Single Market Index
.AEX	Amsterdam Exchanges Index	.TRXFLDEUPU	Thomson Reuters Europe Index
.IBEX	IBEX 35 Index		

The daily market data was retrieved from Thomson Reuters Eikon by developing a VBA script to automate the process of stock-specific data retrieval. The VBA script takes a list of desired indices as input, retrieves their constituents and then, for each stock, it returns the daily market data (i.e. opening, high, low, and closing prices, and trading volume) for the 15 years covered by this study (i.e. 1st January 2000 – 10th May 2015). The data set was further extended using these stocks' primary Reuters Identification Codes (RICs) and attaching their MTF (i.e. Multilateral Trading Facilities) RICs for the following MTFs: BATS, CHI-X, and Turquoise. Then, we retrieved the daily prices and trading volumes for each new MTF RIC.

The market data ranges from 1st January 2000 to 10th May 2015, comprising the daily summary of corporate actions-adjusted volumes (e.g. controlling for stock splits, stock dividends, mergers and acquisitions, spinoffs, rights issues etc.).

The price and volume data is shown in Figure 3.1, using some illustrative time series snapshots of market data for Barclays PLC. There are three charts for market data: Panel A is a candlestick (or OHLC) chart for the daily price data for September 2013, where a solid body candle stick shows that the closing price is greater than the opening price, and a hollow body candle stick indicates the closing price is less than the opening price; the corresponding trading volume data for September 2013 is illustrated in Panel B; and, panel C shows a multi-year plot for price and volume, ranging from 4th January 2000 to 8th May 2015.

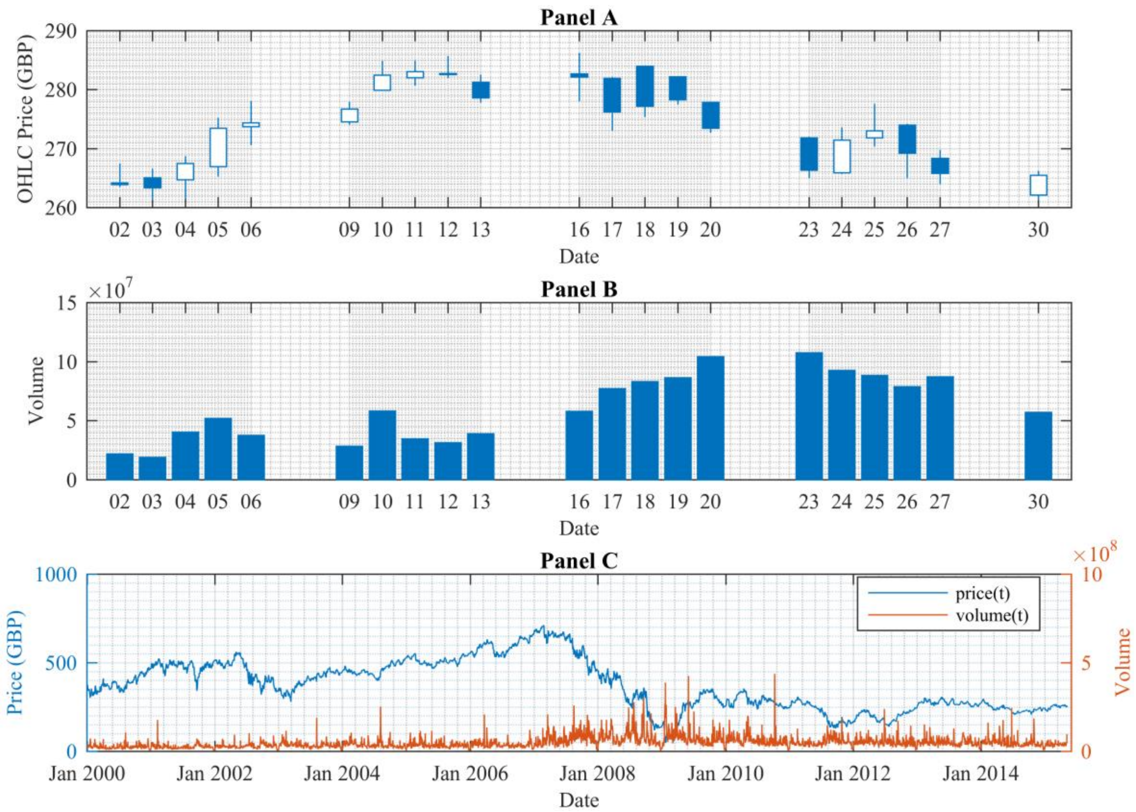

Figure 3.1. Price and volume data for Barclays PLC (BARC.L).

3.3.2 Data Pre-Processing

The market data pre-processing stage starts by eliminating the instruments for which there is no available market data. There were 595 stocks without MTFs, and 194 MTF RICs with no market data. Missing data points are checked in the primary exchange volume and price data (e.g. zero volumes and Thomson Reuters data retrieval errors), and in the MTF volume data only, as the MTF's market data is only used for computing the consolidated volume and the MTF prices are not of interest. The consolidated volume is then computed for each stock by summing up the primary volume and the MTF trading volume. The primary exchange volume is hereafter replaced by the consolidated volume for all stocks. The market data is further processed by discarding the stocks whose number of days of available market data is less than 100 trading days.

We include South Africa in the stock universe due to its liquidity and high level of similarity with European stocks, as it is sharing the same time zone with Eastern Europe.

Throughout the volume analysis, we will be using the logarithmic trading volume due to the high non-normality and outliers of linear volume; from this point forward, we will refer to log-volumes only. Taking natural logarithm of the volumes and price ratios helps normalise the errors, as it reduces skew. Figure 3.2 shows the graphical evidence of the change in the data distribution for the entire Barclays trading volume data set containing 3,878 observations, where the log-volume becomes closer to Gaussian; Panel A illustrates the highly skewed distribution of trading volume. The Barclays volume data was chosen as an illustrative example of a 'liquid stock' and it is not particularly important whether it is Barclays, Vodafone, BP or another stock. The histogram can be generated for an individual stock only and cannot aggregate all of the stocks used in the experiment because of their different price and volume magnitude. Figure 3.3 depicts the distribution of the regression raw residuals for the models of two stocks, i.e. Telefonica SA (TEF.MC) in Panel A and Total SA (TOTF.PA) in Panel B, allowing for the visualisation of the residual rescaling.

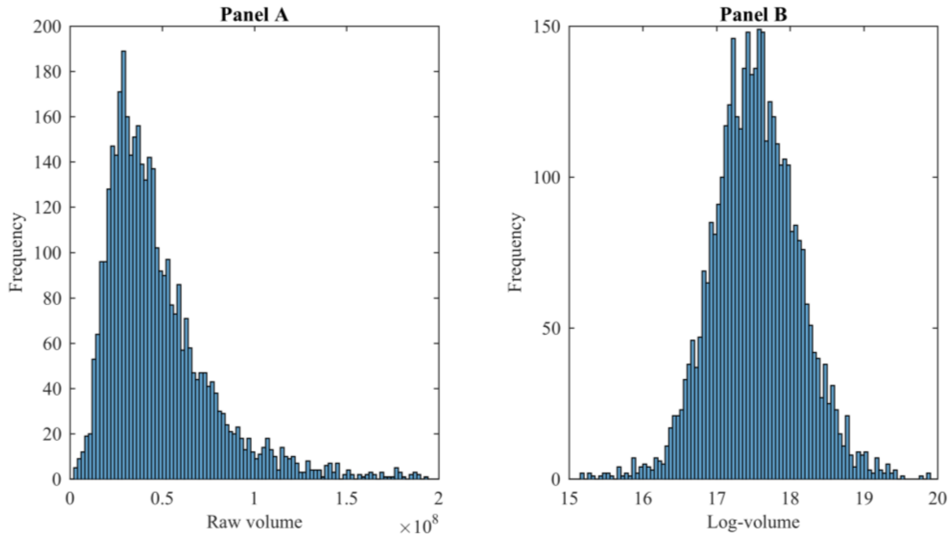

Figure 3.2. Histograms of the raw volume data (Panel A) and the corresponding logarithmic volume data (Panel B).

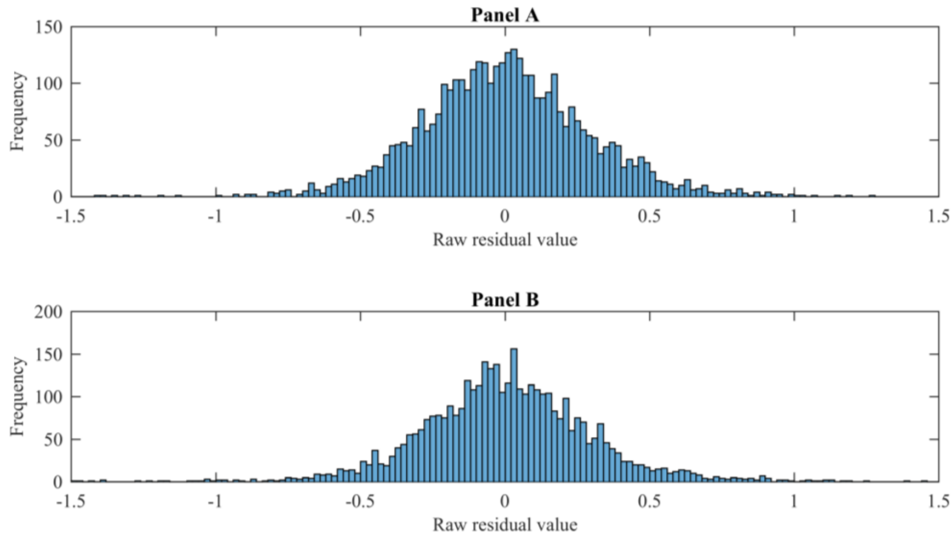

Figure 3.3. The distribution of OLS raw residuals for the historical dynamic model for Telefonica SA (TEF.MC) between 25th January 2000 and 8th May 2015 (3,874 observations) in Panel A, and the historical dynamic and day-of-week model for Total SA (TOTF.PA) between 24th January 2000 and 8th May 2015 (3,908 observations) in Panel B.

Given that the market data provided by Thomson Reuters does not cover the auction volumes, it is impossible to compute a highly accurate breakdown of trading volume breakdown, although these could be approximated by getting tick data and aggregating their values based on the millisecond timestamp, e.g. same time and price values for the first points of the day for the opening auction, and only the ticks at 16:35 (UK time zone) for the closing auction. Therefore, we use the total daily volume as the dependent (or response) variable in our analysis. It includes all the trades executed for the day and it disregards the overnight and off-market trades.

Furthermore, since the data for opening auction volume is unavailable, we define the overnight return as a proxy for the opening auction volume in order to quantify the improvement of more recent data. The overnight return is divided by the number of intervening nights in order to account for non-trading days (i.e. bank holidays and weekends). We explore two variants of defining the overnight return, one that applies a correction (by dividing by the number of intervening nights) and one that is not corrected, which includes an additional variable for the number of extra nights.

3.4 Aims of Study and Analysis Approach

The objective of this study is to explore a prediction framework to understand what drives the trading volume. This study's linear regression framework tests a variety of hypotheses using various factors, which are ultimately reduced through feature selection. This is an exploration of the endogenous and exogenous factors affecting trading volume and it is important to note that the effect size is not our main concern in this study. For each stock in our pan-European universe, different models are compared in order to accomplish the best explanation, while keeping as few predictors as possible and eventually identifying a parsimonious model.

The proposed framework conducts a stock-specific analysis, where each stock is investigated by fitting different stock-specific models, performing feature selection and model comparison. Eventually we report the overall findings and provide a summary of the pan-European stock universe analysis, despite having a per-stock approach.

The stock-specific analyses were normalised by representing the different effects for each stock and account for idiosyncrasies; models vary for each stock independently. The normalised results were aggregated across the 2,353-stock universe.

The methodology for model comparison consists of 10-fold cross-validation (CV), where the objective function seeks to minimise the average mean squared error (MSE). We used cross-validation even for nested models, as it is more robust (instead of an F-test, which assumes Gaussian errors and is sensitive to non-normality) and avoids overfitting. The CV folds are defined at the beginning of the analysis and they are constant throughout the various models that are fit for each stock. After defining the 10 CV folds, we perform stepwise regression on the various models (i.e. multiple linear regression, followed by sequential feature selection). Feature selection reduces dimensionality by producing a reduced model fit on fewer variables, while minimising the predictive error. Whenever a feature is added to or removed from a model, feature selection performs 10-fold CV at each step in order to guarantee that the overall model error is reduced. The objective function of the sequential feature selection minimises the average MSE across the cross-validation folds. Therefore, features are sequentially added (for forward selection) or removed (for backward elimination) at each step, until no other features can be added or eliminated, while decreasing the criterion (i.e. MSE). Because of the unfeasibility of following an exhaustive approach and fitting all of the possible feature subsets, the sequential feature selection technique moves only in one direction, meaning that the candidate feature set is always growing (in the case of forward selection) or shrinking (in the case of backward elimination).

3.4.1 Randomisation Tests

When looking at the return asymmetry and the magnitude of the overnight return depending on the number of intervening nights, we aim to evaluate whether two data vectors are significantly different or whether they come from a similar distribution. The randomisation (or permutation) tests were employed mainly because they make no assumptions about the data distribution, unlike other popular parametric tests such as the student's t-test, where data points are assumed to come from a normal distribution. The randomisation test is a robust and rigorous statistical significance test, and it is appropriate for this study especially since the log-ratio returns and the volumes are not exactly Gaussian (although they are significantly normalised). Nonparametric tests, e.g. the Mann-Whitney U test, could be used alternatively, but since the p-values are based on approximations and using rankings reduces the information inferred from the data (i.e. information loss), randomisation tests are considered a superior methodology (Edgington, 1964). For two vectors (X and Y), the permutation test computes the observed statistic as the absolute difference of the two vectors. The labels of the data points from vectors X and Y are randomised 1,000 times and, for each reshuffling, we compute the randomised statistics using the same equation as the initial statistic. Finally, the test assesses whether the randomised absolute differences are more extreme than the observed absolute difference, resulting in an empirical p-value; this value represents the percentage of times when the observed absolute difference is greater than the randomised absolute differences for a significance level $\alpha = 0.05$. The randomisation test rejects the null hypothesis if the empirical p-value is less than the significance level (5%).

3.4.2 Model Outline

In this study, we are investigating several factors that could potentially drive the trading volume. Therefore, for each stock, the analysis consists of a number of volume prediction models for hypothesis testing and effect quantification, starting from a basic volume model and expanding it subsequently. All of the models in this study include an intercept unless stated otherwise. For a given date (t_{-1}), the target (or dependent) variable is the next trading day (t_0) logarithmic volume, while the model is trained on past data up to the test date ($t_{-n} \dots t_{-1}$). The regression design matrix is computed for each target day (t_0) and then the cross-validation partitions the target date vector accordingly. The structural breaks we investigate in connection with the financial crisis of 2007-08 do not destroy the cross-validation process, since the feature matrix is computed before partitioning the data and hence it does not interfere with the subsequent data partitioning (e.g. structural breaks or cross-validation). When two or more predictors are linearly dependent, the linear regression sets the maximum number of coefficients to zero in order to obtain a basic solution.

In order to test the statistical improvement of the various potential endogenous and exogenous determinants of trading volume, we start by defining a basic prediction model for trading volume (i.e. the 'volume model') based on time-lagged observations, both raw (i.e. autoregressive past observations) and smoothed (i.e. moving average of the last observations). We employ 10-fold CV to find the optimal orders for the time lags of the autoregressive volume and the time windows for the moving average volume. These volume features, as well as the intercept, are kept in all of the subsequent models when performing feature selection.

Next, the price features for the previous day (i.e. intraday range and intraday return) are added to the volume model and we perform feature selection on these price features. The best model in this state is called the 'state A model'.

Then, we add more recent data in the form of overnight return (as a proxy for opening auction volume) to the full 'state A model' (i.e. the model with the full feature set), and perform feature selection on all price features. The model with the lowest MSE is called the 'state B model'.

Up to this point, we use endogenous variables to fit a volume prediction model. We then switch to exogenous variables (i.e. day-of-the-week) and start from the best model up to this point, i.e. the best model among the state A and state B models, which is called 'the historical dynamic model'.

Two day-of-the-week models are defined. The first one (i.e. the 'raw day-of-week model') is a basic model consisting of five dummy variables for each workday, with feature selection applied to them; there are no volume or price features – this is the traditional model employed in the behavioural finance literature. The second model, called the 'historical dynamic and day-of-week model', adds five indicator variables for Monday-Friday on the top of the historical dynamic model (i.e. the best model between state A and B, based solely on price and volume features) and then performs feature selection on the day-of-week features, while forcing the historical dynamic features to remain in the model.

The feature selection consists of backward elimination for the volume and price models (i.e. state A and state B models). Considering these models have a large feature set, the runtime of backward elimination is significantly faster than the forward selection technique, although their final results (i.e. the reduced models) are almost identical. However, in the case of the day-of-the-week models, we used forward selection due to the design of the dummy variables for the day-of-the-week categorical variable. A categorical variable with n possible values is normally encoded as $n-1$ dummy variables. In this study, the day-of-the-week dummy variables are mutually exclusive since the aim is to perform feature selection and extract the days that are the most statistically significant for volume prediction and this is performed in a feature selection framework. Forward selection is preferred to backward elimination because matrix inversion would not be possible for determining the optimal beta when adding a collinear variable. Using backward elimination and starting from the full feature set would result in collinearity issues.

Table 3.2 includes a preview of the models fit for each stock and their full candidate feature sets. The historical dynamic model includes the model with the lower CV MSE among state A and state B.

Depending on the chosen model, the intraday returns and overnight returns can be defined either symmetrically or asymmetrically. The aim of these models is to determine if a certain variable improves the forecast significantly. The main concern is not to evaluate the effect size (e.g. whether a certain feature improves the prediction accuracy by n shares). The models are stock-specific and their coefficients cannot be generalised to the entire data set. Moreover, we did not standardise the variables as their interpretation would not be meaningful in terms of standard deviations and we believe each stock exhibits slightly different correlations.

Table 3.2
Regression models with full candidate feature sets.

	Volume model	Volume and price models						Historical dynamic model	Historical dynamic and day-of-week model	Raw day-of-week model
		State A models		State B models						
Intercept	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Volume lag ($1 \dots n_1$)	✓	✓	✓	✓	✓	✓	✓	✓	✓	
Volume window ($1 \dots n_2$)	✓	✓	✓	✓	✓	✓	✓	✓	✓	
Symmetric intraday return		✓		✓	✓			✓*	✓*	
Asymmetric intraday returns			✓			✓	✓			
Symmetric overnight return				✓		✓		✓*	✓*	
Asymmetric overnight returns					✓		✓			
Day-of-the-week (1 … 5)									✓	✓

3.5 Volume Analysis

In this section, we describe the methodology for the volume analysis study, while introducing a number of particular aspects we are investigating.

3.5.1 Volume Model

In the initial volume model, we explain observed volumes in terms of recent time series. Essentially, this model aims to predict today's (i.e. t_0) volume using the trading volume of the previous days. Therefore, the model contains a lagged time series part, comprising contiguous volume lags, and a lagged smoothed time series part, comprising contiguous moving average volume windows. The lagged time series part reveals the persistence or autocorrelation in the time series assuming stationarity and the lagged smoothed time series part is a low-pass filter effect in the data; they are both used later to evaluate the potential importance of other features in the model. Equation (3.3) defines the volume model, where V_{t_0} refers to the logarithmic trading volume of the current day and $V_{t_{-1}}$ refers to the log-volume of the previous day; β_0 is the intercept coefficient, β_i^{lag} is the coefficient for the lag(i) feature (i.e. the volume lags), and β_j^{smooth} is the coefficient for the smooth(j) feature (i.e. the moving average volume window). Hence, the volume model is composed of three types of terms: the intercept, the autoregressive lagged predictors, and the moving average lagged smoothed predictors; the model has two parameters, i.e. recentTimeSeries(p, q), where p is the autoregressive lagged order and q is the moving average lagged smoothed order. The model can also be represented using the feature names that correspond to the 'lag' and 'smooth' terms, i.e. volume lag (V_i^{lag}) and volume window (V_j^{win}), respectively, as shown in Equation (3.4), or, even simpler, as in Equation (3.5):

$$V_{t_0} = \beta_0 + \sum_{i=1}^{p} \beta_i^{\text{lag}} V_{t-i} + \sum_{j=2}^{q} \beta_j^{\text{smooth}} \frac{1}{k} \sum_{k=1}^{j} V_{t-k} \tag{3.3}$$

$$V_{t_0} = \beta_0 + \sum_{i=1}^{p} \beta_i^{\text{lag}} V_i^{\text{lag}} + \sum_{j=2}^{q} \beta_j^{\text{smooth}} V_j^{\text{win}} \tag{3.4}$$

$$V_{t_0} = \beta_0 + \text{lag}(i) + \text{smooth}(j). \tag{3.5}$$

The ‘lag’ and ‘smooth’ underlying features are contiguous. The ‘lag’ model comprises autoregressive orders from *1* to p and the ‘smooth’ model comprises the moving average orders from *2* to q. For example, lag*(6)* contains all the volume lags from 1 to 6, and no in-between lag can be excluded by the subsequent feature selection.

In order to build an optimal volume model from time-lagged volume observations, we first need to identify the optimal ‘lag’ and ‘smooth’ orders. These are identified by fitting autoregressive (‘lag’) and moving average (‘smooth’) models up to order 15, i.e. lag(1)... lag(15) and smooth(2)... smooth(15). These models consist of the constant term and the autoregressive or moving average terms. Each of these models is cross-validated and the average MSE is returned, which is then used as the criterion of comparing two nested models at a time in an incremental manner, e.g. lag(1) against lag(2); lag(2) and lag(3); lag(14) and lag(15) etc. We start with the lowest possible order and increment it by one; we compare the pair of models with consecutive autoregressive or moving average orders and, if the full model (i.e. the one with the greater order) statistically improves the reduced model (i.e. the one with the lower order), then we increment the order once again and compare the pair of models with the two largest orders at this point. We repeat this process until incrementing the order does not statistically improve the reduced model anymore. For the autoregressive model, i.e. lag(p), we start with $p = 1$ and compare it to the next integer value, i.e. $p = 2$. If lag(2), i.e. the full model, improves lag(1), i.e. the reduced model, then we increment p and compare lag(2) to lag(3); if it does not improve the reduced model, then we stop the incremental process and pick the lower p of the last comparison pair as the optimal ‘lag’ order. Similarly, we determine the optimal order for ‘smooth’, but we start from $q = 2$, i.e. smooth(2), since smooth(1) is the same as lag(1). Equations (3.6) and (3.7) show the models with consecutive orders (i.e. p and $p + 1$) for the comparison of statistical improvement for detecting the optimal ‘lag’ order:

Reduced model
lag(p)

$$V_{t_0} = \beta_0 + \sum_{i=1}^{p} \beta_i^{\text{lag}} V_{t-i} \tag{3.6}$$

Full model
lag($p + 1$)

$$V_{t_0} = \beta_0 + \sum_{i=1}^{p} \beta_i^{\text{lag}} V_{t-i} + \beta_{p+1}^{\text{lag}} V_{t-(p+1)}. \tag{3.7}$$

This comparison of nested models based on 10-fold CV average MSE tells whether a higher order statistically improves the goodness of fit of the model. The comparison of the CV average MSE is performed in the initial phase instead of stepwise regression in order to enforce the 'lag' and 'smooth' predictors to contain contiguous features (i.e. successive volume lags/windows).

3.5.2 Contribution of Volatility and Volume-Price Asymmetry

After identifying the optimal lag and smooth orders and defining the volume model, we extend it by adding a couple of price metrics for the previous trading day, namely: intraday return (i.e. the difference between a trading day's closing and opening prices) and intraday range (i.e. the difference between a day's high and low prices). Each of these metrics can be represented as percentages or log-ratios. The log-ratio representation was preferred to percentages because the percentage returns cannot drop under -100%, but they can go up over 100% (due to the non-negative nature of prices), and therefore price percentages can lie on the interval $(-100\%, +\infty)$, whereas log-ratio returns can in principle belong to $(-\infty, +\infty)$, providing a better representation for price returns that is closer to the Gaussian distribution. Therefore, we used logarithmic price ratios to compute the intraday return and intraday range price metrics. The log-ratios were calculated by taking the log of the raw price ratios (i.e. p_1/p_2). The formulae for the log-ratio returns, which were used for the features added to the volume baseline model, are:

$$\text{intraday return log ratio: } p_{\text{intradayRtn}} = \log \frac{p_{t-1}^{\text{close}}}{p_{t-1}^{\text{open}}} \tag{3.8}$$

$$\text{intraday range log ratio: } p_{\text{intradayRng}} = \log \frac{p_{t-1}^{\text{high}}}{p_{t-1}^{\text{low}}}. \tag{3.9}$$

It is worth clarifying that the target variable of the model is the trading volume for t_0 based on previous information, i.e. up to and including t_{-1}:

The potential collinearity of the predictor variables (also called multicollinearity, i.e. high correlation) can cause numerical instability in the regression coefficient estimates, as there are large fluctuations in the estimates when a few observations are added or removed. The assumption of regression analyses for a model to have no collinearity can refer to the absence of a perfect collinearity (i.e. an exact linear relation between predictors). More advanced techniques, such as PCA or ridge regression, can handle any potential multicollinearity effortlessly. We check for the absence of perfect collinearity by investigating the correlation between intraday range and absolute intraday return using the 15-year data set for Barclays as an illustrative example, whose scatter plot is shown in Figure 3.4. The correlation coefficient of the two vectors is 0.775 and the plot shows that they are somewhat correlated in the middle, but the models tend to diverge.

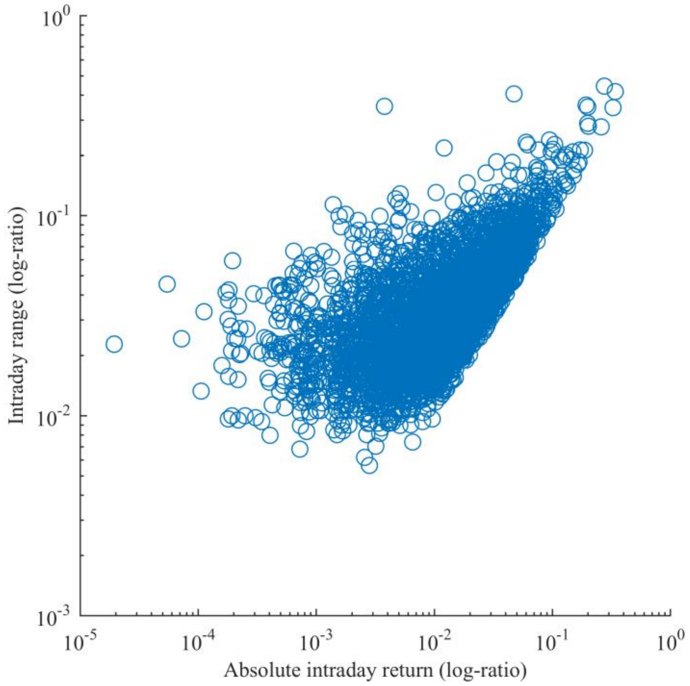

Figure 3.4. Scatter plot of logarithmic scaled absolute intraday return against intraday range for Barclays PLC (BARC.L), between 25th January 2000 and 8th May 2015 (3,863 observations).

Heteroscedasticity in regression analysis consists of the existence of different variability (e.g. variance) across some of the random variables sub-populations. One example is when the variance scales with the magnitude of a random variable, which requires variable transformations (e.g. log-values or square roots) in order to avoid the classical linear regression model's assumption of no heteroscedasticity. The presence of heteroscedastic random variables in the model does not bias the OLS coefficient estimates. However, it can bias the variance (and therefore the standard errors) of the coefficients. More sophisticated regression analyses with heteroscedasticity can be modelled with autoregressive conditional heteroscedasticity (ARCH) models. It is important to remark that we do not employ a GARCH model because the models in this study are not trained only on historical volumes; instead, we bring in different predictors besides the volumes (e.g. returns, which have different variability as they come from a different data population).

The intraday return and the overnight return, the latter of which will be introduced in the next model, allow for both positive and negative results. Given the literature findings on the volume-price relation asymmetry, we define each of these metrics in two ways. The first method regards the intraday and overnight returns as being symmetric in terms of magnitude, and therefore each of them corresponds to a single feature taking the absolute value of the log-ratios (the features are generally called 'abs', e.g. 'abs' intraday return or 'abs' overnight return). The other method is based on the fact that the magnitude of price returns is asymmetric, depending on the sign of the price return; instead of having a single feature, this method generates two features based on the price movement direction (e.g. positive or negative); this allows the positive and negative returns to be represented by two features, which will potentially result in having different coefficients when being fit into the regression model. These two features are called 'absPos', representing the absolute value of positive returns only, and 'absNeg', standing for the absolute value of negative returns only (e.g. 'absPos' intraday return, 'absNeg' intraday return etc.). These indicators split the log-ratio returns at zero, into positive absolute values and negative absolute values.

We extend the volume model, which was constructed based on the optimal values of the *lag* and *smooth* orders, with two more price-related feature sets: symmetric intraday price features and asymmetric intraday price features. The former includes the 'abs' intraday return, and the intraday range log-ratios. The latter model includes the 'absPos' intraday return, 'absNeg' intraday' return, and the intraday range log-ratios.

Once these full models consisting of volume and symmetric/asymmetric intraday prices have been linearly fit using OLS regression, we perform feature selection while enforcing the volume-related features (i.e. 'volLag' and 'volWin') to be kept in the model, along with the constant term.

Table 3.3 outlines the component features for each model in the current state, called 'state A'. The question marks in the table represent a feature that might be selected or not after the feature selection process. The tick means that the feature is definitely present in the model and a cross notes its absence. The model with the lowest CV MSE in this state is called 'the state A model'.

Table 3.3
The features of state A models.

	Volume and symmetric intraday prices model	Volume and asymmetric intraday prices model
intercept	✓	✓
$volLag_p$	✓	✓
$volWin_q$	✓	✓
intraday range	?	?
'abs' intraday return	?	×
'absPos' intraday return	×	?
'absNeg' intraday return	×	?

3.5.3 Contribution of Overnight Return

We extend the state A full models (i.e. the entire set of features) by adding the overnight return (i.e. the difference between today's open and yesterday's close). This variable is employed in order to test whether information that is more recent improves the volume prediction model significantly. The overnight return indirectly measures trading volume and can be expected to be a leading indicator of the day's volume because the opening price is associated with the opening auction. The overnight return incorporates the information accumulated during the non-trading period, when investors rebalance their portfolios. The opening auction plays a major role in the daily price discovery process and reflects the private and public information flowing while the market was closed. This theory has been argued in the price formation models formulated in established literature, as surveyed by Gerety and Mulherin (1994).

The overnight return uses the opening price of t_0. This feature incorporates the after-hours trading (i.e. market-moving events occurring overnight, between yesterday's close and today's open, such as earnings reports, pre-earnings announcements or M&A activity, which drive prices) and does not reflect any of the trading activity for t_0; therefore, there is no look-ahead bias. The reason for adding this feature is to investigate whether more recent information proves to be beneficial to the prediction of the following day's trading volume. Due to licencing constraints, we did not have access to the opening auction volume, and the overnight return is implemented as a proxy for the opening auction volume (and hence for the more recent data).

As with the state A models, the models in this state, which is called 'state B', are defined in two ways, either with asymmetric price ratios or with symmetric price ranges. Each of the intraday return and the overnight return has its own asymmetry, resulting in four models in state B, whose features are listed in Table 3.4. Feature selection is performed on the price features of these models (e.g. intraday range, intraday return abs/absPos/absNeg, and overnight return abs/absPos/absNeg), while keeping fixed the features of the volume model. The model with the minimum CV average MSE is called 'the state B model'.

Table 3.4
The features of state B models.

	Volume, symmetric intraday prices, and symmetric overnight prices model	Volume, symmetric intraday prices, and asymmetric overnight prices model	Volume, asymmetric intraday prices, and symmetric overnight prices model	Volume, asymmetric intraday prices, and asymmetric overnight prices model
intercept	✓	✓	✓	✓
$volLag_p$	✓	✓	✓	✓
$volWin_q$	✓	✓	✓	✓
intraday range	?	?	?	?
'abs' intraday return	?	?	×	×
'absPos' intraday return	×	×	?	?
'absNeg' intraday return	×	×	?	?
'abs' overnight return	?	×	?	×
'absPos' overnight return	×	?	×	?
'absNeg' overnight return	×	?	×	?

At this stage, we performed an intermediate analysis in order to decide whether to correct the overnight return (i.e. dividing by the number of intervening nights) or not. Therefore, we define two variants for each of the four state B models outlined in Table 3.4. The first method provides a correction factor for the higher coefficient magnitude associated with a larger number of intervening nights, dividing by the number of intervening nights; this is the corrected overnight model type. The alternative model considers that the corrected model might have artificially small overnight returns and does not divide by the number of intervening nights (although this could potentially result in artificially high returns). Table 3.5 outlines the frequency table for the intervening nights up to 9 nights; the most common successive trading days have one intervening night (e.g. two trading days during the same week) or three intervening nights (e.g. a Friday and a Monday, which are separated by two additional nights because of the weekend).

Table 3.5
Frequency table of intervening nights.

Class (Number of intervening nights)	Count	Percentage
1	5590714	77.71
2	72852	1.01
3	1369188	19.03
4	84731	1.18
5	52805	0.73
6	13354	0.19
7	3867	0.05
8	1478	0.02
9	655	0.01

Since the data points with 1, 2, 3 and 4 intervening nights account for 98.93% of this study's data set, we used them as four different classes of overnight returns in order to assess whether the overnight returns are significantly different for each pair of these overnight return classes. For each stock, we performed a randomisation test with 1,000 reshufflings for each of these pairs: class 1 and class 2, class 1 and class 3, class 1 and class 4, class 2 and class 3, class 2 and class 4, and finally class 3 and class 4. Using the empirical p-values, we evaluated which pairs of classes are significantly different (i.e. where the null hypothesis is rejected) and then aggregated the results across the entire stock universe by computing the percentage of tests where the null hypothesis is rejected. The results outlined in Table 3.6 show that each pair is predominantly found as coming from the same population, as there is no pair where the two overnight return classes are considered significantly different in more than 50% of the cases. Based on these pairwise randomisation tests, we report that the overnight return magnitude is not different based on the number of intervening nights. Therefore, the number of intervening nights is not a salient factor in determining the trading volume and we did not include it in the list of predictors for the state B models.

Table 3.6
Statistical significance of the overnight returns based on the number of intervening nights.

Class (Number of intervening nights)	Null hypothesis rejected
1-night vs. 2-night overnight returns	36.36%
1-night vs. 3-night overnight returns	46.87%
1-night vs. 4-night overnight returns	34.03%
2-night vs. 3-night overnight returns	25.13%
2-night vs. 4-night overnight returns	5.86%
3-night vs. 4-night overnight returns	20.66%

In order to compare the performance of the two model types in conjunction with trading volume, we fit four models (all combinations of symmetric and asymmetric intraday returns/overnight returns) for each of the corrected and uncorrected overnight returns, for each stock in our universe. Then we compare the corrected models with their corresponding uncorrected models and we choose the model with the lowest 10-fold CV average MSE. There are 2,353 stocks, resulting in 9,412 fit models, which are to be compared in terms of MSE between the corrected overnight return variant and the uncorrected one. The overnight return models with corrections applied to them for the intervening nights perform better than their uncorrected alternatives in 77.54% of the cases.

Based on these results, we correct the overnight return in this study. Therefore, it is divided by the number of intervening nights in order to avoid having higher magnitude coefficients as a correction factor for the trading days following one or more non-trading days (e.g. weekends or bank holidays). Equation (3.10) shows the log-ratio representation of the overnight return, where n is the number of intervening nights:

$$\text{overnight return log ratio: } p_{\text{overnightRtn}} = \frac{1}{n}\log\frac{p_{t_0}^{\text{open}}}{p_{t_{-1}}^{\text{close}}}. \tag{3.10}$$

Figure 3.5 is an adjusted response plot for Barclays (with 3,863 observations between 2000 and 2015), which illustrates the fit response as a function of overnight return log-ratio 'absPos' and 'absNeg', respectively, while adjusting for all the other terms; this consists in averaging out the other predictor variables (i.e. the volume time-lagged features and the other intraday prices features, except the overnight return 'absPos' or 'absNeg'), by taking the mean of the fit values over the data used in the linear fit. The residuals are added to each observation's adjusted fit value in order to compute the adjusted data. The visual representation of these regression models suggests a different magnitude related to price asymmetry.

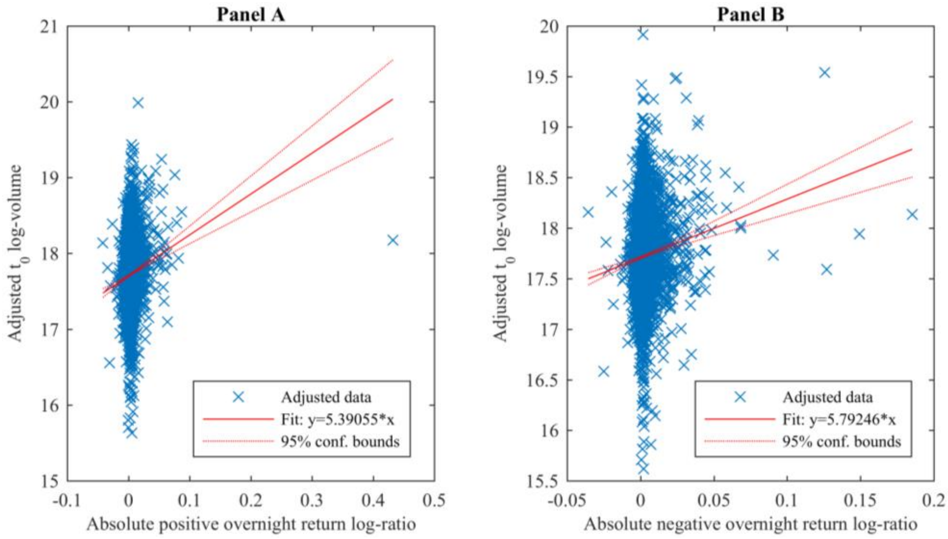

Figure 3.5. Added variable plot for the asymmetric overnight return log-ratios (positive in Panel A and negative in Panel B) in the state B model, i.e. volume, intraday prices and asymmetric overnight prices, after adjusting for all the other terms in the model.

Two added variable plots are illustrated in Figure 3.6, showing the volume model in Panel A and the state B model (i.e. consisting of volume, intraday prices and overnight prices) in Panel B. The figure illustrates the incremental effect on the response of the predictor variables for the entire data for Barclays (3,863 observations); the fit line slope represents the coefficient of the linear combination of the predictor variables, which is projected on the best-fitting direction. The added variable plot allows for model visualisation, which would not be possible otherwise due to the high-dimensional data.

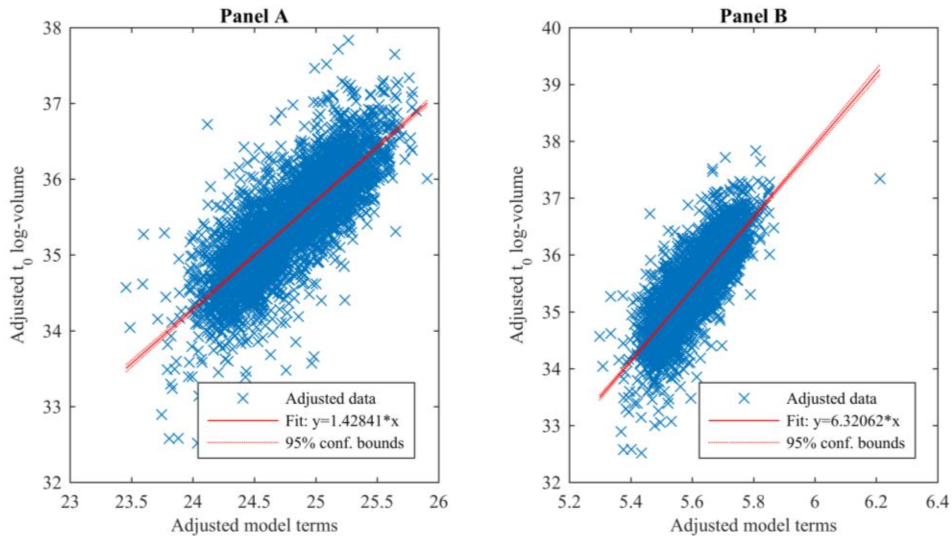

Figure 3.6. Added variable plots for the volume model in Panel A, and the state B model (i.e. the volume, intraday prices and overnight prices model) in Panel B.

3.5.4 Asymmetry Randomisation Analysis

Asymmetry can be either strong (with absPos or absNeg features that are explicitly picked by the stepwise regression), or weak. The weak asymmetry is determined through a procedure based on randomisation tests, which are described below.

First, a randomisation test is performed for the intraday return absPos and intraday return absNeg, excluding the zero-valued observations in order to evaluate the null hypothesis that the data in vectors absPos and absNeg come from independent random samples from distributions having equal means and variances. If the null hypothesis is rejected at the 5% significance level, then the vectors absPos and absNeg are significantly different, coming from populations with unequal means. Similarly, a second randomisation test is conducted on the same samples (i.e. intraday return absPos and intraday return absNeg) for the models where both asymmetric features are picked by the feature selection process, but this time they are multiplied by their regression coefficients. If the randomisation test does not reject the null hypothesis, then it means that the vectors absPos and absNeg would have different magnitudes, which are eventually corrected by the regression coefficients in order to ultimately get a symmetric representation of intraday return.

For the intraday return asymmetry, we consider each stock's state A model (among symmetric and asymmetric intraday price features) where an intraday return feature is present in any way (e.g. abs, absPos, or absNeg); the stocks whose state A model has no intraday return feature kept in the model after feature selection are disregarded. Similarly, the overnight return asymmetry analysis only considers the stocks whose state B model has at least an overnight return feature in the reduced model (e.g. abs, absPos or absNeg).

There are two scenarios where we investigate whether the lack of significant difference provided by the randomisation test is consistent with the regression symmetry.

In the first scenario, the model with the best error is asymmetric. If the presence of absPos and absNeg is mutually exclusive (i.e. if absPos is present and absNeg is absent; or absNeg is present and absPos is absent) or both absPos and absNeg are present and their randomisation test is negative (meaning that the vectors absPos and absNeg come from populations with equal means, but their regression coefficients provide empirical evidence that their impact on volume is different depending on their sign), then there is return asymmetry. However, if both absPos and absNeg are present and their randomisation test is positive (i.e. the two vectors are significantly different), we need to check whether the regression coefficients might reverse their asymmetry, causing them to behave symmetrically. In this case, if the randomisation test of the two vectors multiplied by their regression coefficients is positive, then it means that the two vectors are still significantly different even after accounting for their coefficients; otherwise, if the randomisation test is negative, it means that the coefficients act as a correcting factor for the apparent asymmetry, transforming the two vectors into a symmetric vector.

In the second scenario, the model with the lowest error is symmetric. If the randomisation test of the absPos and absNeg vectors is positive, then these vectors are significantly different and the model is actually asymmetric. Otherwise, it means that they come from populations with equal means and the model is indeed symmetric.

3.5.5 Temporal Context: Day-of-the-Week Effects

The starting point of this stage consists of the best model (i.e. having the lowest CV MSE) among the state A model and the state B model. The resulting model is called 'the historical dynamic model'. In order to investigate the day-of-the-week effect, we define a couple of models: one that extends the historical dynamic model with additional dummy variables for each working day (i.e. Monday-Friday) and an elementary one with features for each of the five workdays only (without any volume or price features), due to the broad implementation of this model in the behavioural finance literature on the day-of-the-week effect and weekend effect.

In order to be able to assess which day-of-the-week features are the most salient, either in improving the historical dynamic model or simply in the raw day-of-the-week model, we performed feature selection on the two models. The intercept and the other features of the historical dynamic model are kept fixed in the model, while we attempt to reduce only the day-of-the-week feature set. Table 3.7 summarises the day-of-the-week models and their potential features, along with the historical dynamic model, which acts as a benchmark.

In theory, if a model has a categorical variable with k possible values, one should assign $k-1$ dummy variables if the model has an intercept and k dummy variables if there is no intercept; although such a model would result in identifiability issues due to assigning dummy variables to all of the five days (and not to only four of them), we use the five dummy variables in order to define the full model that is to be reduced using feature selection. This approach is necessary for the interpretability of each day-of-the-week. Feature selection is applied to this model, having fixed the volume and price features (i.e. the features of the historical dynamic model), in order to find the most significant days of the week.

Table 3.7
The features of the day-of-week models.

	Historical dynamic model	Historical dynamic and day-of-week model	Raw day-of-week model
intercept	✓	✓	✓
volume features volLag$_p$ volWin$_q$	✓	✓	×
price features intraday range intraday return overnight return	?	?	×
Monday	×	?	?
Tuesday	×	?	?
Wednesday	×	?	?
Thursday	×	?	?
Friday	×	?	?

3.6 Results

This section first introduces the results on the volume-price relation and then switches to the day-of-the-week findings. We start by investigating the effects for each stock and then provide a summary of the model performance across the entire stock universe. 10-fold cross-validation is performed on all of the models in states A and B and the day-of-the-week models in order to get the average MSE. We provide goodness-of-fit illustrative examples while introducing each driver of trading volume (i.e. volatility, overnight return, and day-of-the-week features).

3.6.1 Contribution of Volatility and Asymmetry

Based on the aggregated results for the entire data set in Table 3.8, we report that the volatility features (i.e. the previous day's price changes, i.e. intraday range and intraday return) generally improve the autoregressive volume model, where 60% of the models are trained on asymmetric intraday returns. These results are computed by selecting the state A model for each stock in our 2,353 stock universe and examining the feature presence of each model. A representative model is exhibited in Figure 3.7 for Fenerbahce Futbol AS, which also contains a zoomed plot of the time series spanning the last 6 months for better visualisation. The improvement over the volume model for this stock is 5%. Figure 3.8 depicts the different distributions of the vectors 'intraday return absPos' and 'intraday return absNeg' for PCC Rokita SA.

Table 3.8
Volatility findings.

Hypothesis	Observation Percentage		
	Entire Data Set 2000-2015	Structural Break Subsets 2000-2007	2008-2015
Volatility improves the volume model	86.95%	82.08%	84.62%
Asymmetry of volume – intraday return relation (State A models)	60.16%	59.64%	62.52%
More recent price information improves the model	87.21%	81.57%	86.87%
Asymmetry of volume – overnight return relation (State B models)	54.09%	60.27%	47.11%
Historical dynamic model: intraday return asymmetry	56.53%	58.61%	57.81%
Historical dynamic model: overnight return asymmetry	54.07%	60.26%	47.11%
Historical dynamic model: total asymmetry	73.59%	75.70%	70.02%

There are no significant structural changes around the financial crisis of 2007-08. However, the overnight return asymmetry is salient in the pre-crisis subset, but becomes rather neutral in the post-crisis data set, with a performance that is similar to the symmetric overnight return.

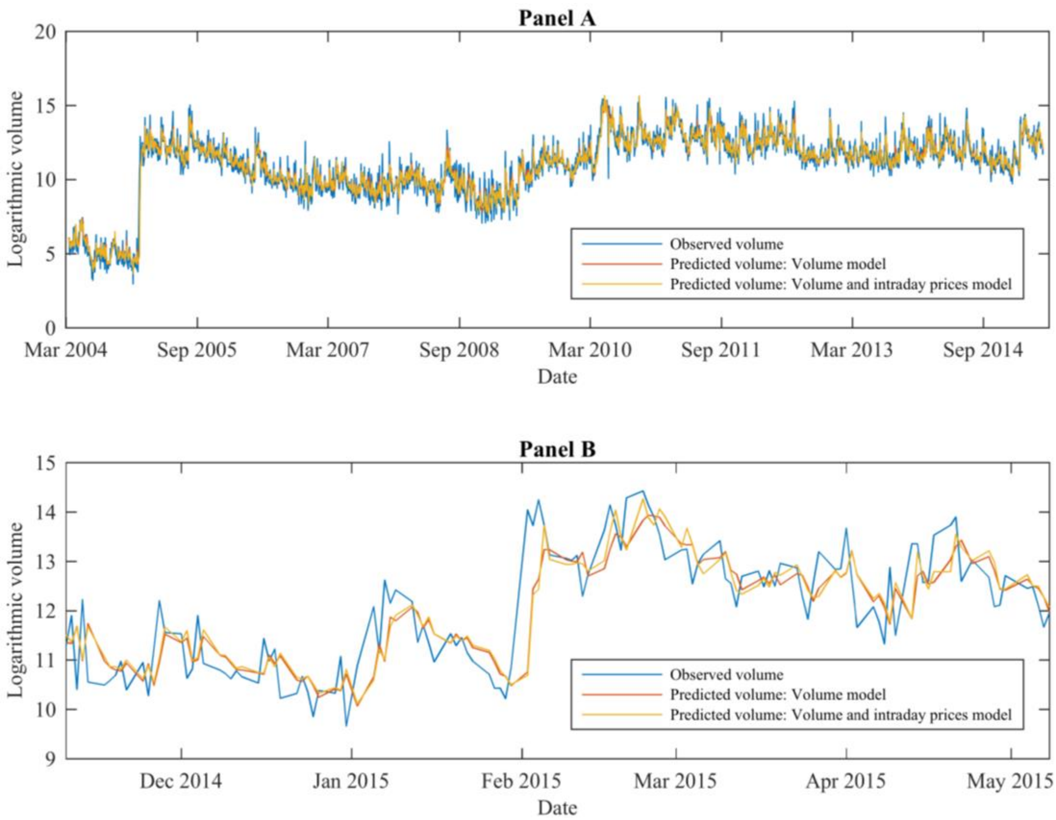

Figure 3.7. Improvement of intraday prices (intraday return and intraday range) over volume by 4.72%. Both panels illustrate the improvement for Fenerbahce Futbol AS (FENER.IS). Panel A shows the entire time series between 12th March 2004 and 8th May 2015, while Panel B provides a zoomed time series for the last 6 months of data.

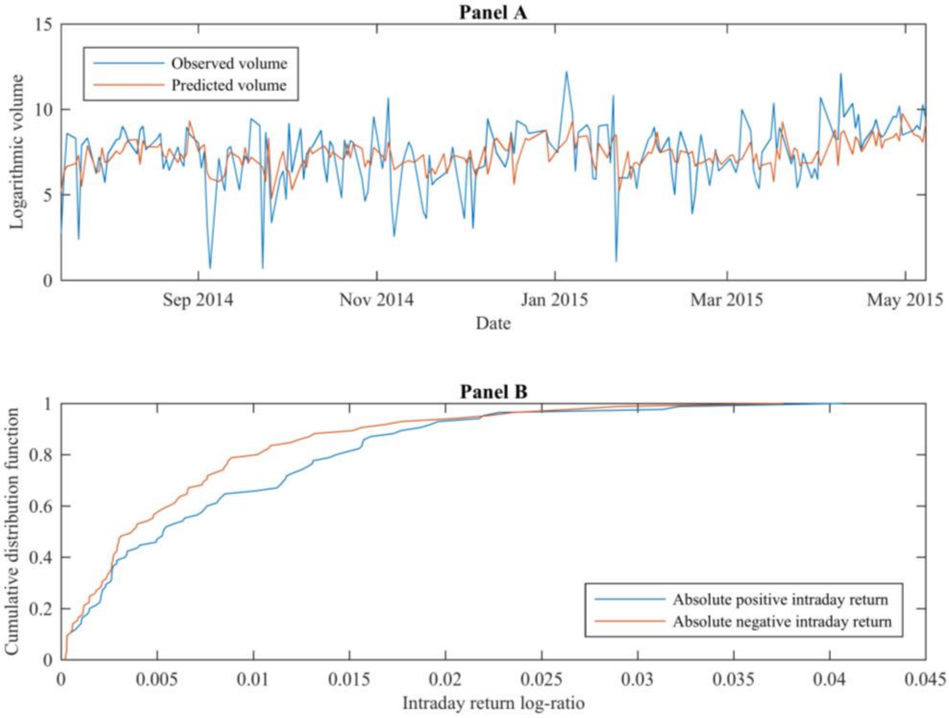

Figure 3.8. Intraday return asymmetric distribution. Panel A illustrates the observed volume against the predicted volume using the asymmetric intraday return model for PCC Rokita SA (PCR.WA) from 16th July 2014 to 8th May 2015 (201 trading days). Panel B illustrates the cumulative distribution breakdown asymmetric intraday return for this volume prediction model.

By using the state B model for each stock, similar results are reported for the contribution of more recent information, in the form of overnight returns, which act as a proxy indicator for the opening auction volume in this study. The opening auction volume represents the most recent piece of information that could be publicly available at the opening of trading, since the continuous trading phase begins based on the conclusion of the opening auction. The overnight return improves the volume model predominantly, with a slightly lower asymmetry for the overnight return of 54%. Figure 3.9 shows the recent data improvement (16%) over the state A model for H & M Hennes & Mauritz AB; Panel B is a magnified view of the last 6 months of the same time series for easier visualisation. The distribution of the asymmetric overnight return vectors is shown in Figure 3.10, along with the predicted and observed volume time series for Aeffe SpA.

Considering that the intraday return asymmetry and overnight return asymmetry provide better performance in more than 50% of the stocks, we argue that the volume-price relation should be modelled with asymmetry.

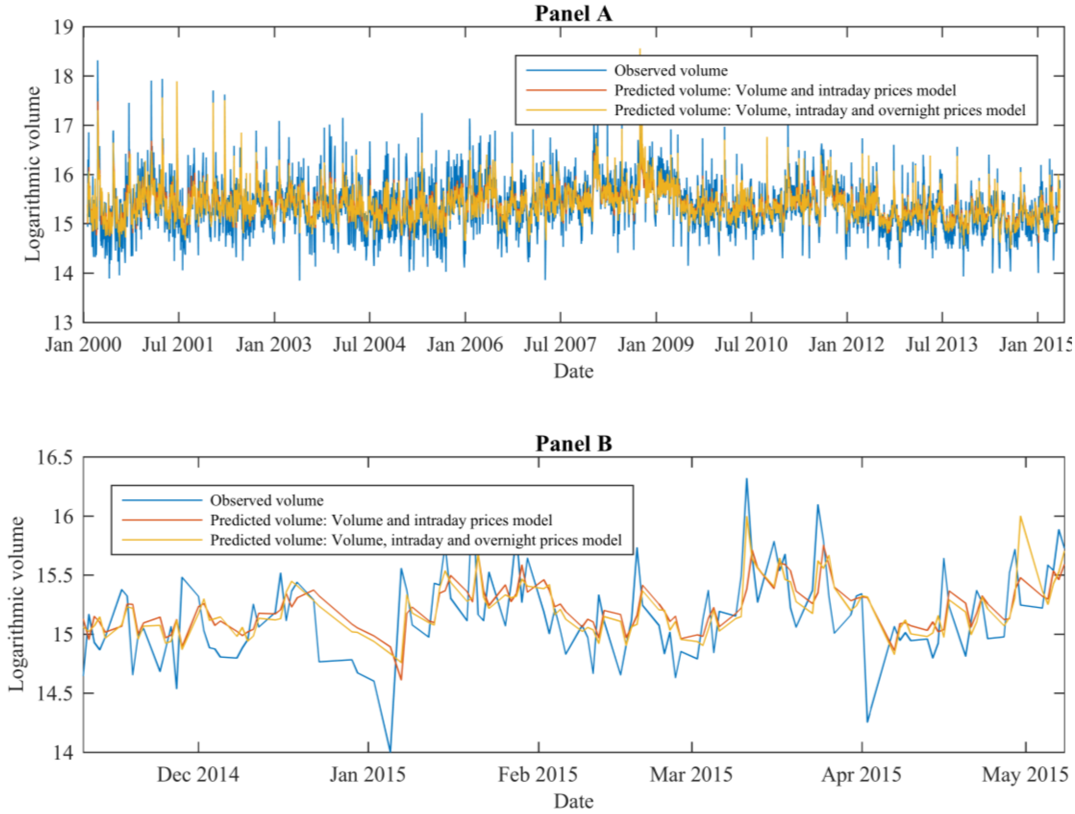

Figure 3.9. Overnight return improvement. Both panels show the overnight prices improvement over volume and intraday prices (16.26%) for H & M Hennes & Mauritz AB (HMb.ST). Plot A contains the entire time series between 25th January 2000 and 8th May 2015 (3,837 observations) and Plot B provides a magnified view of the last 6 months.

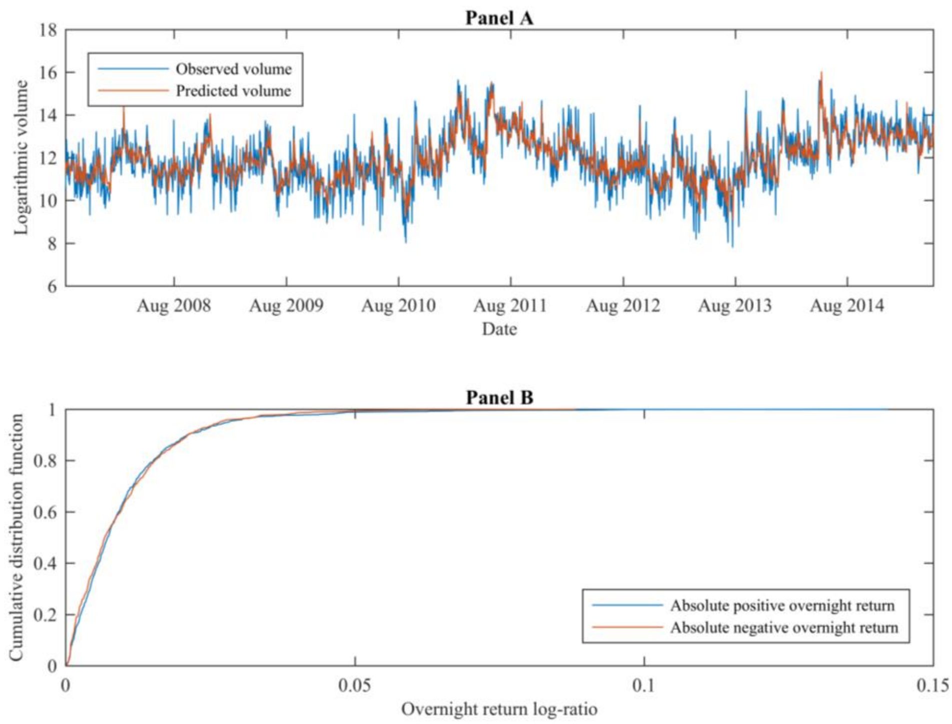

Figure 3.10. Overnight return asymmetric distribution. Panel A shows the observed volume against the predicted volume using the asymmetric overnight return model for Aeffe SpA (AEF.MI) from 14th August 2007 to 8th May 2015 (1,949 trading days). Panel B illustrates the cumulative distribution breakdown of the asymmetric overnight for this volume prediction model.

Further analysis on the asymmetry is computed on the historical dynamic model (i.e. the model with the lowest CV MSE throughout states A and B, fit using volume and price features). Here, we use the historical dynamic model across both types of asymmetry. The total asymmetry for each stock is the logical disjunction (logical or) between the two types of asymmetry for that stock. When examining the historical dynamic model, we report moderate intraday return asymmetry (56%) and overnight return asymmetry (54%). However, the majority of models (74%) exhibit some type of asymmetry, be it in the form of intraday returns or overnight returns.

We explored the transition of intraday price features across states A and B, and investigated whether having the overnight return feature selected in state B could potentially cancel any intraday price features previously selected in state A. Table 3.9 shows the occurrence frequency of each intraday price feature across the two states for the entire data set, whereas Table 3.10 includes the feature presence for the two structural break data subsets (i.e. the pre-crisis and post-crisis data). The results do not exhibit any significant frequency changes for any volatility feature, apart from a general increase in the number of models with the intraday range predictor selected, which is also consistent for our structural break.

Table 3.9
Volatility feature presence across states A and B for the entire data set.

Volatility Feature	State A Count	State B Count
Intraday range	1,185	1,453
Intraday return abs	810	894
Intraday return absPos	649	646
Intraday return absNeg	840	734

Table 3.10
Volatility feature presence across states A and B for the sectional break data subsets.

Volatility Feature	Pre-Crisis Data (2000-2007)		Post-Crisis Data (2008-2015)	
	State A Count	State B Count	State A Count	State B Count
Intraday range	1,012	1,110	1,157	1,322
Intraday return abs	648	662	734	799
Intraday return absPos	488	485	549	558
Intraday return absNeg	543	493	802	726

3.6.2 Contribution of Day-of-the-Week Effects

The study further investigates the temporal context of the volume time series, analysing the day-of-the-week effect. We compared the historical dynamic model (resulting from states A and B) to the day-of-week model that is traditionally employed in the calendar effect literature (i.e. the raw day-of-week model). Based on the results outlined in Table 3.11, we find that the historical dynamic model fit with volume and price features clearly dominates the traditional raw day-of-the-week model in terms of performance (in almost 100% of the analysed stocks). Moreover, we augmented the historical dynamic model with day-of-week features, which, after performing feature selection, improved the historical dynamic model with at least one day-of-week feature in approximately 91% of the models – an illustrative day-of-week improvement (7%) is shown in Figure 3.11 for E.ON SE.

Table 3.11
Day-of-the-week findings.

Hypothesis	Observations Percentage		
	Entire Data Set 2000-2015	Structural Break Subsets	
		2000-2007	2008-2015
Historical dynamic model is significantly better than the raw day-of-week model	99.87%	99.44%	99.79%
The day-of-week features improve the historical dynamic model	90.57%	88.29%	88.61%

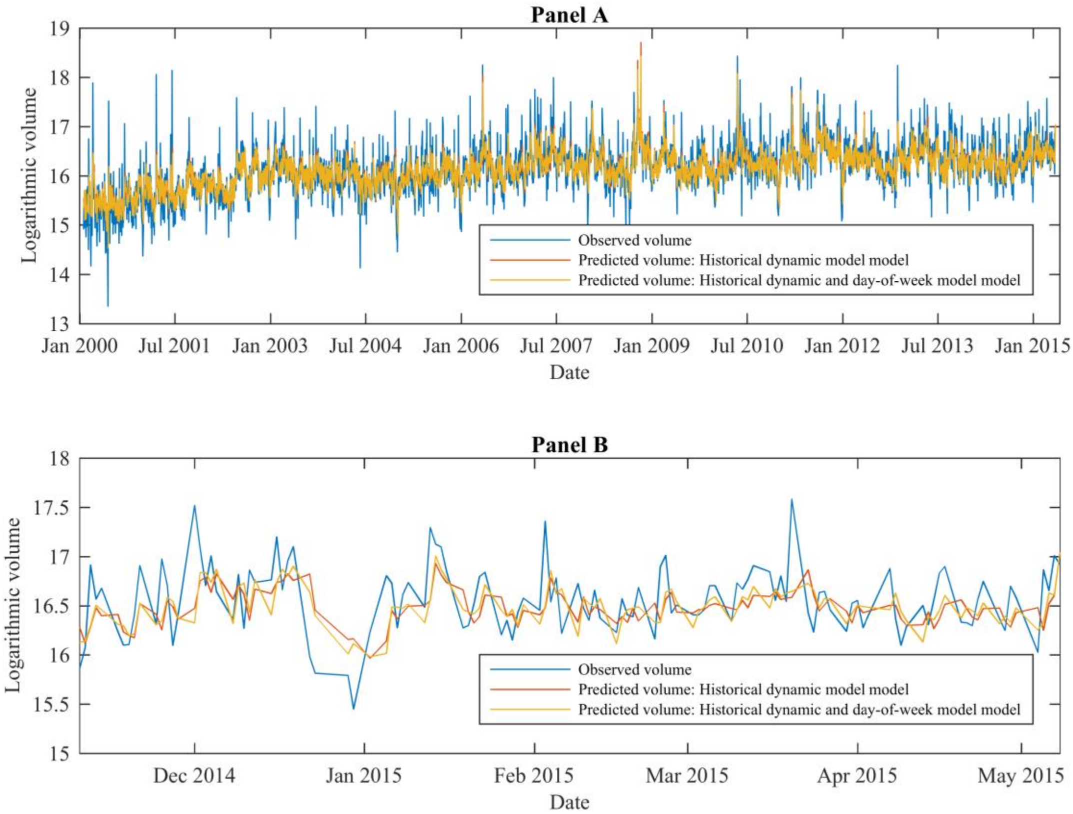

Figure 3.11. Day-of-week improvement over the historical dynamic model (7.03%) for E.ON SE (EONGn.DE). Panel A shows the complete time series between 24th January 2000 and 8th May 2015 (3,880 observations), whereas Panel B provides a zoomed view of the most recent 6 months.

Table 3.12 outlines the day-of-the-week feature selection process for the two models, namely the raw day-of-week model (i.e. a model consisting only of five dummy variables for each workday), and the historical dynamic and day-of-week model (i.e. the model extending the historical dynamic model, which consists of endogenous variables, namely volume features and price features, by adding a dummy variable for each workday). There is a breakdown of the coefficient sign distribution (i.e. positive and negative) across the stock universe for each workday, which is included under the presence proportion of each day-of-the-week (resulting from the stepwise regression). Monday is a notable day-of-the-week feature, which is consistently picked in the raw day-of-week model and in the historical dynamic and day-of-week model, where it is generally negatively correlated with the trading volume. The Monday coefficient is consistently negative despite the overnight return correction (i.e. dividing the overnight return by the number of intervening nights), which suggests that generally there is less trading activity on Mondays. This confirms the potential existence of a weekend effect on trading volumes. We also report predominantly negative coefficients for Fridays, although the Friday day-of-the-week feature is significantly picked only in the historical day-of-week model. Although the weekend effect is documented as having higher than usual Friday returns and, hence higher volumes (according to the literature on the volume-price relation), we observe a mostly negative Friday coefficient, which is associated with lower volumes.

Table 3.12
Day-of-the-week feature selection – presence percentage for each day of the week along with the distribution of coefficient signs.

	Monday	Tuesday	Wednesday	Thursday	Friday
Panel A: Raw day-of-the-week model					
Occurrence	64.39%	17.81%	11.77%	12.92%	18.78%
Positive coefficient	4.49%	41.77%	76.53%	74.34%	24.89%
Negative coefficient	95.51%	58.23%	23.47%	25.66%	75.11%
Panel B: Historical dynamic and day-of-the-week model					
Occurrence	75.69%	30.41%	21.16%	21.87%	45.33%
Positive coefficient	9.11%	90.28%	67.41%	35.19%	14.29%
Negative coefficient	90.89%	9.72%	32.59%	64.81%	85.71%

The Monday and Friday feature presence and coefficient sign distribution for the structural break data subsets are outlined in Table 3.13. We observe a constant Monday effect for both day-of-the-week models throughout the pre-crisis and post-crisis periods. The Friday effect is more volatile though and its coefficient becomes positive for more than 50% of occurrences in the raw day-of-the-week model trained on the post-crisis data.

Table 3.13
Day-of-the-week feature selection for the structural break subsets.

	Pre-Crisis Data (2000-2007)		Post-Crisis Data (2008-2015)	
	Monday	Friday	Monday	Friday
Panel A: Raw day-of-the-week model				
Occurrence	58.56%	21.81%	63.54%	14.02%
Positive coefficient	4.22%	18.29%	4.62%	53.03%
Negative coefficient	95.78%	81.71%	95.38%	46.97%
Panel B: Historical dynamic and day-of-the-week model				
Occurrence	70.73%	44.48%	72.09%	35.11%
Positive coefficient	8.49%	11.83%	9.58%	21.04%
Negative coefficient	91.51%	88.17%	90.42%	78.96%

The historical dynamic and day-of-week model was generally the most accurate model in this analysis. Figure 3.12 and Figure 3.13 illustrate the predicted volume time series along with the observed volumes for two stocks (i.e. Royal Dutch Shell and Siemens); a zoomed plot for the most recent 6 months of modelling data accompanies these figures for better visualisation.

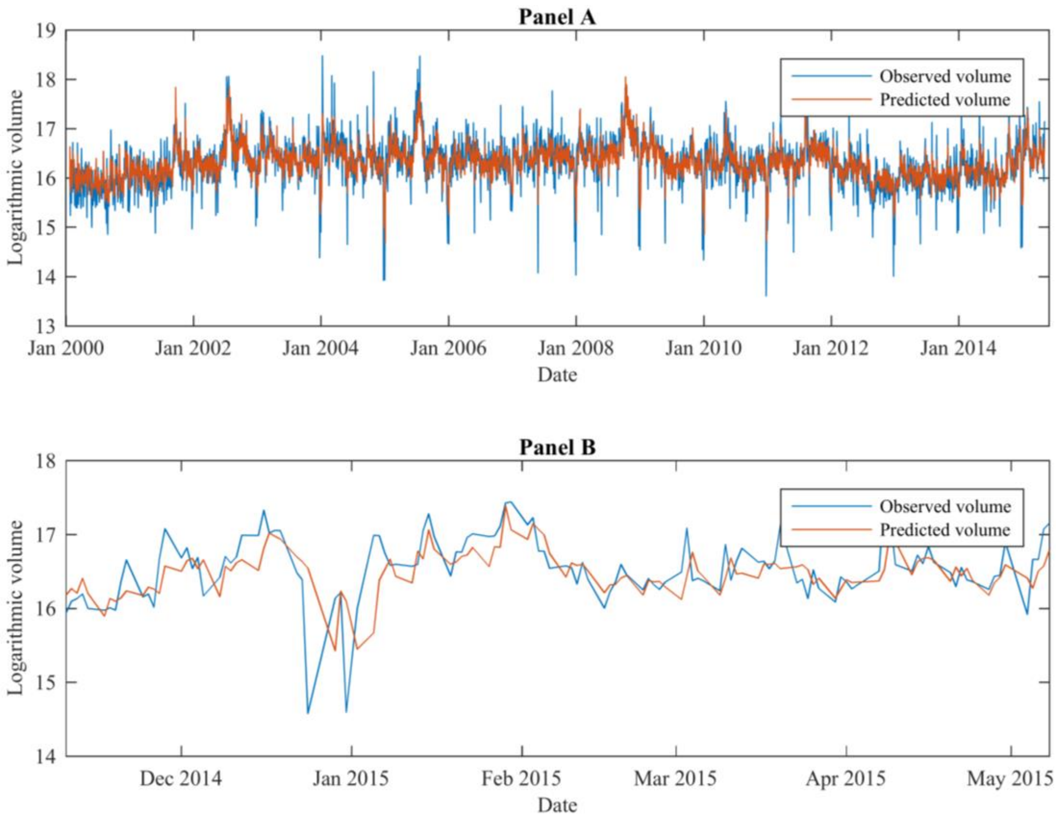

Figure 3.12. Observed volume and predicted volume using the historical dynamic and day-of-week model for Royal Dutch Shell PLC (RDSa.AS) for 3,909 daily observations (24th January 2000 – 8th May 2015). Panel B is a zoomed in plot of the most recent 6 months of data.

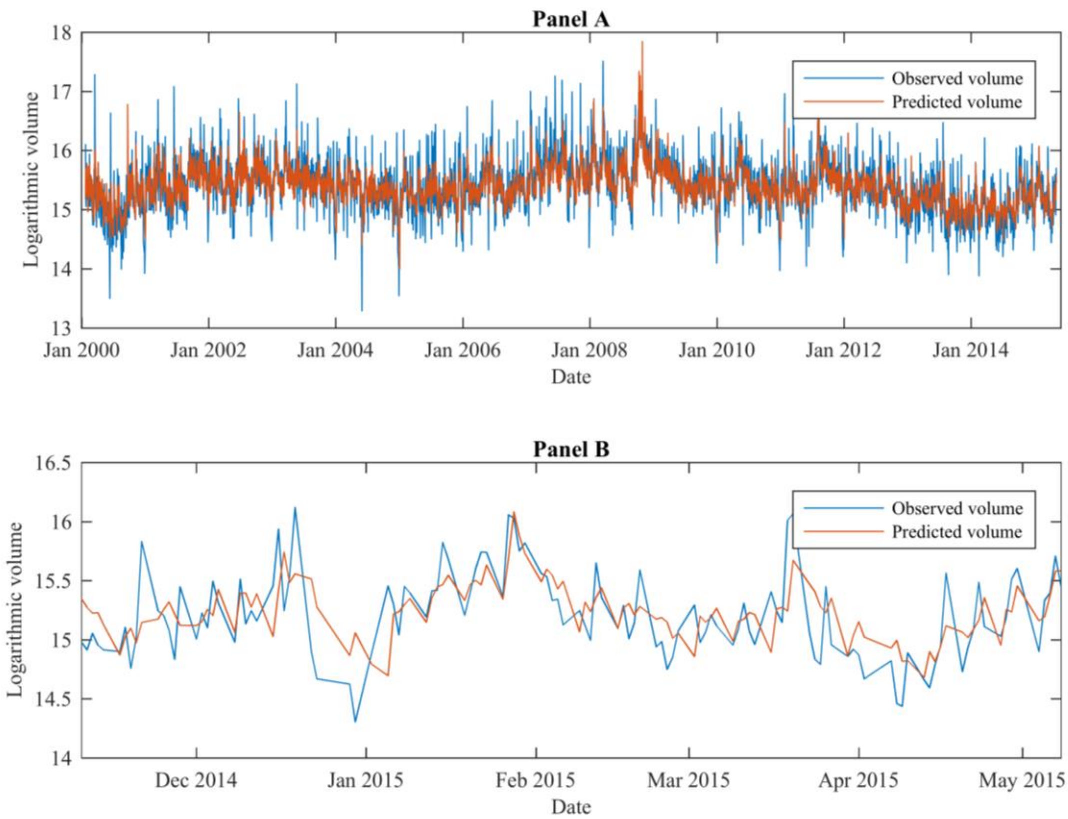

Figure 3.13. Observed volume and predicted volume using the historical dynamic and day-of-week model for Siemens AG (SIEGn.DE). Panel A illustrates the entire time period being studied, between 24th January 2000 and 8th May 2015 (3,880 trading days), whereas Panel B shows a magnified view of the most recent 6 months.

Figure 3.14 depicts the error percentage change from the historical dynamic model to the raw day-of-the-week model, showing predominantly positive observations, meaning that the average MSE increases, making the model worse. We conclude that the traditional raw day-of-the-week model is inferior to the historical dynamic model. Similarly, Figure 3.15 illustrates the error percentage change between the historical dynamic model and the augmented model that adds day-of-the-week features on top of the historical dynamic model, with a dominantly negative distribution suggesting that the historical dynamic and day-of-week model lowers the average MSE and provides a better fit.

At this phase, the day-of-the-week analysis provides a discussion point, which leads to a further study on special events (e.g. cross-market holidays and stock index futures expiries), which could potentially impact on the Friday and Monday volumes.

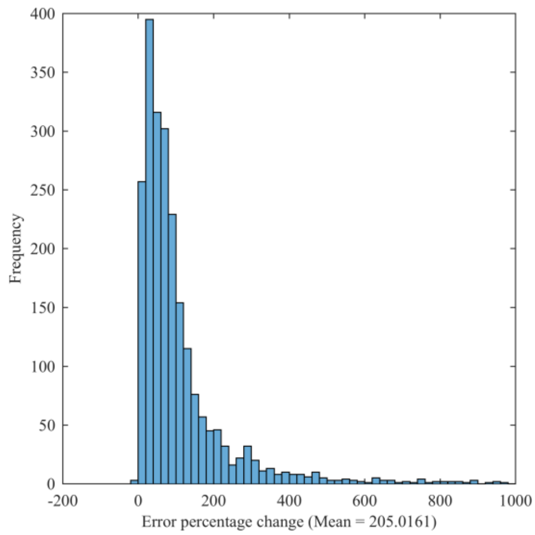

Figure 3.14. Histogram of error percentage change from historical dynamic model to raw day-of-week model for the entire stock universe (2,353 stocks).

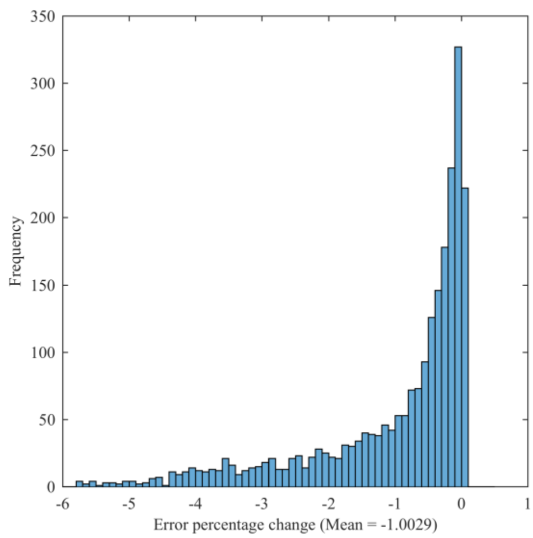

Figure 3.15. Histogram of error percentage change from historical dynamic model to historical dynamic and day-of-week model for the 2,353 stocks studied.

3.7 Discussion

This study provides a broad exploration of endogenous and exogenous factors driving trading volume, while investigating a number of relevant aspects, such as the volume-price relation asymmetry and the existence of structural breaks. The effect size is not part of the scope of this study mainly because we fit the models independently for each stock. The rationale is that there are strong stock-specific variability and magnitude levels that could not allow for a unified model across stocks. Instead, the aim is to identify the variables that help predict the trading volume of the following day. To the best of our knowledge, we provide the largest pan-European stock universe in any academic study. The extended data universe provides robust validation of our results.

We investigate potential structural breaks and non-stationarity around the financial crisis of 2007-08 as a method of validating the results, which assume strong homogeneity. We split the data set into two folds: the pre-crisis data set (2000-2007) and the post-crisis data set (2008-2015).

The study considers single stock modelling and eventually aggregates the results across a data universe of 2,353 stocks. We provide empirical evidence of a significant improvement over the autoregressive volume model using volatility features (i.e. intraday range and intraday return for the previous trading day), more recent price information (i.e. overnight return as a proxy for the opening auction volume), and day-of-the-week features. The only constant day-of-the-week exerting a dominant influence over trading volumes is Monday, which improves the historical dynamic model in over 75% of the sample stocks. The coefficients are predominantly negative for Mondays, even though we divide the overnight return by the number of intervening nights; Monday's coefficient is not a corrective factor and it suggests that there is less activity on Monday. Friday is the second most selected day-of-the-week feature, but it improves the model in only 45% of the times; its regression coefficient is mostly negative as well, although it is positive in more than 50% of the observed models for the raw day-of-the-week model using the post-crisis data subset.

The empirical evidence suggests a stronger day-of-the-week effect in conjunction with the endogenous variables. More notably, there is a Monday effect and a less salient Friday effect, both days exhibiting negative returns. This confirms the weekend effect literature with regard to lower Monday returns. However, the Friday returns tend to be negatively correlated with the volume. We have not specifically addressed special events in the context of this analysis. The reasoning behind this decision consists of the insufficient number of observations of special events in each fold (i.e. futures expiries and cross-market holidays, which have very few observations anyway), given the context of our stock-specific modelling. The day-of-the-week modelling provides a discussion point, which leads to a separate investigation of the special events (i.e. cross-market holidays and stock index futures expiries) that potentially affect the Monday and Friday volumes. The next two studies of this thesis will examine the Monday effect, exploring whether this is actually a cross-market holiday effect or whether it is indeed a Monday effect, and the Friday effect, which could potentially be an effect caused by the stock index futures expiries that typically occur on Fridays.

We also examined the accuracy of the raw day-of-the-week model, which is traditionally employed in the weekend effect and day-of-the-week literature. It constantly underperformed compared to the historical dynamic model by a large factor and this evidence suggests that fitting a dummy variable model for a particular effect in complete isolation from other variables, especially endogenous, would provide questionable results.

Another interesting aspect being analysed in this study is the asymmetry of the price-volume relation. We proposed a different framework for exploring the price-volume relation asymmetry. The empirical results suggest that we first need to discriminate between two types of price-volume relations; there is an intraday return (i.e. the previous day price returns) – volume relation, which manifests asymmetry in approximately 60% of the stocks, and there is an overnight return (i.e. newer data) – volume relation, which exhibits asymmetry in 54% of the stocks. Combining these two relations, we find an overall asymmetry in approximately 70% of the analysed stocks. We report a structural break with regard to the overnight return asymmetry, which is salient for the pre-crisis data, but then it becomes rather neutral and reaches similar performance as the symmetric overnight return. Apart from the Friday effect and the overnight asymmetry, where we found notable structural breaks, the study confirms data homogeneity for all of the other aspects being examined throughout the entire sample period, i.e. 2000 - 2015.

The empirical results show that more recent information (i.e. overnight return) improves the volume prediction model and having the overnight return being used as a proxy for the opening auction volume confirms the price-volume relation. This could be further improved by performing an intraday volume prediction model to further analyse the price-volume relation based on tick data.

4. European Trading Volumes on Cross-Market Holidays

There is anecdotal evidence of reduced trading volume in equity markets when other external markets are not trading. This phenomenon can be called the 'cross-market holiday effect' and this study investigates it in detail, providing evidence for the existence of a strong cross-market holiday effect in the pan-European equity markets. The analysis provides an in-depth examination of other aspects like lagged volumes, market capitalisation or multi-step ahead modelling. The trading volumes on dates when there is at least one cross-market holiday are on average 8.5% lower than the volumes of the previous period. There are salient effects when the holiday takes place in a dominant market or when most of the European markets are shut. We test whether the lower trading activity on Monday cross-market holidays is a consequence of the weekend effect, or whether the Monday bank holidays push down the Monday trading volume. We report a significantly lower volume associated with the Monday bank holidays and we argue that the weekend effect has an insignificant impact on the Monday volumes where there is at least one regional cross-market holiday.

4.1 Introduction

This study investigates the anecdotal evidence of lower trading volume associated with one or more external markets not trading. We propose naming this phenomenon as the 'cross-market holiday effect'. We test the hypothesis that the trading volume is lower on cross-market holidays than usually. The rationale behind these hypotheses relies on the fact that markets are event-driven and are typically in a rather constant state. However, certain events occur (e.g. expiries, trading holidays, earnings announcement, news etc.) and markets are consequently transitioning to a different state. It is this event-driven nature that we would like to exploit in this volume analysis. This study considers the impact of cross-market holidays on trading volume and explores this effect among European countries.

The motivation of this study is threefold: first, the cross-market holiday effect has not been investigated sufficiently in the literature, nor has it been examined on European market data; second, the studies on the European equity markets and trading volume are very scarce, with a large majority focusing on the price returns instead; and third, planning a multi-day trade is extremely important for practitioners and this is why this study proposes a multi-step ahead prediction model. An example of a very common use case consists of traders and portfolio managers, who want to size a multi-day order allocation with the aim of minimising the market impact based on the available liquidity. For example, they could ask how the trading volume would be in a few days' time in the UK given that it will be the 1st May and the mainland Europe is not trading; they need to have the ability to quantify and forecast the volume trends in order to be able to plan multi-day trades. This practical problem has not yet been properly addressed.

To the best of our knowledge, the cross-market holiday effect has been investigated in only a couple of studies, using data sets from over ten years ago; one of the studies investigates the effect of US holidays on the European markets' returns and volumes (Casado, et al., 2013), while the other study explores the cross-market holiday effect on volumes between the USA and Canada (Cheung & Kwan, 1992). It is the first time such an analysis is performed on the European markets on a huge data set and this is the main contribution of this study. It is the first analysis to employ an accurate trading calendar for more than 20 countries (i.e. covering the USA and the vast majority of the European Union) in order to produce a unified region-wide trading calendar.

We surveyed 80 relevant papers and found that very few studies include European stocks in their analysis. There are 10 studies focusing exclusively on one European country, 16 studies with international data sets including a few European countries, and there are 3 studies that focus on and cover more European countries; the largest data set is employed in a study on the emerging Central and Eastern European financial markets (Dodd & Gakhovich, 2011), covering 14 European markets for almost 20 years. Moreover, only 7 of the surveyed papers include recent market data after 2005 and little attention has been paid to the volume dimension.

The holiday effect consists of rather rare events, but the study is conducted on a comprehensive pan-European data set with sufficient observations (1,343,636 observations). The aim of this study is to introduce a number of in-depth pan-European in-sample analyses for volume prediction in the context of special events, such as the cross-market holidays.

The study consists of two main methodological components: first, we conduct randomisation testing in order to explore the existence of the potential cross-market holiday effect and investigate whether the lower volume corresponding to cross-market holidays is in reality caused by the Monday effect or by the Monday bank holidays, since the United Kingdom is the largest market in Europe and its bank holidays fall predominantly on Mondays; we also analyse whether there is a differentiating effect magnitude across small-, mid-, and large-cap stocks; second, based on these hypotheses, we propose a number of predictive models for trading volume in order to assess the out-of-sample performance of a forecasting model based on this effect and the other relevant aspects. It is important to note the rigour associated with the (pairwise) randomisation tests to determine the outcome of the various hypotheses based on controlled rearrangements, and the novelty of the application of ridge regression on financial time series in this study.

The data analysis consisted of a series of challenges, ranging from unavailable trading calendars to high coefficient variability due to multicollinearity. We constructed from scratch a highly accurate non-trading calendar for the USA and the European markets included in this analysis, which allowed us to validate the hypotheses investigated in this study. The scope of this phenomenon is new and we provide an extensive study of its existence and effect size.

The study proceeds as follows: section 2 surveys the relevant literature on calendar effects in order to provide the fundamental knowledge on the relevant calendar effects; section 3 describes the data sample of this study, including the stock universe, the market data, and the calendar data; section 4 presents the analytical approach of this experiment; section 5 introduces an examination of the existence of the cross-market holiday effect and other potential drivers of decreased volume using randomisation tests; section 6 proposes a number of volume prediction models using ridge regression and presents the results of the cross-market holiday models and their interpretation, while section 7 provides a brief discussion of the randomisation and regression results and a conclusion of this study.

4.2 Background

We start surveying the finance literature on comovement in international markets in order to motivate the cross-market holiday study. We provide further context to the temporal exogenous variables being investigated in this volume prediction analysis with a review of the behavioural finance literature on a variety of calendar effects. Since most of the calendar effects have been previously studied in conjunction with price returns, the finance literature review concludes with a summary of the empirical findings on the volume – price relation, which will be ultimately used to infer a direct relation between the calendar effects and trading volume.

4.2.1 Comovement of Returns and Volatility in International Markets

The analysis of information flow across international markets stems from the previous stock market crashes and the way price changes have diffused throughout international markets. King and Wadhwani (1990) provide empirical evidence for a contagion effect during the crash of October 1987, when investors inferred information from the price changes in other markets, causing the world stock markets to fall uniformly. The authors argue that volatility is positively correlated with the contagion effect magnitude. Similarly, Hamao et al. (1990) investigated price volatility spillover effects in three international markets, namely Tokyo, London, and New York. The spillover effect exhibits asymmetry, with a significant spillover effect on the Japanese market and considerably weaker effects on the UK and US markets. The spillover effect asymmetry is also shared by the findings of Becker et al. (1990), who found that the open-to-close returns of US stocks from the previous trading day are highly correlated with the current day returns of Japanese stocks, whereas the Japanese market has a minor impact on the US returns, and Eun and Shim (1989), whose nine-market vector autoregression model exhibited a significant transmission of US innovations (or residual returns) to other markets, while none of the other eight markets could explain the US price movements. Connolly and Wang (2003) argue that the intraday and overnight return comovement in international equity markets cannot be explained by public information on economic fundamentals; instead, it is rather driven by contagion and trading on private information.

Domestic comovements are found across the US asset classes, i.e. the stock, bond, money and currency markets (Darbar & Deb, 2002).

4.2.2 Calendar Effects

There is a wide range of studies looking at the calendar effects on prices, with little focus on their effects on trading volume. Therefore, it is important to understand these findings and then combine them with the results of the literature looking at the volume correlation with prices, in order to infer a connection between calendar effects and trading volume. We further extend this and investigate the cross-market holiday effect since there is extremely little literature investigating this hypothesis (i.e. the literature looks at the cross-market holiday effect on trading volume in the US and Canada only), whose importance is crucial for predicting major liquidity changes.

Calendar effects are market anomalies or economic effects that are related to the calendar. They involve a seeming change in the stock markets' behaviour; their granularity varies from intraday and day-of-the-week effects to turn-of-the-year and multi-year effects. Many calendar effects have vanished or reversed since they were discovered and documented (Dimson & Marsh, 1999). These anomalies have been researched ex post, since their existence is inferred from past empirical data. Therefore, the market inefficiency theories cannot be predicted ex ante due to the data-driven nature of such theoretical studies documenting a calendar effect and the ambiguity of the economic variables inter-dependencies.

A survey of the most illustrative calendar effects is included below in order to understand the impact of these calendar anomalies on prices, and, consequently, on volume. The following review of calendar effects proves that markets have event-driven irregularities. One of the most popular calendar effects being investigated by the behavioural finance literature is the weekend effect. The main literature findings that are synthesised below prove the inconclusiveness of the research on calendar effects, where this study contributes by providing further evidence on the cross-market holiday effect in a pan-European setting.

Noise and outliers are salient features of financial data, making many of the studies on calendar effects prone to biases. Sullivan et al. (2001) argue that the significance of individual calendar effects is weaker when they are evaluated in the context of the full universe containing all the calendar effects and rules (and their inter-dependencies) than when they are assessed in isolation. Moreover, they draw attention to the potential data mining biases resulting from the common practice of using the same data set to both formulate and test hypotheses.

4.2.2.1 Monday Effect

The Monday effect (or weekend effect) consists of a lower closing price on Monday than on the previous Friday, as first reported by Cross (1973) and confirmed by other authors (French, 1980) (Gibbons & Hess, 1981) (Jaffe & Westerfield, 1985) (Pettengill, 2003). Moreover, the empirical evidence of Berument & Kiymaz (2001) confirms the lowest returns on Monday and finds a day-of-the-week effect on volatility.

4.2.2.2 Holiday Effect

The holiday (or pre-holiday) effect consists of high mean returns on the trading day before a holiday, with a mean of nine to fourteen times the average return during the remaining days of the year (Ariel, 1990). This effect is not related to any other calendar anomaly (Meneu & Pardo, 2004) and its magnitude is related to the level of economic activity and firm size (Liano & White, 1994). Fabozzi et al. (1994) reported that the trading volume of futures contracts is lower than average on the day prior to a holiday. Kim and Park (1994) reported that the holiday effect in the UK stock market is independent of the holiday effect in the US stock market. Chan et al. (1996) found that the effect of cultural holidays is stronger than the effect of state holidays. Chong et al. (2005) show that the pre-holiday effect has declined in the US and Hong Kong markets, and more significantly in the US; the period between 1991 and 1997 witnessed a reverse pre-holiday effect (with negative mean returns) and the subsequent period between 1997 and 2003 marked the elimination of the pre-holiday effect. A few authors have investigated and confirmed the presence of the holiday effect in the European returns, such as Arsad and Coutts (1997), who investigated UK's FT 30 Index over a 60-year timeframe, Krämer and Runde (1996) with their study on Germany's DAX Index, where the average return over holidays is more than 10 times larger than the non-holidays average return, Dodd and Gakhovich (2011), who analysed 14 Central and Eastern European markets, or Dumitriu et al. (2011), who analysed the Romanian market and found abnormal post-holiday returns along with the pre-holiday high returns. Vergin and McGinnis (1999) reported that the positive pre-holiday returns have disappeared for large firms and diminished for small firms between 1987 and 1996. Conversely, the hypothesis that the holiday effect has diminished or disappeared is rejected by Brockman and Michayluk (1998), whose results reveal a robust and persistent holiday effect after 1987. Hong and Yu (2009) confirm that the trading volume is lower during the summer because market participants are on holiday. Similarly, Al-Ississ (2010) reported significantly lower trading volume and changes in daily stock returns in 17 Muslim financial markets during the Muslim holy days of Ramadan and Ashoura. However, Bialkowski et al. (2010) found higher stock returns and no change in the trading volume during Ramadan.

The pre-holiday effect has been widely studied in an intra-market context, but very little attention has been paid to holidays in a cross-market context. This motivation introduces the review of the literature on the cross-market holiday effect.

4.2.2.3 Cross-Market Holiday Effect

Cheung and Kwan (1992) were the first authors to bring the volume dimension into the literature on the transmission of information across international markets. By computing the average Canadian daily volume during US holidays and US trading periods, and computing the ratios of these volume averages, they found evidence that the US trading holidays impact both volatility and trading volume in Canada's Toronto Stock Exchange (TSE); the Canadian trading volume drops when there is a holiday in the US. Similarly, they investigated the reverse causality, looking at the effect of Canadian holidays on the US market; despite finding a decrease in volumes, this was a significantly less blatant response. Cheung and Kwan's study concludes that the information originating from the US has a major impact on other markets, while the converse might not be valid. Casado et al. (2013) reported a significant US holiday effect on the European markets, with return rates above average and volatility/trading volume below average. The lower volume could be caused by the absence of US institutional investors and a lower macroeconomic information volume with less investor disagreement since the world's largest stock market and economic news source is closed; these factors change the public information flow and the European investor mix. There are significantly positive returns in the European stock markets when there is a holiday on the NYSE and their magnitude depends on the sign of the previous day's NYSE closure. They used recent financial data for the European stock market indices ranging from 1991 to 2008 and defined three measures for returns: off-market return (i.e. close-to-open return), intraday return (i.e. open-to-close return) and ordinary return (i.e. close-to-close return); in order to assess the impact of the six US holidays that occur on European trading days and compared the average of the sample return with the average of the returns during NYSE holidays before proceeding to fitting a regression model with dummy variables, the previous day's return and the ordinary trading volume on Mondays only. The authors decomposed the European returns into off-market returns and intraday returns in order to test whether the NYSE information is not totally reflected in the European prices before the markets shut, but found that the previous day's NYSE information is fully incorporated in the European opening prices and therefore it is irrelevant to the cross-market holiday effect. The US holidays that are non-holidays in Europe are:

1. Labour Day on the first Monday in September;
2. Presidents Day on the third Monday in February;
3. Memorial Day on the last Monday in May;
4. Independence Day on 4th July;
5. Thanksgiving Day on the fourth Thursday in November;
6. Martin Luther King Day (since 1998) on the third Monday in January.

On a partially related note, Meneu and Pardo (2004) examined the cross-market pre-holiday effect. Using the five most traded stocks in the Spanish Stock Exchange, which are also traded on the New York Stock Exchange and the Frankfurt Stock Exchange, they analysed a pre-holiday effect in the Spanish market prior to a US or German holiday, and found no such effect in their analysis sample.

The only significant pre-holiday effect in the Spanish market was domestic (i.e. prior to Spanish holidays) and not international.

There is research investigating and confirming all of the calendar effects (i.e. the weekend effect, turn-of-the-month, turn-of-the year, and holiday effects) and their persistence (Lakonishok & Smidt, 1988) (Barone, 1990) (Agrawal & Tandon, 1994) (Mills & Coutts, 1995). Other papers report the recent diminishing (or absence) of calendar effects for large-firm stocks starting from the late 1980s (Hansen, et al., 2005) (Pearce, 1996). Many popular anomalies do not hold up in different sample periods (Schwert, 2003).

4.2.2.4 Other Calendar Effects

Another part of the literature on calendar effects investigates the *intraday patterns of the bid/ask spreads*, having higher spreads at the trading day's open and close relative to the interior period (McInish & Wood, 1992) (Abhyankar, et al., 1997). A *day-end transaction price anomaly* causes a large mean price change on the last transaction, suggesting that stock values are not well represented by the closing price (Harris, 1989). Although summer months are profitable, September is the least profitable month of the year, being the only month in the US with a negative mean return, and is known as the *September effect* (Siegel, 2008). The *daylight saving effect* is an anomaly that causes large negative returns after the daylight saving weekends as a consequence of the sleep pattern changes; this effect is argued to be between two and five times larger than the regular weekend effect (Kamstra, et al., 2000), although other studies used extensive data sets and econometric techniques (Pinegar, 2002) to contradict the hypothesis that daylight saving weekends' mean returns are significantly lower than returns on the regular weekends, on the basis that its methodology is not robust (Müller, et al., 2009). The presidential *elections effect* is a particular example of a multi-seasonal calendar effect. Nippani and Medlin (2002) found that the delay in publishing the results of the USA presidential election in 2000 had a short-term negative impact on the market performance. The *Mark Twain effect* has been documented by Cadsby (1989), who found evidence of this effect on the Canadian stock returns. The name of this effect was coined by Cadsby, who used a citation from Mark Twain's novel, 'Pudd'nhead Wilson', to define the Mark Twain effect: "October. This is one of the peculiarly dangerous months to speculate in stocks. The others are July, January, September, April, November, May, March, June, December, August, and February" (Cadsby, 1989).

Bouman and Jacobsen (2002) found that the *Halloween effect* (or sell in May effect), which states that stock returns are significantly lower between May and October than during the rest of the year (i.e. from November to April), is present in 36 out of 37 countries from their sample. Another study extended Bouman and Jacobsen's research and confirmed that there is a significant Halloween effect up to the point when Japanese financial markets were internationalised in the mid-1980s, after which the Halloween effect disappeared (Maberly & Pierce, 2003).

Behavioural finance speculates other interesting hypotheses on a variety of potential exogenous determinants of trading volume, such as the *weather effect*. A defining study conducted by Saunders

(1993) argued that stock returns are systematically affected by weather due to investor psychology. Hirshleifer and Shumway (2003) also exploited this psychology concept and found that sunshine is strongly correlated with positive daily stock returns, allowing the investors who use a weather-based strategy to outperform the market, although the profits are negligible. Another study (Kamstra, et al., 2003) focused on the *seasonal affective disorder (SAD) effect*, i.e. a depressive condition affecting many people during the seasons with fewer daylight hours, and, after considering popular market seasonality factors, found a SAD effect in the seasonal cycles of the stock returns. On the other hand, Cao and Wei (2005) show evidence for a significantly negative correlation between temperature and stock returns, linking lower temperatures to higher returns and higher temperatures to higher or lower returns. Other authors, such as Krämer and Ralf (1997), Loughran and Schultz (2004) and Goetzmann & Zhu (2005), found no systematic relationship between weather and stock returns. The weather effect is unlikely to be applicable to the equity market nowadays due to round-the-clock trading and the high availability of electronic trading. The contradicting perspectives regarding the weather effect might also stem from the extent to which other well-known temporal anomalies are considered in the models (e.g. the monthly effect). Because of these reasons, the weather effect is not particularly studied in this thesis.

4.2.3 The Volume-Price Relation

The positive correlation between volume and price changes has been extensively studied in the finance literature (Harris & Raviv, 1993) (Hong & Stein, 2007). There are two forms that price changes can take in their positive correlation with volume: the magnitude (or absolute value) of the price change, i.e. $|\Delta p|$ (Assogbavi & Osagie, 2006) or the price change per se (or the raw price change value), i.e. Δp (Karpoff, 1987), where one can define the price change as either the log-price difference or the percentage price change.

4.3 Data Set

This section introduces the data sample retrieval and processing. The study employs a data sample that contains an extensive sample of 2,353 stocks. The financial market data set was complemented by a data set containing a representative temporal exogenous determinant of volume and non-stationarity, namely bank holidays. The data sample spans from 1st January 2000 to 10th May 2015 and investigates the markets during stable periods, but also during the financial crisis of 2007-08, which motivates a subsequent analysis of structural breaks before and after the crisis.

4.3.1 Stock Universe

We start from a list of indices, outlined in Table 4.1 along with their RICs (Reuters Identification Codes). The constituent list for each of these European indices is generated and is valid as of 11th May 2015 in order to create an optimal representation of the pan-European market.

Table 4.1
Market data European indices.

RIC	Name	RIC	Name
.STOXX	STOXX Europe 600 EUR Price Index	.PSI20	Euronext Lisbon PSI 20 Index
.FTSE	FTSE 100 Index	.OMXS30	OMX Stockholm 30 Index
.FTMC	FTSE Mid 250 Index	.OBX	Oslo Stock Exchange Equity Index
.FTLC	FTSE 350 Index	.OMXHPI	OMX Helsinki_PI
.FTSC	FTSE Small Cap Index	.BFX	BEL 20 Index
.FTAS	FTSE All Share Index	.OMXC20	OMX Copenhagen 20 Index
.GDAXI	Deutsche Boerse DAX Index	.ATG	Athex General Composite Share Price Index
.MDAXI	MDAX Performance Index	.ISEQ	ISEQ Overall Price Index
.SDAXI	SDAX Share Index	.JTOPI	Johannesburg Stock Exchange Top 40 Tradeable Index
.FCHI	CAC 40 Index	.ATX	Austrian Traded Index
.CN20	CAC Next20 Index	.FTMIB	FTSE MIB Index
.CACMD	CAC Mid 60 Index	.MSPE	MSCI International Pan Euro Price Index
.CACS	CAC Small Index	.MCX	MICEX Composite Index
.SSMI	Swiss Market Index	.WIG20	Warsaw SE WIG-20 Single Market Index
.AEX	Amsterdam Exchange Index	.TRXFLDEUPU	Thomson Reuters Europe Index
.IBEX	IBEX 35 Index		

We added the largest 42 South African stocks to our pan-European universe for a number of reasons: the trading of South African stocks is closely connected to the European stocks, South Africa operates on the same time zone as the Eastern Europe, i.e. UTC (Coordinated Universal Time) + 2 hours, and the Johannesburg Stock Exchange is a liquid trading venue. The frequency table of the market data sample in Table 4.2 shows the stock distribution by country. The country codes are encoded in the two-letter format specified by ISO 3166-1 alpha-2. A stock is assigned to a country based on the exchange this stock is trading at, e.g. a Spanish stock trading on the London Stock Exchange is associated with the United Kingdom.

Table 4.2
Market data sample country breakdown.

Country Code	Country Name	Stocks	Percent	Country Code	Country Name	Stocks	Percent
AT	Austria	32	1.36	HU	Hungary	4	0.17
BE	Belgium	62	2.63	IE	Ireland; Republic of	43	1.83
CH	Switzerland	104	4.42	IT	Italy	111	4.72
CZ	Czech Republic	5	0.21	NL	Netherlands	46	1.95
DE	Germany	176	7.48	NO	Norway	69	2.93
DK	Denmark	43	1.83	PL	Poland	65	2.76
ES	Spain	61	2.59	PT	Portugal	18	0.76
FI	Finland	130	5.52	SE	Sweden	158	6.71
FR	France	346	14.70	TR	Turkey	130	5.52
GB	United Kingdom	647	27.50	ZA	South Africa	42	1.78
GR	Greece	61	2.59				

4.3.2 Market Data

Daily market data containing OHLC (open, high, low, close) prices and end-of-day volume are retrieved from Thomson Reuters using a VBA script that automates the data retrieval. The stocks' daily market data is retrieved and augmented by computing their consolidated volume, i.e. the sum of a stock's main exchange trading volume and its volume on MTFs (multilateral trading facilities).

The consolidated volume has been used in this study in order to better reflect the actual volumes. Therefore, any subsequent reference to trading volume in this study indicates the consolidated volume.

The market data was pre-processed to account for missing data or incorrect data. For instance, the constituents of the FTSE MIB index could not be retrieved along with all the other indices' constituents, and, in this instance, manual intervention for data cleansing was required. The stocks whose number of market data available days is less than 100 days have been discarded. The stocks are augmented by metadata that includes exchange location, currency, company market capitalisation, economic sector, business sector name, industry/subindustry name, activity name etc. The final number of daily observations across all 2,353 stocks is nearly 7.2 million.

4.3.3 Construction of the Calendar Data Set

The calendar data consists of the European non-trading calendar for the 21 countries being analysed and for the USA, and was laboriously constructed from scratch in order to provide an accurate reflection of the special events occurring in the equity markets. A total number of 3,039 bank holidays are included in the calendar data. Due to data unavailability, half-trading days are not included in this study.

The non-trading calendar contains public holidays and bank holidays when the stock exchanges are closed. The data set covers the same period as the market data and contains the trading holidays for European countries and the United States of America. The non-trading calendar for the 21 European countries outlined in Table 4.2 and the USA (as a dominant financial market) was elaborated using multiple sources, ranging from the trading calendar on the exchange websites, and public holidays from www.timeanddate.com, to the empirical trading calendar inferred from this study's daily market data. The rationale of manually constructing the holiday calendar is twofold: first, high accuracy is crucial for identifying the extent to which volume is correlated to cross-market holidays; second, there is no available trading calendar that mirrors the observed activity for the European exchange venues and there are major differences between the non-trading calendars and the official holiday calendars that are publicly available.

A non-trading calendar CSV file was created for each of the trading countries. We started by getting a list of expected trading holidays by getting the zero-volume business days for each country. This ensured that the non-trading calendar is accurate from a financial market viewpoint. It is important to distinguish from the public holidays calendar of a country and the non-trading calendar for an exchange venue, since the latter might be owned by an international company (e.g. Euronext), which enforces a different trading calendar, or it might be located in a region with additional holidays, or unforeseeable events might occur (e.g. Hurricane Sandy, 11th September Terrorist Attacks etc.). No external source was able to accurately reflect the trading holidays for the entire study period and necessitated thorough implementation of a non-trading calendar. For the countries with few and potentially illiquid stocks that are listed in Table 4.3, the expected list of holidays was significantly larger than in reality due to many zero-volume days when the markets are actually open. Therefore, we got the data from the main indices from these countries and cross-checked the expected trading days as the methodology based on zero-volume would output more non-trading dates than the actual number. Additional data cleansing was performed for very few incorrect stocks that were trading during their exchanges' trading holidays.

Table 4.3
Low liquidity countries.

Country	Number of stocks	Main stock index	RIC
Czech Republic	5	Prague Index	.PX
Hungary	4	Budapest Index	.BUX
Portugal	18	Euronext Lisbon PSI20	.PSI20

4.3.3.1 Country-Specific Calendar Peculiarity

Each country has its own 'hidden' methodology of generating the public holidays calendar. When a public holiday falls on a weekend, it is substituted by the previous trading day in some countries (e.g. New Year's Eve in Austria and Belgium) or the next day in others, or it is not substituted at all. Some other countries have additional 'bridge' holidays when a holiday falls on a Tuesday or Thursday, in order to get a four-day weekend (e.g. Hungary and Poland).

Additional holidays are observed despite not being officially declared as public holidays. It is worth noting that periodic holidays can cease at certain times and others can be introduced. For example, the Swedish National Day started being celebrated from 2005; the Swiss National Day was not a public holiday for a few consecutive years, between 2001 and 2005; Good Friday in Hungary and Czech Republic was observed from 2012 and 2013, respectively; Christmas Eve is a non-trading day in Ireland until 2005 etc. On the other hand, a few countries like Norway and South Africa have a well-defined periodic structure for their bank holidays, although South Africa has a few one-off holidays for General Elections and Municipal Elections.

Surprisingly, the Greek exchanges are not trading on both catholic and orthodox Easters, apart from the years when they fall on the same date (i.e. 2001, 2004, 2007, 2010, 2011, and 2014) and 2013.

The longest holidays are Turkey's Festival of Sacrifice (or Eid al-Adha) and End of Ramadan (or Eid ul-Fitr).

Despite the fact that 1st May is not a bank holiday in the Netherlands, it is actually observed on the Amsterdam stock exchange after it merged with the Brussels and Paris stock exchanges to form Euronext in 2000; it became a non-trading day since 2002 though. In the Netherlands, 1st May is not an official day due to the Queen's Day, which was a public holiday in its own right until 2013, falling on 30th April. It was replaced in 2014 by the King's Day (falling on 27th April). After joining Euronext, the public holidays in Belgium occurring between 1st May and Christmas Eve became regular trading days on the exchange, starting from 2002. Similarly, the Portuguese trading calendar changed from 2003 after Lisbon joined Euronext in 2002, and the French trading calendar changed from 2001.

4.3.3.2 Normalised Trading Calendar

When constructing each country's calendar from scratch, the post-processing step ensured that each periodic holiday observation is identified by the same name (i.e. bank holiday normalisation). However, slight variations were identified when using the calendar of the European countries and USA altogether, and we had to define regular expression-based nomenclature transformations so that the alternative names of the various events are mapped to the same holiday concept. Some representative examples from the 56 rules include:

- From *May Day* (except Ireland), *Labour Day*, *Labor Day* (except USA), *Workers' Day*, *Labor and Solidarity Day* to *May 1st*;
- From *St. Stephen's Day*, *2nd Christmas Day*, *Second Day of Christmas*, *Synaxis of the Mother of God*, *Day of Goodwill*, and *Christmas Day* (occurring on 26th December) to *Boxing Day*;
- From *Independence Day* to *[country name] Independence Day*; from *Constitution Day* to *[country name] Constitution Day*; from *National Day* to *[country name] National Day*;
- From *Family Day* to *Easter Monday*;
- From *Spring Bank Holiday* (UK) and *Memorial Day* (US) to *GB US Spring Bank Holiday/Memorial Day*;
- From *Pentecost Monday* to *Whit Monday*;
- From *Dormition of the Holy Virgin* to *Assumption of Mary*;
- From *Independent Czechoslovak State Day* (CZ) and *The Ochi day* (GR) to *CZ GR Independent Czechoslovak State Day/The Ochi day*.

In some cases, we are discriminating between two or more holidays that have the same name, but are essentially referring to different holiday concepts (e.g. the generic *Independence Day* holiday, which occurs on different dates in US, Finland, Poland, Greece etc.). In other cases, we are aggregating holidays with different names which refer to the same concept to some extent, ranging from semantically identical holidays (e.g. *Pentecost Monday* and *Whit Monday*) to similar holidays falling on the same date annually (e.g. *Boxing Day*, *Second Christmas Day*, *St. Stephen's Day*).

The calendar also exhibits additional holidays issued by certain governments. Since these holidays are sparse observations that are not periodic, we appended 'Additional' in order to distinguish from the main holiday which they follow.

All of the holidays that are observed exclusively in one country are prefixed by the country code. Early May Bank Holiday is observed in the UK and Ireland.

Table 4.4 shows all of the normalised bank holidays, both periodic and non-periodic. This bank holiday normalisation found 95 unique pan-European non-trading events from 3,039 non-normalised bank holidays.

Table 4.4
Normalised bank holidays.

Holiday Name	Closed Markets	Holiday Date Observations	Trading Observations	Trading Stocks	Trading Countries
AT Austria National Day	1	10	17879	2127	20
All Saints' Day	6	11	19008	2132	18
Ascension Day	9	15	20106	1832	17
Assumption of Mary	7	13	20856	2190	20
BE Belgian National Day	1	1	1353	1353	19
Boxing Day	21	12	1206	130	1
Boxing Day Additional	4	8	11680	2236	21
CH Berchtold Day	1	11	18561	2246	20
CH Swiss National Day	1	8	15079	2232	20
CZ GR Independent Czechoslovak State Day, The Ochi day	2	10	18543	2276	19
CZ Jan Hus Day	1	10	18109	2143	20
CZ Saints Cyril and Methodius	1	11	19804	2200	20
CZ St. Wenceslas Day	1	10	18275	2146	20
CZ Struggle for Freedom and Democracy Day	1	10	18608	2339	20
CZ Victory in Europe Day	1	12	22377	2347	20
Christmas Day	21	11	1097	130	1
Christmas Eve	16	11	12069	1508	12
Corpus Christi	3	15	26562	2209	19
DE Day of German Unity	1	2	3435	2154	20
DK Ascension Day Additional	1	6	12435	2251	20
DK Denmark Constitution Day	1	10	17411	2252	20
DK Great Prayer Day	1	16	27545	2265	20
ES Assumption of Mary Additional	1	1	1577	1577	20
ES Hispanic Day	1	3	4379	1636	20
ES Spain Constitution Day	1	4	5543	1556	19
GB IE Early May Bank Holiday	2	16	19044	1664	20
Easter Monday	20	16	1755	169	2
Epiphany	6	11	16805	1955	17
FI Finland Independence Day	1	11	18922	2114	20
FR Bastille Day	1	1	1159	1159	19
GB Golden Jubilee Bank Holiday	1	1	1086	1086	20
GB Royal Wedding Bank Holiday	1	1	1512	1512	20
GB Summer Bank Holiday	1	15	20043	1691	21
GB The Queen's Diamond Jubilee	1	1	1500	1500	19
GB US Spring Bank Holiday, Memorial Day	3	17	22559	2241	21
GR Clean Monday	1	16	28667	2292	20
GR Greece Independence Day	1	11	18999	2292	20
GR Holy Spirit Monday	1	15	23115	2224	20
GR Orthodox Easter Monday	1	10	15073	2288	20
GR Orthodox Easter Tuesday	1	2	3577	2120	20
GR Orthodox Good Friday	1	11	17922	2292	20
Good Friday	21	16	1841	178	4
HU 1848 Revolution Memorial Day	1	11	19502	2171	20
HU 1848 Revolution Memorial Day Additional	1	5	9160	2130	20
HU 1956 Revolution Memorial Day	1	11	20110	2340	20
HU 1956 Revolution Memorial Day Additional	1	6	11456	2338	20
HU All Saints' Day Additional	1	5	9381	2156	20
HU Hungary National Day	1	11	20556	2333	20
HU Hungary National Day Additional	1	3	5687	2203	20
IE August Bank Holiday	1	1	1365	1365	20
IE June Bank Holiday	1	15	24780	2252	20
IE October Bank Holiday	1	2	2845	1541	20
Immaculate Conception	3	10	18328	2317	20
Maundy Thursday	3	16	28200	2240	20
May 1st	19	12	6518	878	6

Holiday Name	Closed Markets	Holiday Date Observations	Trading Observations	Trading Stocks	Trading Countries
May 1st Additional	3	9	15031	2285	21
Midsummer Eve	2	15	24377	2031	19
NL Queen's Birthday	1	1	1422	1422	20
NO 17 May Constitution Day (1814)	1	9	15372	2131	20
New Year's Day	22	16	5891	1517	20
New Year's Day Additional	4	5	8140	2301	21
New Year's Eve	18	11	12240	1510	13
New Year's Eve Additional	3	5	8477	2113	21
PL Poland Constitution Day	1	11	18155	2128	20
PL Poland Independence Day	1	10	18687	2279	20
PT Carnival/Shrove Tuesday	1	3	4220	1550	20
PT Liberty Day	1	3	4261	1552	20
PT Portugal Day	1	1	1520	1520	20
PT Republic Implantation	1	2	2841	1508	20
PT Restoration of Independence	1	1	1424	1424	20
SE Sweden National Day	1	8	14981	2145	20
TR Commemoration of Ataturk, Youth and Sports Day	1	10	17305	2158	20
TR National Sovereignty and Children's Day	1	12	21599	2221	20
TR Ramadan Feast	1	41	67097	2203	20
TR Sacrifice Feast	1	52	87713	2212	20
TR Turkey Republic Day	1	11	19688	2215	20
TR Victory Day	1	10	15826	2081	20
US 11 September Terrorist Attacks	1	4	5784	1542	21
US Independence Day	1	15	27570	2326	21
US Labour Day	1	15	27831	2333	21
US Markets closed - Hurricane Sandy	1	2	4148	2156	21
US Martin Luther King Day	1	16	29615	2351	21
US National Day of Mourning for President Gerald R. Ford	1	1	1706	1706	19
US National Day of Mourning for President Ronald Reagan	1	1	1608	1608	21
US Presidents Day (Washington's Birthday)	1	16	29453	2351	21
US Thanksgiving Day	1	15	28065	2350	21
Whit Monday	10	15	21890	2112	17
ZA Day of Reconciliation	1	13	24339	2310	20
ZA Freedom Day	1	14	25490	2309	20
ZA General Elections	1	3	5736	2240	20
ZA Heritage Day	1	13	23653	2294	20
ZA Human Rights Day	1	12	21306	2223	20
ZA Municipal Elections	1	3	5187	2052	20
ZA National Women's Day	1	12	21381	2164	20
ZA Youth Day	1	12	21573	2256	20

4.4 Analysis Approach

In this section, we outline the analytical approach we followed throughout this study and we outline the experiment methodology. We start by asking the question whether the cross-market holiday effect is real. Then, we check whether it can be modelled and we are looking at the feasibility of modelling this effect with the view of predicting using ridge regression, along with additional features (besides the bank holiday-specific indicator variables). Finally, we raise the question of how well the cross-market holiday effect can be modelled the further we go into the future. We investigate how accurate the predictions of the volume are n days ahead of time. This is motivated by the use case of multi-step ahead analysis, where traders and portfolio managers who do not work on bank holidays and try to size the allocation of a trade with a view to gauging the available liquidity and minimising the market impact.

For each cross-market holiday of every stock, we would like to compare the trading volume on the special event (i.e. 'target date' or t_0) with its benchmark volume. The benchmark volume is defined as the median of the previous 20 trading days' volumes, since the median helps dampen the effect of outliers. Besides the default 1-step ahead analysis, a multi-step ahead forecasting is provided for step size *n*, ranging from 2 to 6; for example, if *n = 6*, we could use the 20 most recent trading days' volumes up to today in order to predict the impact on the trading volume in 6 days' time. Figure 4.1 illustrates the lower trading volumes on cross-market holidays, compared to the median of the previous 20 trading days (i.e. benchmark volume) on the logarithmic scale.

The data necessitates further processing in order to compute the relative volume. It is the log-ratio between the t_0 volume on the cross-market holiday and the median of the benchmark volumes, which can be lagged depending on the step size for the step ahead analysis, where $l = n - 1$:

$$V = \log \frac{V_{t0}}{\text{median}(V_{t-l-1}, V_{t-l-2}, \dots, V_{t-l-20})}. \tag{4.1}$$

We are dealing with data that is sparse and there is a limited number of holidays. Consequently, we normalise the analysis data in order to increase the number of observations and find effects that are common to a basket of stocks and a particular event.

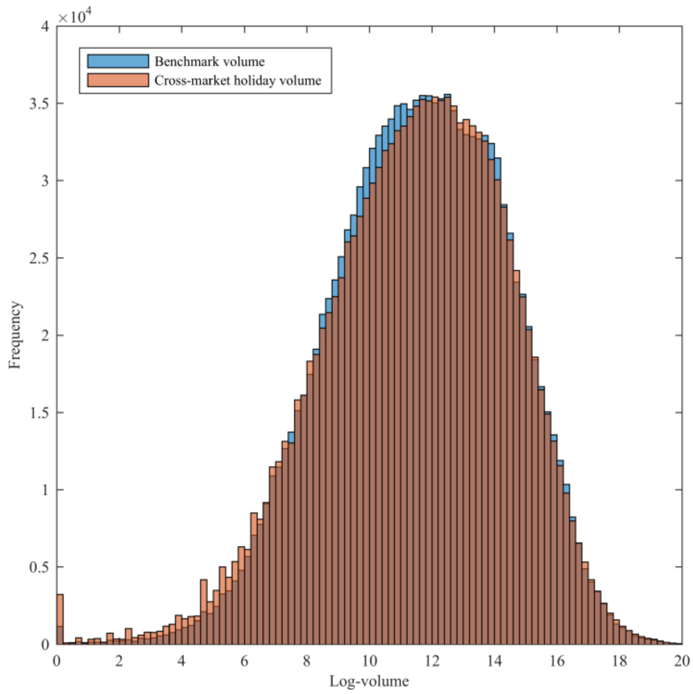

Figure 4.1: Histograms of the logarithmic volume data for the entire stock universe on cross-market holidays and on the benchmark period.

Figure 4.2 shows three alternative methods of aggregating the benchmark volumes and exhibits the relative volume, whose formula is shown in Equation (4.2), where V_x is the volume on a day when there is a cross-market holiday, and the aggregating function is the median in Panel A, the arithmetic mean in Panel B, and the geometric mean in Panel C:

$$V = \log \frac{V_x}{f_{\text{aggr}}(V_{t-1}, V_{t-2}, \dots, V_{t-20})}. \tag{4.2}$$

The median and the geometric mean output similar results in terms of skewness, although the geometric mean outliers are more dispersed. Arithmetic mean is appropriate for independent variables; however, especially in the financial world, there are many instances when the arithmetic mean is inappropriate for computing averages. We use the median as a measure of central tendency since volumes are rather volatile and the median is extremely robust to outliers.

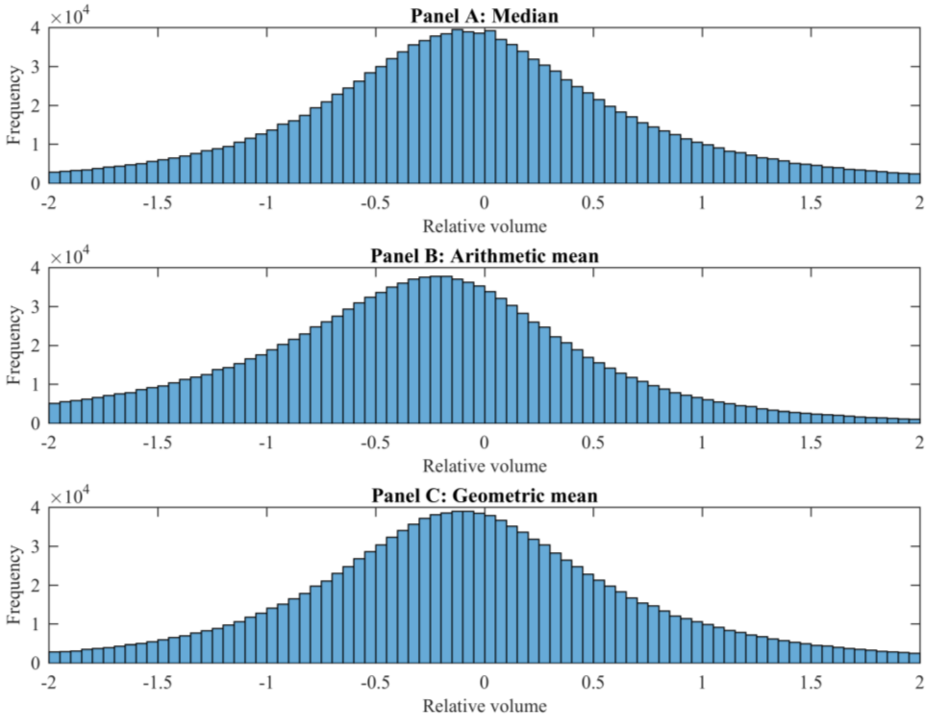

Figure 4.2: Histograms of the relative volume on cross-market holidays across the entire stock universe.

The bar charts in Figure 4.3 show the median cross-market holiday effect, expressed in linear space percentage values, for the trading volume in the United Kingdom, Germany, France and Switzerland, for each external holiday country. These represent the volume decrease percentage, compared to the benchmark volume.

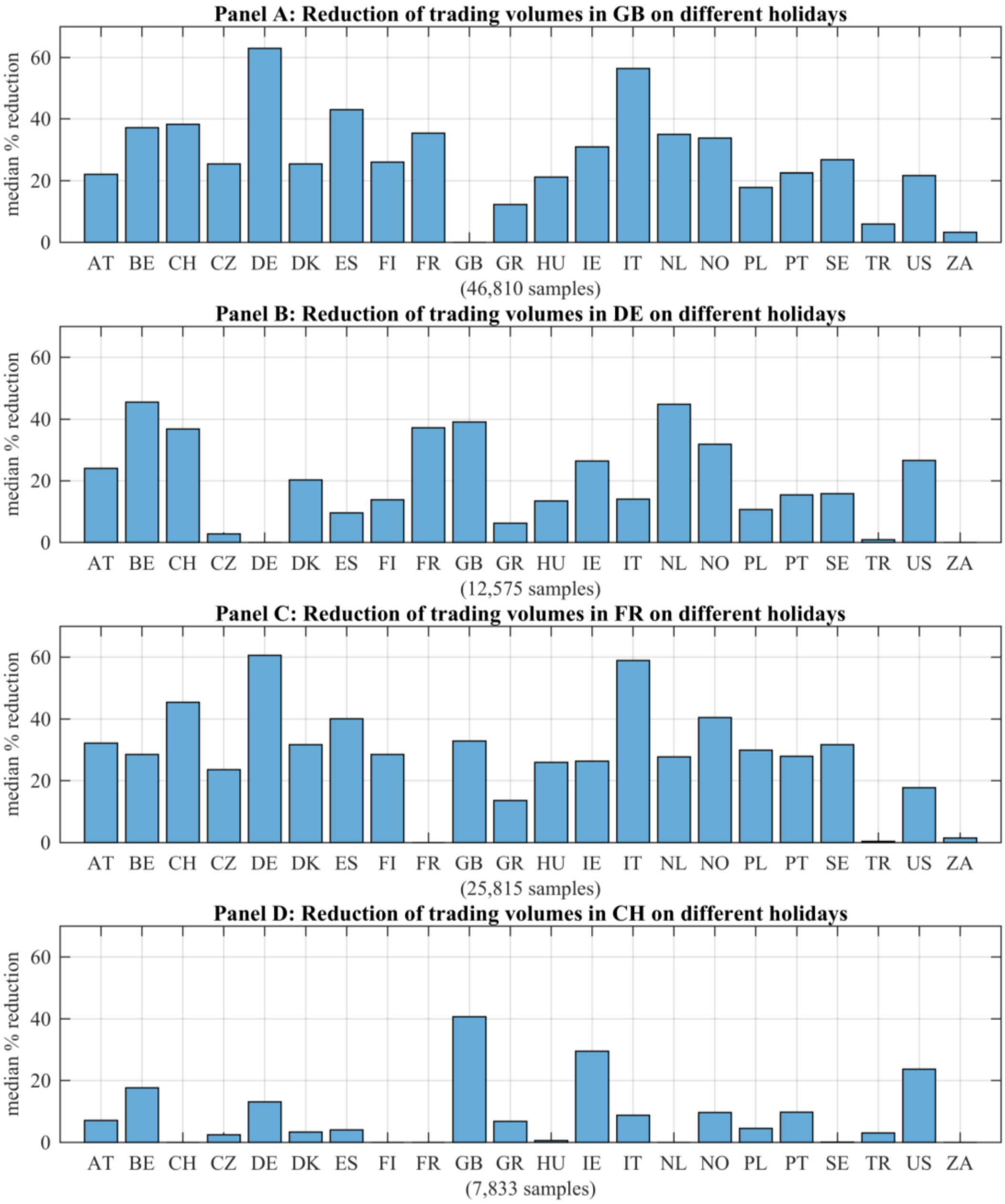

Figure 4.3: Cross-market holiday effect on the trading volume in the United Kingdom, Germany, France and Switzerland, shown for each external holiday country.

Table 4.5 shows the median percentage reduction in volume on cross-market holidays for every pair of countries, expressed in linear space; the columns represent the holiday countries, while rows represent the trading countries.

Table 4.5
Cross-market holiday effect showing the median percentage reduction in volume for each pair of countries.

	AT	BE	CH	CZ	DE	DK	ES	FI	FR	GB	GR	HU	IE	IT	NL	NO	PL	PT	SE	TR	US	ZA
AT	0.0	-43.5	21.0	9.0	8.9	3.4	0.0	-2.0	32.9	35.4	9.1	1.5	25.4	0.0	16.4	1.0	10.6	-8.1	-10.0	1.6	24.4	1.5
BE	37.4	0.0	53.9	21.3	68.7	37.4	50.6	32.7	31.1	35.9	14.3	28.4	28.6	65.8	29.4	47.8	32.7	32.8	35.8	1.3	22.9	1.5
CH	7.1	17.7	0.0	2.4	13.1	3.3	4.0	-3.0	-20.7	40.7	6.8	0.5	29.5	8.8	-7.8	9.7	4.5	9.8	0.1	3.1	23.7	-0.2
CZ	19.4	54.4	31.9	0.0	44.8	20.4	26.6	15.2	54.4	43.8	10.9	13.4	33.9	33.8	49.2	23.1	18.4	44.4	23.2	5.4	25.9	12.4
DE	24.1	45.5	36.8	2.8	0.0	20.4	9.5	13.8	37.2	39.1	6.2	13.5	26.5	14.0	44.8	31.9	10.7	15.4	15.8	0.9	26.6	-0.8
DK	3.5	44.0	14.5	10.8	32.4	0.0	14.5	10.9	46.1	30.3	3.8	5.0	24.1	11.8	46.7	27.9	3.9	22.6	9.6	3.9	25.5	6.4
ES	24.1	2.6	26.4	6.7	44.9	19.1	0.0	20.0	26.8	40.6	14.8	17.8	26.0	53.2	17.5	29.1	20.9	12.2	15.5	2.6	25.9	2.3
FI	7.9	42.7	17.5	12.3	14.7	16.8	0.0	0.0	28.8	30.6	6.4	4.3	22.0	0.0	29.1	29.0	4.4	5.9	19.2	0.0	23.1	-2.5
FR	32.2	28.5	45.4	23.6	60.7	31.7	40.1	28.5	0.0	32.9	13.6	26.0	26.3	58.9	27.8	40.5	29.9	27.9	31.7	0.4	17.7	1.5
GB	22.0	37.1	38.2	25.4	63.0	25.4	43.1	26.0	35.4	0.0	12.2	21.1	31.0	56.4	35.0	33.8	17.8	22.5	26.8	6.0	21.7	3.2
GR	14.0	37.6	24.0	18.0	38.3	10.0	25.7	17.6	29.5	20.0	0.0	8.8	18.6	48.6	34.7	16.3	14.5	7.7	19.3	-4.0	11.2	6.5
HU	10.6	83.5	42.2	13.8	71.6	21.9	39.1	10.2	86.0	55.5	1.1	0.0	40.6	47.7	86.0	32.3	6.1	43.5	17.5	4.7	44.5	14.1
IE	30.8	67.0	49.2	26.9	76.9	27.8	50.1	29.8	66.0	70.1	19.4	29.0	0.0	66.9	61.5	45.3	23.3	29.1	33.1	3.0	32.7	7.0
IT	22.2	21.9	19.4	5.2	22.2	8.4	34.4	9.6	18.0	37.3	10.2	15.3	30.9	0.0	20.4	13.5	10.4	28.0	10.3	1.4	21.3	4.5
NL	25.9	23.8	49.7	21.0	76.2	31.3	39.5	26.3	23.1	46.7	10.0	21.4	36.0	70.2	0.0	44.2	18.0	15.5	29.4	3.4	32.7	1.5
NO	-2.4	-7.4	14.7	11.3	-2.1	-4.6	1.0	2.1	-25.0	35.9	1.3	1.3	28.7	-20.9	-38.0	0.0	-0.4	-0.5	1.9	-0.6	26.0	0.2
PL	13.0	15.0	23.1	9.1	38.8	15.3	8.0	3.6	42.0	30.4	9.0	18.5	24.4	40.5	29.3	15.3	0.0	13.5	8.9	-1.2	25.0	5.0
PT	31.1	17.8	37.8	17.2	56.1	29.1	46.6	22.7	24.1	39.3	17.8	25.1	29.7	59.6	25.5	38.8	31.7	0.0	23.2	2.1	26.8	7.9
SE	3.6	-11.2	24.3	11.4	9.8	23.1	-3.6	0.9	3.5	31.0	5.3	8.9	24.3	-6.9	26.0	37.1	2.0	0.7	0.0	0.9	22.1	-1.3
TR	12.5	22.1	20.2	19.8	23.9	17.1	17.6	14.4	23.7	22.1	16.2	13.4	19.4	22.0	23.2	16.6	14.7	16.9	16.5	0.0	18.5	11.9
ZA	18.0	30.4	47.2	29.6	78.2	23.4	48.9	24.8	39.4	43.2	15.0	20.7	31.2	73.7	20.0	34.0	11.2	6.7	26.3	4.9	28.9	0.0

The cross-market holidays analysis is extended by a further investigation of the discrimination between market capitalisation classes. This particular analysis uses a restricted data set for three countries (UK, France and Germany), each with three indices for small-, mid-, and large-cap stocks, although France contains four indices because there are two CAC indices for large-cap indices, as shown in Table 4.6.

Table 4.6
Market capitalisation indices.

Index	Country	Small-Cap Index	Mid-Cap Index	Large-Cap Index
FTSE	UK	FTSE Small Cap Index	FTSE Mid 250 Index	FTSE 100 Index
DAX	Germany	SDAX Index	MDAX Index	DAX 30 Index
CAC	France	CAC Small Index	CAC Mid 60 Index	CAC 40 Index CAC Next20 Index

The data analysis starts with an exploration of the cross-market holiday effect using rigorous randomisation tests. Once the existence of this phenomenon is confirmed, we build a predictive model for trading volume.

4.5 Randomisation Analysis

In this first part of the analysis, we assess the statistical significance of the existence of the cross-market holiday effect. Then, we examine whether the phenomenon is driven by the Monday effect, and whether its impact on the trading volume is different based on the stock market capitalisation.

The aim of the randomisation (or permutation) tests is to assess whether two vectors X and Y are significantly different. The procedure starts by calculating the observed statistic, which is the difference between the two vector means. If the test is two-tailed, then the observed statistic is the absolute value of this difference. Then, the labels of vectors X and Y are randomised and the randomised statistic is recomputed in the same manner as the observed statistic, based on the newly randomised vectors. This step is repeated 1,000 times. Eventually, the randomisation test checks if the randomised differences are more extreme than the observed data. This allows computing an empirical p-value that corresponds to the percentage of times when the observed difference is larger (for the right-tailed and the two-tailed tests) or smaller (for the left-tailed) than the randomised differences. The significance level for the randomisation tests is $\alpha = 5\%$. The null hypothesis is rejected when the empirical p-value is less than the significance level.

The cross-market holiday and Monday bank holiday randomisation tests involve a pairwise shuffling of labels. In this instance, for each target date, we compute an artificial control date that is conditioned on the original target date. Since the vectors have the same size, we flip a coin for each element and decide whether the elements are to be interchanged.

Throughout these permutation tests, we also investigate the existence of potential structural breaks. We test the validity of structural homogeneity by splitting the data set covering almost 16 years into two folds: the first sample period half is between 1st January 2000 and 31st December 2007, while the second sample half covers 1st January 2008 – 10th May 2015. The motivation stems from the financial crisis of 2007-08, which culminated with the collapse of Lehman Brothers on 15th September 2008.

Three types of randomisation tests are performed in order to test a number of aspects regarding the cross-market holidays, such as the differentiated effect magnitude depending on market capitalisation or determining whether the Monday bank holidays are the drivers of lower trading volumes on Monday or whether it is the Monday effect that impacts on the trading volume.

4.5.1 Cross-Market Holidays vs. Control Dates

For the randomisation between cross-market holidays and their control dates, we defined the target volumes as the relative volume of a stock on the days (i.e. the ‘target dates’) when there is at least on cross-market holiday. For each stock, we iterate each of its unique target dates and compute a pairwise control date such that it is a trading day when there are no cross-market holidays, it falls on the same day-of-the-week as the target date, and it is within a +/- 2-month interval relative to the target date. If more than one control dates are found for a given target date, we pick the control date that is closest to the target date (i.e. the control date whose calendar day difference relative to the target date is the lowest). We define the relative trading volume for these control date as the control volumes of the randomisation test. Since this is an instance of a pairwise randomisation test, each target date has one paired control date. Based on this methodology, we perform a two-tailed and a left-tailed pairwise randomisation test. The null hypothesis of the two-tailed test is that the difference of the relative volumes of the dates with cross-market holidays and the dates without cross-market holidays comes from a distribution with mean equal to zero. The left-tailed randomisation tests the alternative hypothesis that the relative volume mean on cross-market holidays is less than the mean on dates without cross-market holidays.

The volumes on the cross-market holidays are significantly lower than those on the control dates (i.e. with no cross-market holidays), as shown by the randomisation tests' *p*-values in Table 4.7. This is also valid for the multi-step ahead forecasts (*n*-step ahead analyses for *n* between 2 and 6; here, the number of target dates becomes 1,343,485 - 1,343,481). The results are consistent for the two sample period halves. Figure 4.4 reveals the relative volume distribution for dates with cross-market holiday (i.e. target dates) and dates with no cross-market holidays (i.e. control dates), with overlapped histograms. The cross-market holidays show a slightly positive skew compared to the control dates. The median of the relative volume change for the dates with cross-market holidays is -8.882%. This relative change is expressed as a percentage on the natural logarithm scale and, after exponentiation, it corresponds to a reduction of volume in linear space by 8.499% compared to the benchmark volume.

Table 4.7
Randomisation tests: cross-market holidays vs. control dates.

Sample period/s	*n*-step ahead	Stocks	Target dates	Randomisation tail/s	*p*-value	Reject H_0
2000-2007, 2008-2015, 2000-2015	1	2,353	1,343,487	Both	0	Yes
2000-2007, 2008-2015, 2000-2015	1	2,353	1,343,487	Left	0	Yes

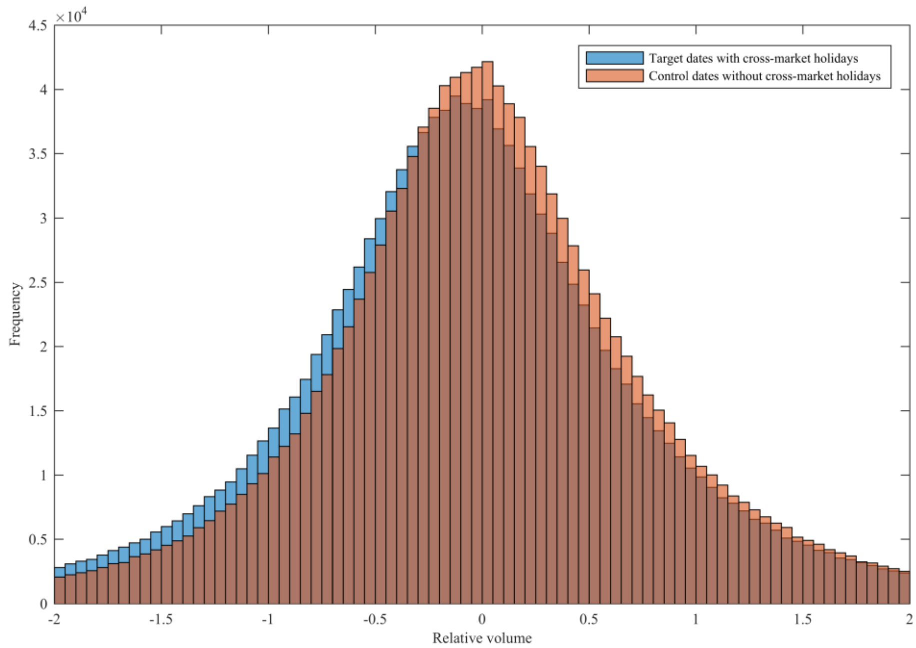

Figure 4.4: Relative volume distribution for the cross-market holiday target and control dates.

4.5.2 Monday Bank Holidays vs. Regular Mondays

The target dates for the Monday bank holiday randomisation test consist of all of the trading Mondays for a given stock when there is at least one cross-market holiday. The pairwise control dates consist of the closest Monday relative to a target date, falling within a +/- 2-month time interval and having no cross-market holidays. The test randomises the relative volumes of the target dates and the control dates. We performed a left-tailed test and a two-tailed test and found that the volumes on the cross-market holidays falling on Monday are significantly lower than the volumes of the Mondays with no cross-market holidays, as indicated by the results in Table 4.8. The results are consistent among the multi-step ahead analyses. There are no structural breaks around the financial crisis of 2007-08, as we observe a significantly lower volume on Monday cross-market holidays throughout the two 7-year time periods. As with the cross-market holiday randomisation test, the null hypothesis of the two-tailed test is that the difference of the relative volumes of Mondays with cross-market holidays and Mondays without cross-market holidays comes from a distribution with mean equal to zero. The left-tailed randomisation tests the alternative hypothesis that the relative volume mean on Monday cross-market holidays is lower than the mean on Mondays without cross-market holidays.

Table 4.8
Randomisation tests: Monday bank holidays vs. regular Mondays.

Sample period/s	*n*-step ahead	Stocks	Target dates	Randomisation tail/s	*p*-value	Reject H_0
2000-2015	1-6	2,353	424,976	Both	0	Yes
2000-2015	1-6	2,353	424,976	Left	0	Yes
2000-2007	1	1,997	188,740	Both	0	Yes
2000-2007	1	1,997	188,740	Left	0	Yes
2008-2015	1	2,353	234,212	Both	0	Yes
2008-2015	1	2,353	234,212	Left	0	Yes

Figure 4.5 illustrates the relative volume distribution for Mondays with cross-market holidays (i.e. target dates) and Mondays with no cross-market holidays (i.e. control dates for the Monday effect), with overlapped histograms.

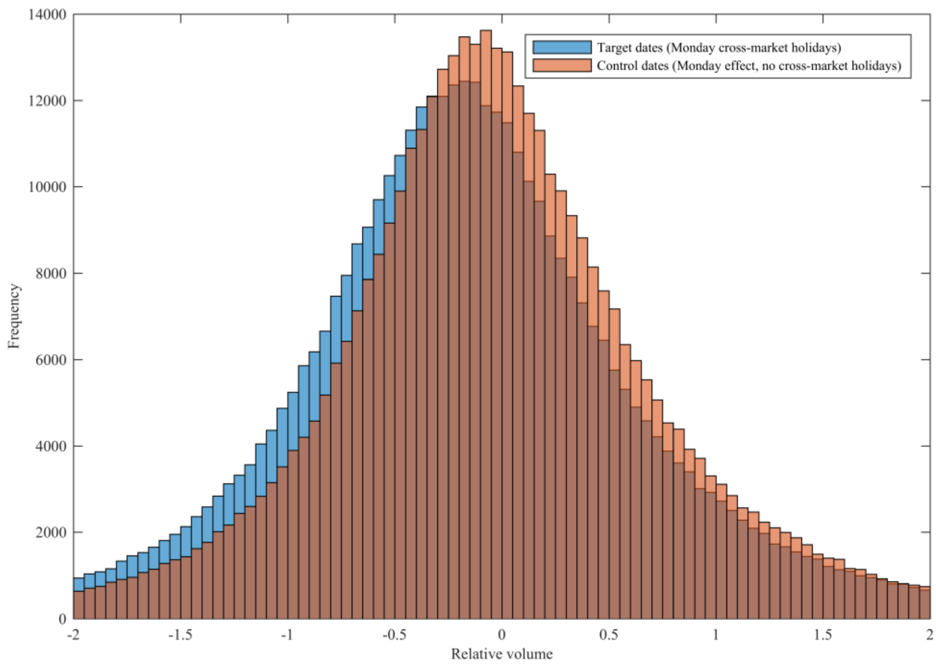

Figure 4.5: Relative volume distribution for the Monday bank holiday target and their control dates.

4.5.3 Small vs. Mid vs. Large Market Capitalisation

We further investigate whether the cross-market holiday effect might be less conspicuous or even absent in any of the market capitalisation classes. The methodology for the randomisation test for market capitalisation is slightly different from the previous pairwise randomisation tests. Given X small-cap stock, Y mid-cap stocks, and Z large-cap stocks, we define the observed test statistic using Equation (4.3), where $\bar{X}$, $\bar{Y}$, and $\bar{Z}$ represent the mean volumes of stocks X, Y, and Z respectively:

$$\text{Observed statistic: } |\bar{X} - \bar{Y}| + |\bar{Y} - \bar{Z}| + |\bar{X} - \bar{Z}|. \tag{4.3}$$

The test randomises the stock capitalisation classes for vectors X, Y and Z, and re-computes the randomised statistic as the new sum of pairwise absolute differences in means for the market capitalisations. We perform a two-sample absolute value randomisation test (i.e. two-tailed) in order to test the null hypothesis that the pairwise market capitalisation differences come from the same distribution, i.e. that there is a sum of absolute differences that is persistent. We expect the statistic on the structured data to be an extreme value and the shuffled values to be much lower. The test is performed for the main European market capitalisation-based indices: FTSE, CAC, and DAX. We test each index individually and then we aggregate the three indices and test them altogether.

Based on the results of the two-tailed randomisation tests in Table 4.9, we report that FTSE, DAX, and the aggregated indices exhibit a market capitalisation-based differentiation of the cross-market holiday effect. The CAC index does not exhibit significant differences across the market capitalisation classes. The multi-step ahead analyses have identical test outcomes, although the p-values increase slightly for the German and French indices, but they support the same null hypothesis rejection decisions.

Table 4.9
Market capitalisation randomisation tests.

Sample period/s	*n*-step ahead	Country	Index/Indices RIC	Stocks	*p*-value	Reject H_0
2000-2015	1	GB	.FTSC	290	0	Yes
			.FTMC	249	0	Yes
			.FTSE	100	0	Yes
	1	DE	.SDAXI	50	0	Yes
			.MDAXI	50	0	Yes
			.GDAXI	30	0	Yes
	1	FR	.CACS	223	0.088	No
			.CACMD	59	0.088	No
			.FCHI and .CN20	56	0.088	No
	1	Pan-Euro	Pan-Euro Small Cap	563	0	Yes
			Pan-Euro Mid Cap	358	0	Yes
			Pan-Euro Large Cap	186	0	Yes
2000-2007	1	GB	.FTSC	240	0	Yes
			.FTMC	206	0	Yes
			.FTSE	93	0	Yes
	1	DE	.SDAXI	40	0	Yes
			.MDAXI	40	0	Yes
			.GDAXI	30	0	Yes
	1	FR	.CACS	191	0.889	No
			.CACMD	53	0.889	No
			.FCHI and .CN20	54	0.889	No
	1	Pan-Euro	Pan-Euro Small Cap	471	0.12	No
			Pan-Euro Mid Cap	299	0.12	No
			Pan-Euro Large Cap	177	0.12	No
2008-2015	1	GB	.FTSC	290	0	Yes
			.FTMC	249	0	Yes
			.FTSE	100	0	Yes
	1	DE	.SDAXI	50	0.349	No
			.MDAXI	50	0.349	No
			.GDAXI	30	0.349	No
	1	FR	.CACS	223	0.021	Yes
			.CACMD	59	0.021	Yes
			.FCHI and .CN20	56	0.021	Yes
	1	Pan-Euro	Pan-Euro Small Cap	563	0	Yes
			Pan-Euro Mid Cap	358	0	Yes
			Pan-Euro Large Cap	186	0	Yes

We find a structural break for the distinctive impact of cross-market holidays on the three market capitalisation classes. For example, during 2000-2007, only FTSE and DAX have a significantly different impact based on market capitalisation; this is similar to the randomisation test performed on the entire sample period, except for the aggregated indices, which do not exhibit a market capitalisation differentiation before the financial crisis. In the second period following the financial crisis, FTSE, CAC and the aggregated indices show a significantly different influence of the cross-market holidays on the market capitalisation classes. We report a reverse effect for the DAX and CAC indices following the financial crisis.

Figure 4.6 shows the cumulative distribution for the relative volume on cross-market holidays for each market capitalisation class. Each market capitalisation class is computed by aggregating the stocks from the three stratified indices - FTSE, DAX, and CAC. It is important to note that a few stocks, which are constituents of an index, might be left out because of missing trading data, e.g. FTSE Mid 250 Index has 250 constituents, whereas the analysis uses 249. The small-cap stocks have a widening

CDF curve, suggesting their high susceptibility to lower volumes caused by cross-market holidays, whereas the mid-cap and the large-cap stocks have a progressively sharper curve.

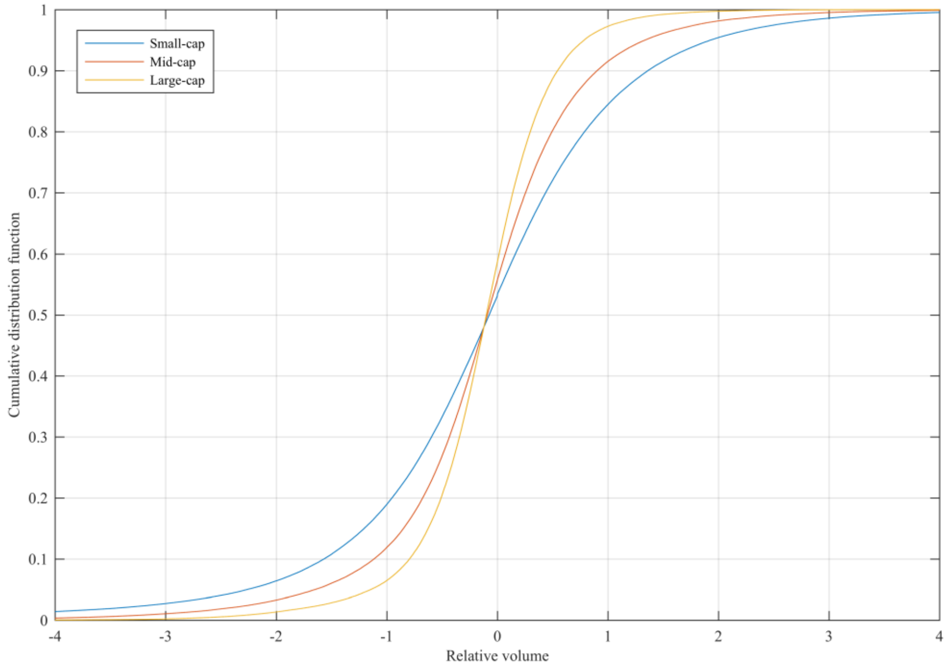

Figure 4.6: Cumulative distribution for the relative volume on cross-market holidays for each market capitalisation.

4.6 Predictive Modelling

We fit a ridge regression model for each variant of the cross-market holiday effect models. All of these contain a constant term (or intercept) and the dependent variable consists of the relative volume. We reduce the variability and numerical instability of these models by identifying an 'appropriate' value for the shrinkage parameter λ, such that it provides the lowest cross-validation MSE based on the proposed 2-section search, and it shrinks and stabilises the coefficients, as illustrated in the ridge trace in Figure 4.9. It is important to note that some coefficients presented in the Results section are very close to zero, but not exactly zero, because ridge regression normalises the data and therefore the zero indicator variables have a noise in the resulting model. However, the results are straightforward to interpret, since such values are negligible, whereas the real effects are reflected in large coefficient sizes.

4.6.1 Ridge Regression

Unlike classical variable selection techniques, where variables are assessed in a discrete manner (i.e. they are either kept in the model or excluded), resulting in a reduced model that is interpretable and might have a lower prediction error than the full model, shrinkage (or regularisation) methods are more continuous and provide less variability (Hastie, et al., 2011). Eliminating 'non-significant' predictors can result in large prediction biases; ridge regression solves this problem by using small proportions of all the variables, instead of using some variables entirely and none of the other ones that are considered insignificant by the variable selection process (Marquardt & Snee, 1975). This is the rationale of biased estimators and provides our motivation for using ridge regression instead of least squares and variable selection. An improved mean squared error (MSE) is achieved at the cost of introducing some bias, while greatly reducing the variance. The bias-variance trade-off balances the following two concepts: increasing the local structure/curvature by making the model more complex, and making the coefficients susceptible to high variance by including more terms in the model. The main problem arises when linear regression models contain many correlated variables and therefore their coefficients are poorly identified and have high variance. For example, a variable with a large positive coefficient can be cancelled by another variable that is correlated with the former and has a similarly large negative coefficient. Therefore, OLS performs poorly on new data (especially outside the training data region) when the data is ill-conditioned.

This study is based on ridge regression (Hoerl & Kennard, 1970), which is similar to least squares, but the regression coefficients are constrained by imposing a penalty on their size. The ridge coefficients minimise a penalised residual sum of squares (Hastie, et al., 2011), outlined in Equation (4.4), and results in an orthogonal system:

$$\beta_{\text{ridge}} = \operatorname{argmin}\left\{\sum_{i=1}^{N}\left(y_i - \beta_0 - \sum_{j=1}^{p} x_{ij}\beta_j\right)^2 + \lambda\sum_{j=1}^{p}\beta_j^2\right\}. \tag{4.4}$$

Lambda ($\lambda \geq 0$), which is called the shrinkage/tuning/complexity/regularisation parameter and commonly denoted by λ, is the complexity parameter controlling the amount of shrinkage (i.e. the strength of the penalty term). When $\lambda = 0$, the solution is the linear regression estimate. The larger the lambda, the greater the amount of shrinkage, i.e. the coefficients are shrunk toward zero and toward each other. If $\lambda = \infty$, then the coefficients are all set to zero and an intercept-only model is obtained. When searching for λ, one balances two ideas, i.e. shrinking the coefficients and fitting a linear model.

An important step in performing ridge regression is to generally standardise the input variables before solving Equation (4.4), since the ridge solutions are not equivariant under scaling of the inputs. This is appropriate whenever the model includes a constant term. Not standardising the predictors causes ill-conditioning due to the arbitrary origins of the scales on which the predictors lie. Centring the data cancels the non-essential ill-conditioning and reduces the variance inflation in the coefficient estimates. In the particular context of linear models, centring removes the correlation between the intercept and the other terms, while scaling allows the equation to be interpreted and used in a straightforward manner (Marquardt & Snee, 1975).

The coefficient of the constant term (i.e. the intercept β_0) is not affected by the penalty term. The rationale is that its penalisation would make the ridge process depend on the origin chosen for Y, i.e. adding a constant c to each target value y_i would not result simply in a shift of the predictions by the same amount as the constant c (Hastie, et al., 2011). Since β_0 is not penalised, one estimates it by the sample mean of the response variables using Equation (4.5). When the input matrix X is standardised and the linear model contains a constant term, this estimation of β_0 is significantly better than estimating β_0 in the model using least squares (Bertie & Cran, 1985). The other coefficients for the p predictors are estimated by a ridge regression without intercept, using the centred x_{ij}. Therefore, we assume that at this step centring was performed and that the input matrix X has p columns instead of $p+1$ columns:

$$\beta_0 \cong \bar{y} = \frac{1}{N}\sum_{i=1}^{N} y_i. \tag{4.5}$$

The ridge regression solutions are determined using Equation (4.6), where I is the p x p identity matrix:

$$\beta_{\text{ridge}} = (X^{\mathrm{T}}X + \lambda I)^{-1}X^{\mathrm{T}}y. \tag{4.6}$$

The ridge regression solution is a linear function of y due to the choice of the quadratic penalty $X^{\mathrm{T}}X$. By adding a positive constant to the diagonal of the penalty $X^{\mathrm{T}}X$ before inversion, the problem becomes non-singular, even if $X^{\mathrm{T}}X$ is not of full rank. This was the main motivation of ridge regression when it was introduced by Hoerl and Kennard (1970).

It is important to conclude that it is a common practice to firstly eliminate an all-constant column in the input matrix X (i.e. the constant term or intercept), and then centre and scale the predictors in order to have mean zero and unit standard deviation, before computing the ridge coefficients. If the predictors have different scales, the shrinking is not fair because the predictors would have different contributions to the penalised term, which is calculated as the sum of squares of all the coefficients. This is the reason why the optimal value of lambda is generally smaller, as it is associated with a smaller sum of squares of the coefficients.

We use ridge regression instead of lasso because there are weak effects in our model for cross-market holidays and lasso is inappropriate in this situation, potentially resulting in an excessively reduced model.

Identifying λ

The most common technique of determining a good value for the shrinkage parameter is to try various values (e.g. using grid search) and cross-validate the models for each lambda value such that the shrinkage parameter minimises the MSE. Too little regularisation might not be able to solve the numerical instability issues (i.e. matrix singularity) and therefore the lambda value has to be increased in order to find a threshold above which lambda solves the multicollinearity problem.

Generally, there is an 'optimum' value for lambda and the practical methodology is to explore potential values of λ between 0 and 1 (Marquardt & Snee, 1975) by investigating a range of 'admissible' values of λ (i.e. having smaller MSE than the OLS). Another empirical finding of Marquardt and Snee suggests that models without a constant term generally require smaller values of λ (i.e. $\lambda \leq 0.01$) than the models with an intercept.

Introduced by Hoerl and Kennard (1970), the ridge trace is a graphical representation of the coefficients' sensitivity to lambda, plotting each coefficient against the chosen values of λ. The ridge trace is a method of showing the non-orthogonality in two dimensions and illustrates one curve per coefficient; it is advised not to plot the trace for more than 10 coefficients at once in order to provide a meaningful and readable visualisation.

The variance of a coefficient is a decreasing function of λ, while the bias is an increasing function of λ. As a result, as λ increases, the coefficient MSE (i.e. variance and squared bias) decreases to a minimum and then increases back (Marquardt & Snee, 1975). Figure 4.7 illustrates the relation between lambda, variance and the squared bias. The main goal is to find a value of λ for a set of coefficients whose MSE is smaller than the OLS solution. Even if increasing λ would also increase the residual sum of squares (RSS), we are more interested in finding a 'stable' set of coefficients, which will perform well on new observations. The stability we aim to find implies that the coefficients are not sensitive to small changes in the estimation data. Initially, if the predictors are highly correlated, the coefficients will change rapidly for small values of λ up to a point where they stabilise and start changing insignificantly for larger values of λ. The goal is to find the λ values where the coefficients stabilise; there is a range of such equivalent values from a practical viewpoint, since plotting the prediction standard deviation of new data against λ usually exhibits a flat minimum (Marquardt & Snee, 1975). However, this method has been criticised by many researchers for not providing an objective basis of determining λ.

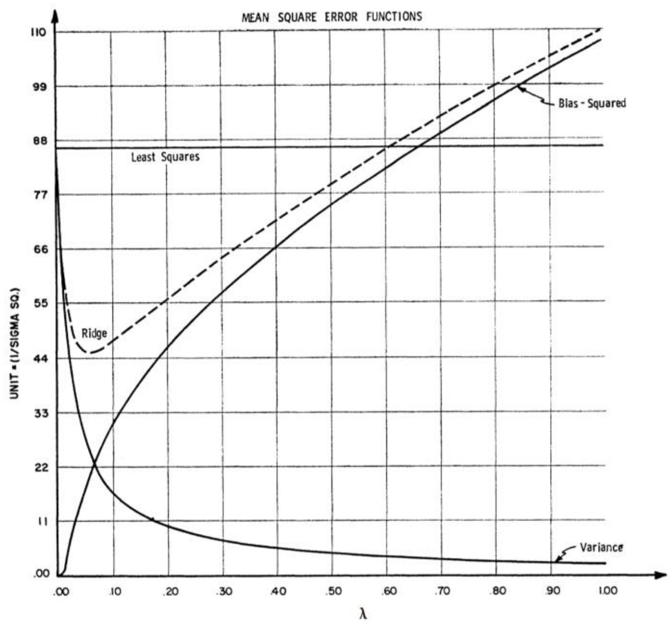

Figure 4.7: The relation between lambda, variance and the squared bias (Hoerl & Kennard, 1970).

4.6.2 Modelling Approach

The target date for the cross-market holidays for a particular stock consists of the days when a particular stock is trading (i.e. its exchange country is on a regular business day) and at least one external market (i.e. the US market or one or more European markets) is shut. The variety of target dates for each stock are multiple observations of the cross-market holiday effect and are aggregated into the regression design matrix. The target variable in the regression models is the relative volume and not the raw volume.

Based on the findings on volume autoregression from the previous study, we fit the models along with 20 lagged volumes in order to assess whether autoregression improves the volume model in the context of cross-market holidays. Since various stocks with different volume magnitudes are plugged into each regression model, the raw lagged volumes are divided by the median of the benchmark volumes in the same manner as the target relative volume, in order to get the normalised lagged volumes. The reasoning behind the normalisation is that stocks are assumed to be different and their lagged volumes need to be normalised in order to account for any differences in magnitude.

The cross-market analyses have two main modelling versions, either as a country-specific regression model, where we fit a separate regression model for each trading country, allowing us to compute the country susceptibility to cross-market holidays, or as a pan-European regression model, where we fit all of the observations across the European countries in a unified regression model and supply additional indicator variables for the trading country of each stock.

For each stock, we know it is traded in a particular European country and is small-/mid-/large-cap. There are two stock-specific predictor variables, namely the trading country and the stock capitalisation.

The pan-European model allows for the identification of small clusters of countries, e.g. the regional effect of US on the whole Europe, the effect of UK on the mainland Europe etc. In the pan-European model, we take some variability away off a country's holiday onto the other individual countries (from the country-specific model). The effect of any country onto the region as a whole is relatively constant. The pan-European model could be considered a reduced model assuming a constant effect, although any particular country might be more or less susceptible to holiday effects from other markets. Despite being interesting for measuring each country's holiday susceptibility, the country-specific model does not provide any insights on clustering.

10-fold stratified cross-validation is applied throughout the analyses of this study, creating random subsamples having roughly equal sizes and roughly the same proportions of observation classes. The class of each observations is defined by encoding each observation's indicator (i.e. binary) variables into a class; for example, an observation in the pan-European cross-market holiday effect model would be encoded by concatenating all the values of the indicator variables (e.g. trading countries, holiday countries, and small-/mid-/large-capitalisation flags) into a binary string. This ensures the model is always trained and tested using subsamples that contain observations from all the classes; furthermore, the representative classes of unbalanced data sets are evenly distributed among the folds. The regression feature matrices are highly sparse and there can be certain levels that are not represented in one of the cross-validation folds. The cross-validation process returns the 10-fold cross-validation estimate of the MSE. The folds are determined before performing the main analysis and they are constant across the various analyses in order to ensure consistency across results.

We used ridge regression (L_2 regularisation), instead of fitting a linear model and subsequently performing forward feature selection. Employing this shrinkage method was motivated by the fact that it deals well with outliers and collinearity, unlike multiple linear regression, which struggles with numerical instability issues, i.e. the design matrix singularity.

The methodology for identifying the value of the ridge parameter λ consists of a two-section search, i.e. grid search, followed by the bisection method. Since too little regularisation causes numerical instability, it is important to find the minimum value of the shrinkage parameter where the regression matrix becomes non-singular. In the first stage, the grid search iterates 21 possible values of λ in log-space, ranging from -10 to 10 (with a step size of 1). For each of this values, the grid search fits a ridge regression and computes the average cross-validation MSE. Eventually, grid search returns the logarithmic value of λ that minimises the cross-validation MSE. The second stage performs the bisection method for the adjacent logarithmic values of the λ identified by the grid search. Therefore, we use the previous and the next values relative to the identified λ and start performing the bisection method, where the grid search optimal λ value is the initial midpoint. We iteratively bisect the interval, by performing cross-validation for the two given end-points of the interval. The interval midpoint successively substitutes the endpoint whose cross-validation MSE is the largest. The process continues unless any of the following criteria fails:

- Minimum delta (i.e. lambda relative change): 0.1;
- Minimum error relative change: 10^{-11};
- Maximum number of iterations: 20.

This two-section search is illustrated in Figure 4.8, where the function between log-space ridge parameters and the cross-validation MSE exhibits a convex interval. Within this convex function, the red circle markers indicate the minimum cross-validation MSE among the search values for λ. The figure contains six illustrative examples of the identification of the regularisation parameter for six country-specific models.

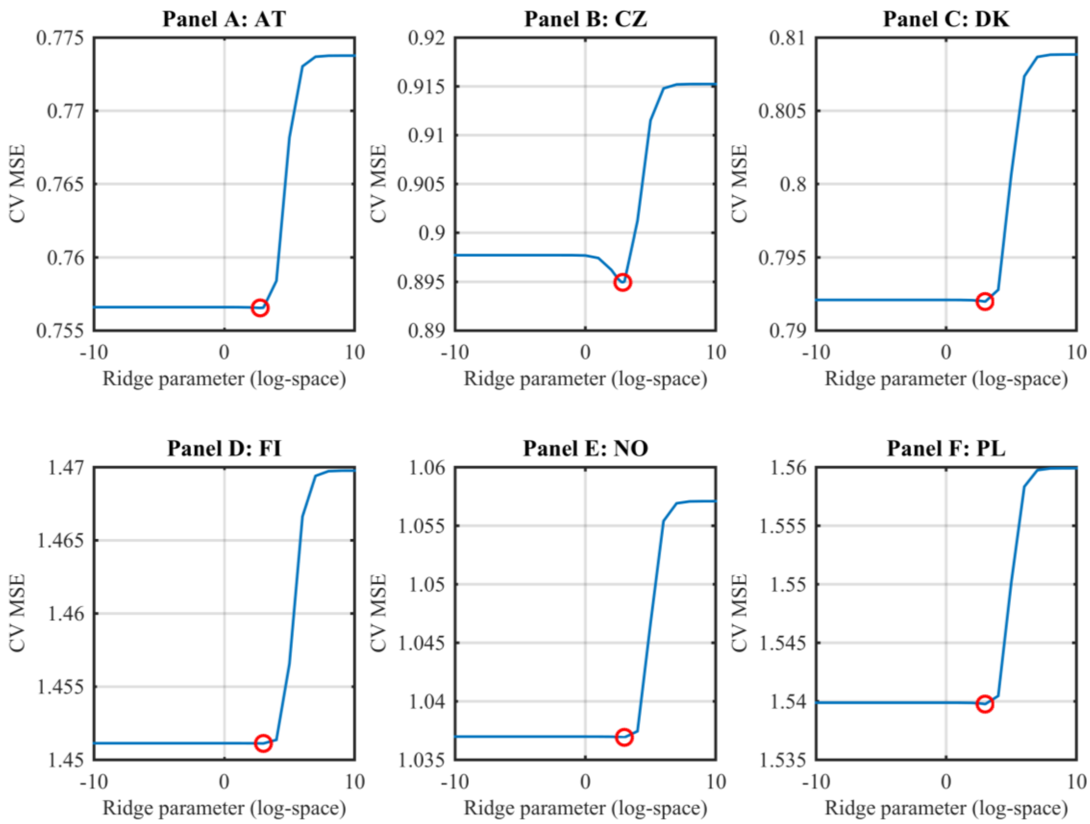

Figure 4.8: Shrinkage parameter vs. cross-validation MSE.

If the independent variables have different scales, then the shrinking is not fair. That is because the penalised term is the sum of squares of all the coefficients, and therefore the predictors will have different contributions to the penalised terms. Hence, the ridge regressions in this study are fit without a constant term and the predictors are centred and scaled to have zero means and unit variance. Finally, the coefficients are restored to the scale of the original data and the constant term is estimated by the sample mean of the response variables.

Ridge regression imposes a penalty on the size of coefficients, shrinking them toward zero and toward each other. Figure 4.9 illustrates the ridge trace, i.e. the shrinkage of the coefficients as a function of λ. For visualisation considerations, we constrained the number of coefficients to 11, including the intercept (or constant term). Panel A shows a large-scale view of the coefficients; a few of them have extremely large values reaching approximately -269 and +269 when lambda is extremely small (i.e. $\lambda = 10^{-10}$). Panel B exhibits the coefficient stabilisation as λ gets close to -6 in log-space.

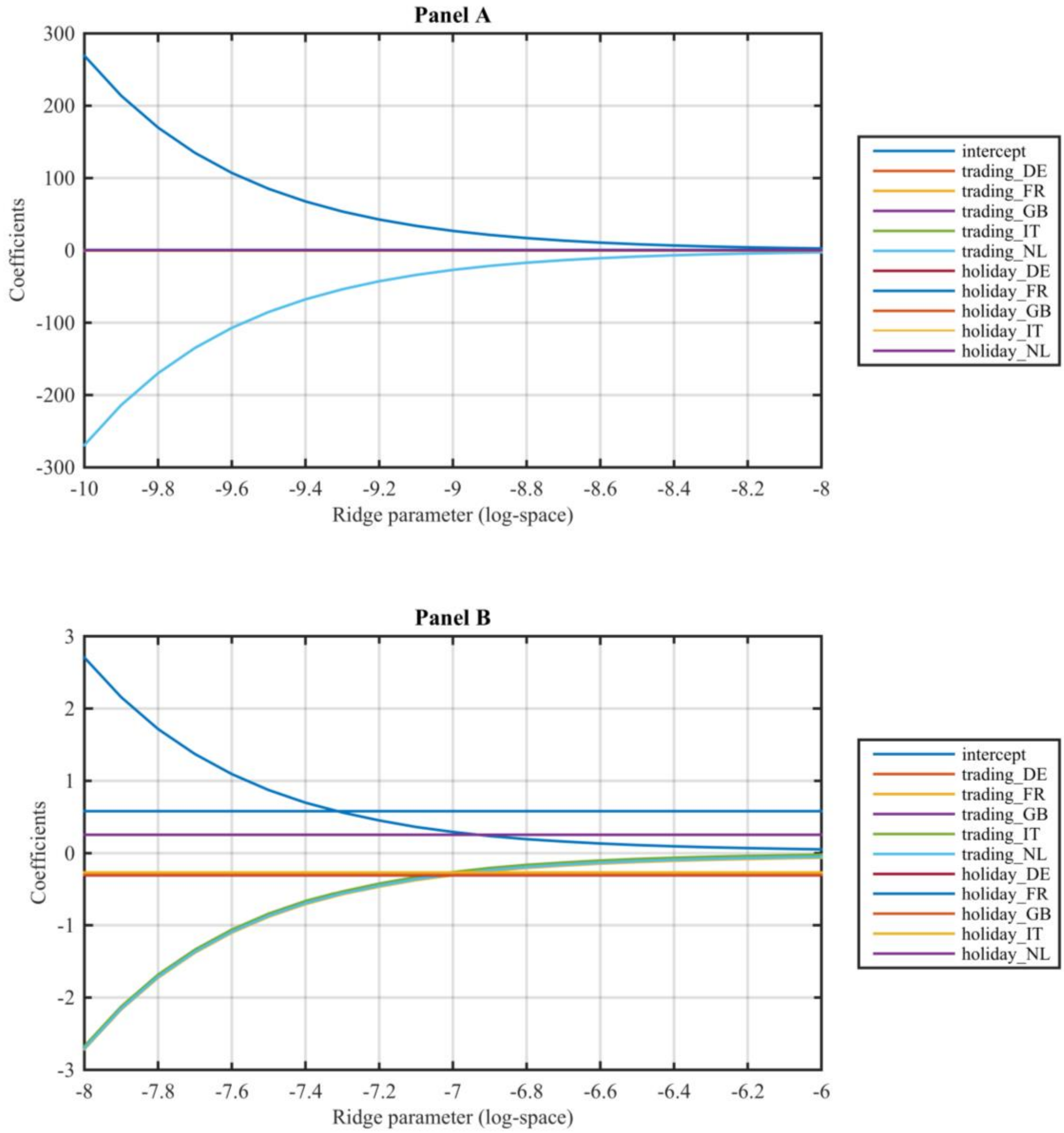

Figure 4.9: Ridge trace: Shrinkage parameter vs. coefficients.

4.6.3 Models Outline

Twelve model variations are fit in this study and their feature sets are outlined in Table 4.10. The words in italics in the feature names on the left-hand side column are generic names and multiple features would exist based on this template feature name depending on the data sample. For example, ‘Trading *country code*’ would be substituted by ‘Trading GB’, ‘Trading DE’, ‘Trading FR’ etc. Moreover, the predictor ’20 lagged normalised volumes’ represents 20 distinct features, and similarly the predictor ‘*Small-/Mid-/Large*-cap’ corresponds to three features.

Table 4.10
Regression models – feature sets.

	Holiday country								Holiday breakdown			
	Country-specific				Pan-European				Country-specific		Pan-European	
Intercept	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Trading *country code*					✓	✓	✓	✓			✓	✓
Holiday *country code*	✓	✓	✓	✓	✓	✓	✓	✓				
Holiday *name*									✓	✓	✓	✓
20 lagged normalised volumes	✓	✓			✓	✓			✓		✓	
Small-/Mid-/Large-cap	✓		✓		✓		✓					

4.6.4 Holiday Country and Holiday Breakdown Models

The first class of models refers to the 'holiday country' models, where the holiday calendar and holiday features correspond to each of the countries in our data set. The aim of this model class is to determine the cross-market holiday effect of each non-trading market and therefore it has 22 'holiday *country*' predictors.

Unlike the 'holiday country' models, the second class of models treats the US and pan-European holidays as a globally unified feature set, and it is called 'holiday breakdown'. Here, we are interested in finding the cross-market holiday effect of individual holidays. These models include 95 distinct 'holiday *name*' predictors, reflecting the pan-European normalised bank holidays occurring during the study period of over 15 years. These are either periodic holidays, typically occurring on an annual basis, or non-periodic holidays, which are one-off events such as UK's Royal Wedding and Diamond Jubilee, or USA's National Day of Mourning for President Gerald R. Ford and National Day of Mourning for President Ronald Reagan.

Each of the two model classes is further split into two model types: country-specific and pan-European.

For the 'country-specific' models, a model is fit for each of the 21 trading markets and the aim is to identify country-to-country holiday effects, while allowing for a better interpretation of a country's susceptibility to the cross-market holiday effect. Because of the highly sparse data, this model is trained on the entire stock universe for a given country, since a stock-specific model would be impractical. For each country-specific regression model, we compute the country susceptibility to cross-market holidays by averaging the non-zero regression coefficients of the 'holiday *country code*' predictors. Using this regression model, one can quantify the magnitude of the cross-market holiday effect among various clusters. For instance, the Swedish holidays' effect is different in magnitude on the UK trading volume than the US holidays' effect on the UK trading volume.

Unlike the 'country-specific' models, only one unified pan-European model is fit in the 'pan-European' model types. The proposed pan-European models are fit similarly to the country-specific models, and, unlike the country-specific models, they include all the trading countries in one single model, having additional 'trading *country code*' indicator variables for each stock's exchange country. Here, the focus is on the regional effect of each non-trading country and we rank the cross-market holiday effect strength of a holiday country within a pan-European context based on the regression coefficients.

We report a concordant negative correlation between trading volume and cross-market holidays. The effect sizes are outlined in this section.

4.6.4.1 Holiday Country Models

The regression coefficients for the cross-market holiday model are summarised in Table 4.11 for the country-specific model and in Table 4.12 for the pan-European model; the predictors whose coefficients are included in these tables are a subset of the entire feature set and only the relevant variables have been included due to space constraints. The pan-European model is fit for a shrinkage parameter whose log-value is 2.96387, and has 1,343,636 observations and a CV MSE of 0.94049.

The results in Table 4.11 raise concerns regarding the Monday effect driving the 'Holiday *country code*' coefficients. This motivates the Monday bank holiday randomisation test in order to identify which effect is driving Monday volumes. It is important to mention that the UK holidays typically fall on a Monday. The randomisation test indicated that the volumes on Monday cross-market holidays are significantly lower than the volumes on regular Mondays, with no cross-market holidays. Therefore, we argue that the cross-market holidays are the main driver of lower Monday volumes and we dispute the role of the weekend effect with regard to lower Monday volumes.

The countries with the highest susceptibility to the cross-market holiday effect are Belgium, Spain, France, Hungary, Netherlands, Portugal, and South Africa, as indicated by the relatively high negative coefficients in Table 4.11.

Table 4.11
Country-specific holiday country model.

Trading Country	Observations	CV MSE	Log Shrinkage Parameter	Intercept	Holiday DE	Holiday FR	Holiday GB	Holiday US	Country Susceptibility
AT	16,746	0.67252	2.75	0.00	-0.18	-0.39	-0.39	-0.20	-0.06
BE	39,984	0.69772	2	0.12	0.01	-0.25	-0.25	-0.27	-0.14
CH	62,070	0.60541	2.5	0.03	-0.18	0.22	0.22	-0.23	-0.05
CZ	2,629	0.78933	2	-0.11	0.03	-0.17	-0.17	-0.02	-0.03
DE	102,500	0.57447	2	0.11		-0.35	-0.35	-0.29	-0.12
DK	25,861	0.68383	2.75	0.02	0.01	-0.31	-0.31	-0.23	-0.05
ES	34,052	0.32781	2	0.05	-0.11	-0.25	-0.25	-0.25	-0.09
FI	70,830	1.32237	2.75	0.02	-0.10	0.12	0.12	-0.20	-0.06
FR	219,181	1.33236	2.942382813	0.06	0.03			-0.14	-0.10
GB	364,239	1.07827	2.75	0.06	-0.49	0.68	0.68	-0.26	-0.05
GR	36,762	0.91295	2	-0.05	0.06	0.21	0.21	-0.04	-0.02
HU	2,743	0.41961	2	-0.04	-0.44	-0.61	-0.61	-0.36	-0.09
IE	18,619	1.99832	2	0.06	-0.79	0.29	0.29	-0.31	-0.12
IT	62,361	0.37758	2	0.03	-0.09	-0.04	-0.04	-0.18	-0.07
NL	30,980	0.38703	2	0.13	-0.39	0.44	0.44	-0.40	-0.08
NO	31,060	0.94061	2.75	-0.03	0.09	0.41	0.41	-0.19	0.02
PL	28,956	1.39764	2.75	-0.02	-0.32	0.04	0.04	-0.20	-0.03
PT	10,319	0.70198	2	0.06	-0.07	0.15	0.15	-0.28	-0.10
SE	83,880	0.69879	2	0.02	0.01	0.08	0.08	-0.19	-0.06
TR	74,946	0.57264	1	-0.05	-0.16	-0.57	-0.57	-0.06	-0.01
ZA	24,918	0.49736	2	0.07	-0.44	-0.05	-0.05	-0.30	-0.11

The 'trading *country code*' coefficients of the pan-European model are low, close to zero. However, the 'holiday *country code*' predictors, such as Germany, UK, Italy, and the USA, exert a strong impact on the trading activity. There are a couple of unexplained positive coefficients for holiday countries such as France and Netherlands.

Table 4.12
Pan-European cross-market holiday model (selected trading and holiday countries exhibited).

Predictor	Coefficient	Predictor	Coefficient	Predictor	Coefficient	Predictor	Coefficient
intercept	0.04	trading_GR	0.04	holiday_AT	-0.04	holiday_HU	-0.09
trading_AT	-0.01	trading_HU	-0.06	holiday_BE	-0.04	holiday_IE	-0.03
trading_BE	-0.03	trading_IE	-0.07	holiday_CH	-0.09	holiday_IT	-0.23
trading_CH	-0.01	trading_IT	-0.01	holiday_CZ	-0.10	holiday_NL	0.22
trading_CZ	-0.04	trading_NL	-0.02	holiday_DE	-0.28	holiday_NO	-0.16
trading_DE	-0.01	trading_NO	0.00	holiday_DK	-0.06	holiday_PL	-0.04
trading_DK	0.01	trading_PL	0.00	holiday_ES	-0.03	holiday_PT	0.03
trading_ES	-0.02	trading_PT	-0.03	holiday_FI	0.04	holiday_SE	-0.03
trading_FI	0.01	trading_SE	0.00	holiday_FR	0.51	holiday_TR	-0.02
trading_FR	-0.01	trading_TR	0.11	holiday_GB	-0.26	holiday_US	-0.21
trading_GB	0.00	trading_ZA	-0.03	holiday_GR	-0.02	holiday_ZA	0.00

Figure 4.10 shows the distribution of relative volume for the two holiday countries whose coefficients are positive (i.e. France and Netherlands) and for two other countries (i.e. UK and USA) exerting a clear subduing effect on the other pan-European trading countries, along with the distribution of all the pan-European stocks' relative volume on any cross-market holiday. The relative volumes on French and Dutch holidays are still slightly positively skewed and we cannot conclude that these countries have a positive impact on the other markets' trading volume.

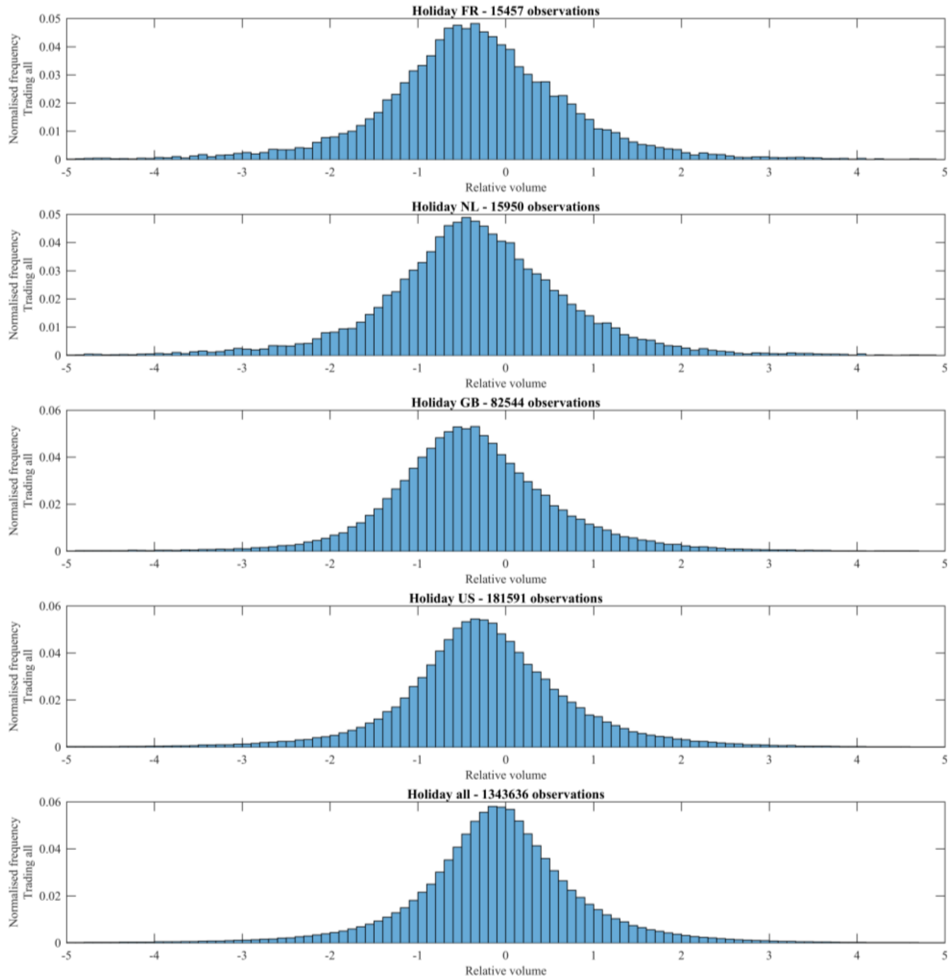

Figure 4.10: Relative volume distribution for the pan-European stocks trading on cross-market holidays occurring in France, Netherlands, United Kingdom, USA, and, eventually, in any country

4.6.4.2 Holiday Breakdown Models

The holiday breakdown models present a set of salient holidays that tend to generally drive the regional trading volume lower. There are two models fit for this model class: the country-specific model, whose results are shown in Table 4.13 for the regional/global holidays, and in Table 4.14 for the significant country-specific holidays, and the pan-European model, outlined in Table 4.15 for the regional/global holidays and in Table 4.16 for the country-specific holidays. These tables contain a small subset of relevant holidays (out of the 95 normalised holidays), excluding other features, such as the lagged normalised volumes.

Table 4.13
Country-specific holiday breakdown model – selected regional/global holiday features.

Country	Samples	CV MSE	Log Shrinkage Parameter	Intercept	New Year's Day	Boxing Day	Christmas Day	Good Friday	Easter Monday	May 1st	New Year's Eve	Country Susceptibility
AT	16,746	0.67	2.94434	-0.03	-0.13	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.03
BE	39,984	0.67	1.91064	0.11	-0.19	-0.02	-0.02	-0.02	-0.02	-0.02	-0.97	-0.21
CH	62,070	0.59	2.00000	0.08	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.15
CZ	2,629	0.78	2.84717	-0.15	0.10	-0.01	-0.01	-0.52	-0.01	-0.01	-1.17	-0.02
DE	102,500	0.57	2.00000	0.08	-0.15	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.13
DK	25,861	0.68	2.00000	0.06	0.09	0.00	0.00	0.00	0.00	-0.65	0.00	-0.08
ES	34,052	0.32	1.93750	0.08	-0.33	-0.01	-0.01	-0.01	-0.01	-0.01	-0.75	-0.16
FI	70,830	1.32	2.95752	0.01	-0.11	0.00	0.00	0.00	0.00	0.00	0.00	-0.04
FR	219,181	1.31	2.00000	0.09	-0.11	-0.02	-0.02	-0.02	-0.02	-0.02	-0.81	-0.17
GB	364,239	1.06	2.00000	0.05	-0.01	-0.01	-0.01	-0.01	-0.01	-0.37	-1.27	-0.12
GR	36,762	0.90	2.00000	0.10	-0.25	-0.01	-0.01	-0.25	-0.49	-0.01	-0.56	-0.12
HU	2,743	0.41	2.00000	-0.06	0.14	-0.01	-0.01	-1.90	-0.01	-0.01	-0.34	-0.06
IE	18,619	1.96	2.75000	-0.08	-0.02	-0.02	-0.02	-0.02	-0.02	-1.03	-1.34	-0.10
IT	62,361	0.37	2.00000	0.04	-0.17	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.11
NL	30,980	0.38	2.00000	0.06	0.05	-0.03	-0.03	-0.03	-0.03	-0.33	-1.10	-0.14
NO	31,060	0.93	2.87500	0.00	-0.15	0.00	0.00	0.00	0.00	0.00	0.00	-0.06
PL	28,956	1.38	2.99561	-0.03	-0.15	-0.01	-0.01	-0.01	-0.01	-0.01	-0.20	-0.07
PT	10,319	0.69	2.50000	-0.06	-0.18	-0.01	-0.01	-0.01	-0.01	-0.01	-0.43	-0.06
SE	83,880	0.68	2.50000	0.01	-0.07	0.00	0.00	0.00	0.00	0.00	0.00	-0.06
TR	74,946	0.55	2.00000	0.04	1.64	-0.16	-0.23	-0.30	-0.20	-0.26	-0.16	-0.06
ZA	24,918	0.47	2.00000	0.00	-0.39	-0.01	-0.01	-0.01	-0.01	-0.01	-1.20	-0.11

Table 4.14
Country-specific holiday breakdown model – selected country-specific holiday features.

Country	GB Spring Bank Holiday, US Memorial Day	GB IE Early May Bank Holiday	DE Day of German Unity	FR Bastille Day	GB Summer Bank Holiday	GB The Queen's Diamond Jubilee	US Independence Day	US Labor Day	US Martin Luther King Day	US Presidents Day / Washington's Birthday
AT	-0.41	-0.21	-0.10	-0.36	-0.28	-0.44	-0.25	-0.21	-0.17	-0.20
BE	-0.47	-0.29	-0.12	-0.23	-0.39	-0.28	-0.46	-0.23	-0.17	-0.31
CH	-0.62	-0.37	-0.19	0.17	-0.46	-0.19	-0.37	-0.33	-0.09	-0.40
CZ	-0.32	0.05	0.33	-0.01	-0.35	-0.05	-0.02	-0.25	0.11	-0.14
DE	-0.55	-0.30	-0.01	-0.02	-0.49	-0.20	-0.38	-0.31	-0.20	-0.32
DK	-0.54	-0.18	0.19	0.04	-0.34	0.00	-0.39	-0.24	-0.21	-0.38
ES	-0.56	-0.39	-0.09	-0.15	-0.53	-0.29	-0.40	-0.29	-0.27	-0.36
FI	-0.40	-0.17	-0.17	0.22	-0.30	-0.26	-0.45	-0.11	-0.16	-0.22
FR	-0.39	-0.24	-0.21	-0.02	-0.27	-0.29	-0.39	-0.02	-0.03	-0.27
GB	-0.20	-0.01	-0.05	0.03	-0.01	-0.01	-0.33	-0.28	-0.11	-0.36
GR	-0.31	-0.09	-0.29	-0.05	-0.22	0.13	-0.40	-0.23	-0.07	-0.35
HU	-0.66	-0.25	-0.32	-0.28	-0.65	0.04	-0.47	-0.30	-0.59	-0.46
IE	-1.20	-0.02	0.01	-0.25	-1.00	-0.75	-0.20	-0.26	-0.07	-0.29
IT	-0.42	-0.25	-0.20	0.08	-0.38	-0.40	-0.36	-0.19	-0.11	-0.22
NL	-0.74	-0.40	-0.06	-0.06	-0.54	-0.42	-0.51	-0.32	-0.23	-0.37
NO	-0.52	-0.22	-0.12	0.38	-0.34	-0.26	-0.38	-0.15	-0.14	-0.29
PL	-0.30	-0.12	-0.45	-0.31	-0.32	-0.30	-0.24	-0.21	-0.10	-0.18
PT	-0.37	-0.29	-0.08	0.40	-0.36	-0.29	-0.27	-0.12	-0.04	-0.27
SE	-0.47	-0.14	-0.02	0.09	-0.26	-0.47	-0.37	-0.18	-0.13	-0.16
TR	-0.17	-0.16	-0.68	-0.68	-0.32	0.29	-0.12	-0.23	-0.12	-0.02
ZA	-0.56	-0.37	-0.15	-0.29	-0.55	-0.29	-0.42	-0.35	-0.06	-0.35

The 1st May bank holiday is observed in most of Europe's markets and it is very salient in the few countries that are trading on 1st May: Denmark (-0.65), UK (-0.37), Ireland (-1.03), Netherlands (-0.33), and Turkey (-0.26). Turkey is the only country trading on the Christmas Day (-0.23) and Boxing Day (-0.16), and therefore these Christian holidays drive the volume lower as the rest of Europe is not trading. Similarly, the Catholic Easter holidays (i.e. Good Friday and Easter Monday) have a strong impact in Greece and Turkey. Hungary experiences low volumes caused by Good Friday only. The New Year's Eve has a strong effect (with coefficients close to -1) in Belgium, Czech Republic, Spain, France, UK, Greece, Ireland, Netherlands, and South Africa.

The Early May Bank Holiday is observed in the UK and Ireland, where this bank holiday substitutes the generally observed May 1st across the other European countries. Therefore, the Early May Bank Holiday, along with other country-specific holidays from the UK, Germany, France, and the USA, have a generally strong negative impact on the other European markets. These holidays include one-off holidays (i.e. single occurrence bank holidays issued by the governments for certain reasons), e.g. the Queen's Diamond Jubilee.

Some countries exhibit incredibly low volumes associated with the Christmas Eve, which is certainly caused by the fact that most of the trading markets on the Christmas Eve have a half-trading day schedule and therefore the volume is significantly lower. These countries include Belgium (-1.52), Spain (-1.40), UK (-1.92), Ireland (-3.26), Netherlands (-1.80), Poland (-2.29), Portugal (-1.70), and South Africa (-1.95).

Table 4.15 shows the significant regional/global holidays for the pan-European model, which has 1,343,636 observations, CV MSE 0.93182, and shrinkage parameter 2 (log-space). The 'trading *countryCode*' coefficients are very low (close to zero). Some notable regional/global cross-market holidays include Boxing Day (both the main day and the additional day), Christmas Day, Good Friday, Easter Monday, May 1st, New Year's Eve (both the main day and the additional day), Christmas Eve, Whit Monday, Ascension Day, Assumption of Mary, All Saints' Day and the additional day for the New Year's Day. The additional days are issued by certain countries, especially as 'bridge holidays' (i.e. when the main holiday falls on a Thursday or on a Tuesday, and the governments transform the in-between Friday or Monday into a bank holiday in order to have a 4-day break, including the weekend).

Table 4.15
Pan-European cross-market holiday breakdown model – full candidate feature set (selected regional/global holiday features exhibited).

Holiday	Coefficient	Holiday	Coefficient
Boxing Day	-0.33	Whit Monday	-0.39
Christmas Day	-0.42	Ascension Day	-0.43
Good Friday	-0.42	Assumption of Mary	-0.26
Easter Monday	-0.34	All Saints' Day	-0.22
May 1st	-0.37	Boxing Day (Additional Day)	-0.40
New Year's Eve	-0.94	New Year's Day (Additional Day)	-0.21
Christmas Eve	-1.51	New Year's Eve (Additional Day)	-0.43

Table 4.16 shows country-specific holidays that exhibit a strong impact to the pan-European trading volume. Unlike the regional/global holidays outlined in Table 4.15, which are usually bank holidays in most of the European countries, the magnitude of the country-specific holidays is incredibly high since these are official bank holidays in only one or two countries, while the other markets are trading on these days. Important country-specific holidays are identified mainly from the UK (e.g. Early May Bank Holiday, Spring Bank Holiday, or Summer Bank Holiday) and the USA (e.g. Memorial Day, Independence Day, Labor Day, or Presidents Day), including some one-off holidays with conspicuous effects on the trading volume, e.g. Golden Jubilee, Diamond Jubilee, Royal Wedding, Hurricane Sandy, and National Day of Mourning for President Ronald Reagan.

Table 4.16
Pan-European cross-market holiday breakdown model – significant country-specific holidays extracted from the full candidate feature set.

Holiday	Coefficient	Holiday	Coefficient
(GB, US) Spring Bank Holiday, Memorial Day	-0.45	(PL) Poland Independence Day	-0.23
(GB, IE) Early May Bank Holiday	-0.24	(PT) Portugal Day	-0.27
(DE) Day of German Unity	-0.17	(US) Independence Day	-0.36
(GB) Golden Jubilee Bank Holiday	-0.33	(US) Labor Day	-0.22
(GB) Royal Wedding Bank Holiday	-0.21	(US) Markets Closed (Hurricane Sandy)	-0.24
(GB) Summer Bank Holiday	-0.38	(US) National Day of Mourning for President Ronald Reagan	-0.33
(GB) The Queen's Diamond Jubilee	-0.25	(US) Presidents Day (Washington's Birthday)	-0.29

4.6.5 Volume Autoregression

All of these models are fit with and without 20 lagged normalised volumes in order to determine whether the volume autoregression is improving the volume prediction. Strong volume autoregression is observed across the model variants. Fitting the models with 20 lagged normalised volumes considerably outperforms the models without lagged volumes. The lower MSE achieved by the models trained with lagged volume is outlined in Table 4.17, where we show the cross-validation MSE for the two pan-European models, fit with and without lagged volumes. These 1-step ahead models are fit for the entire sample period of the study.

Table 4.17
Pan-European models – comparison of the presence and absence of the lagged volumes.

Model	Lagged volumes	Observations	CV MSE	Shrinkage parameter (in logarithmic space)
Pan-European holiday country	Yes	1,343,636	0.94049	2.96386
	No	1,343,636	1.04997	3
Pan-European holiday breakdown	Yes	1,343,636	0.93182	2
	No	1,343,636	1.03797	2

4.6.6 Market Capitalisation

The 'holiday country' models provide a market capitalisation variant, where we further investigate a potential relationship between the cross-market holiday effect and the market capitalisation of the stocks. Using only the constituents of the FTSE, DAX and CAC market capitalisation-stratified indices, we can discriminate between small-, mid-, and large-cap stocks.

Figure 4.11 illustrates the relative trading volume for the three capitalisation indices in each column panel: UK's FTSE in Panel A, Germany's DAX in Panel B, and France's CAC in Panel C. Each of the rows corresponds to a market capitalisation class, with large-cap in the top row, mid-cap in the middle row, and small-cap in the bottom row.

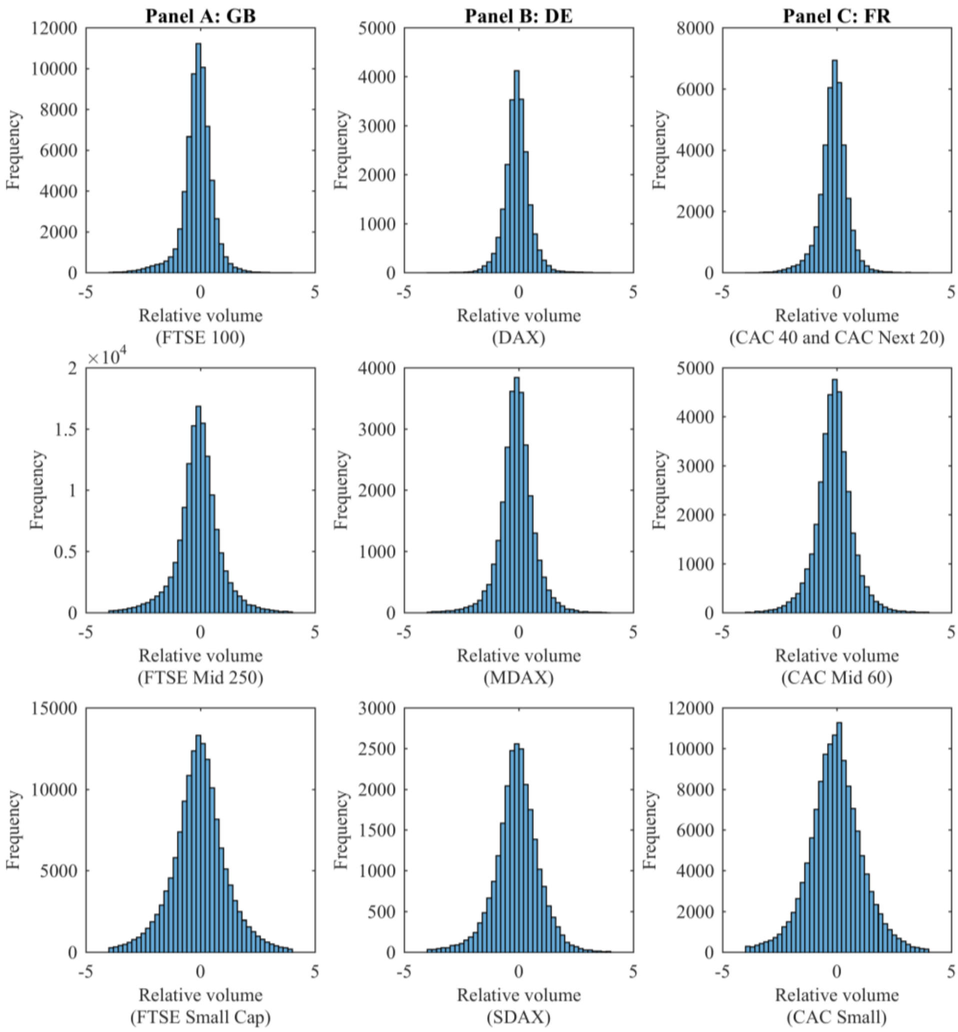

Figure 4.11: Relative volume distribution for the individual market capitalisation-stratified stocks.

The cross-market holiday models have been enhanced with three indicator variables for the market capitalisation class. The country-specific and pan-European models constantly underperform when the market capitalisation is included. For example, consider the two variants of the pan-European holiday country model in Table 4.18. The CV MSE grows from 1.05 to 1.22 when the market capitalisation indicator variables are added to the model, and only the market capitalisation-stratified observations are considered. The coefficients for the small-/mid-/large-cap indicator variables are provided for the three countries, along with the cross-market holiday effect susceptibility for FTSE, DAX, and CAC. FTSE is the less susceptible index among these.

Table 4.18
Country-specific cross-market holiday model with market capitalisation – reduced feature set (selected market capitalisation features exhibited).

Model	Country	Market cap indicators	Samples	CV MSE	Small-cap coefficient	Mid-cap coefficient	Large-cap coefficient	Index susceptibility
Pan-European holiday country	-	No	1,343,636	1.04997	-	-	-	-
	-	Yes	650,957	1.22052	0.00431	0.00172	-0.00916	-
Country-specific holiday country	UK	Yes	361,510	1.15396	0.01452	0.00051	-0.02464	-0.04966
	Germany	Yes	74,679	0.63508	-0.01422	-0.00379	0.01947	-0.13188
	France	Yes	214,768	1.51558	-0.00698	0.01075	0.00047	-0.10024

4.6.7 Multi-Step Ahead Prediction

Besides the default one-step ahead models introduced so far (i.e. $n = 1$), additional n-step ahead analyses are conducted for step sizes ranging from 2 days to 6 days for a subset of models, based on the findings from the one-step ahead models, i.e. the model variations whose CV MSE is minimal. The motivation for a multi-step ahead prediction stems from the real-world scenario where traders could plan their portfolios by predicting a cross-market holiday effect on a stock's trading volume given its exchange country. If we want to predict a cross-market holiday effect n days in advance, then we compute the median of the benchmark volumes between $(t - n)$ and $(t - 20 - n)$ and compare it against the volume on the cross-market holiday (i.e. V_0) in order to train the model.

The models exhibit a constant trend of increasing the MSE between the one-step ahead model, and the multi-step ahead models, where the largest step becomes 6 days (i.e. 6-step ahead model). However, the models and their coefficients perform rather similarly. The cross-validation MSE is directly proportional with the step ahead size, as outlined in Table 4.19. The error increases progressively once the step size is increased due to the lack of more recent data.

Table 4.19
Comparison of MSE between the 1-step ahead model and multi-step ahead models.

Model	Observations	Cross-validation MSE					
		1-step ahead	2-step ahead	3-step ahead	4-step ahead	5-step ahead	6-step ahead
Pan-European holiday country	1,343,636	0.94049	1.00390	1.03766	1.06146	1.07983	1.09700
Pan-European holiday breakdown	1,343,636	0.93182	0.99275	1.02501	1.04818	1.06608	1.08266

4.7 Discussion

This study investigates the anecdotal evidence of lower volumes associated with external markets not trading. This phenomenon was described as the 'cross-market holiday effect' in this study, where we examine it in the European equity markets using a comprehensive pan-European stock universe. It is the first study to investigate the cross-market holiday effect within Europe and it probably has the largest data set employed by any study on the European equity markets and the most accurate European and US trading calendar spanning almost 16 years. As far as we are aware, there are only two studies on the cross-market holiday effect (Cheung & Kwan, 1992) (Casado, et al., 2013), despite its popularity among finance professionals. The study proposes a novel methodology, consisting of ridge regression applied to finance time series. This is complemented by the initial randomisation tests, which provide rigour to our investigation of the phenomenon existence.

Throughout the in-sample analyses, we report compelling evidence of volume autoregression. The empirical results strongly support the existence of a negative cross-market holiday effect in the European markets. The relative trading volume is significantly lower on cross-market holidays. On average, the volume is reduced by 8.5% compared to the volume of the benchmark period. We investigated whether these results are caused by the fact that most of the holidays fall on a Monday in the UK (i.e. Europe's largest market) and it could possibly be the Monday effect driving down the volumes. The results of the randomisation test confirm that the lower trading activity is associated with the cross-market holidays. We do not debate whether the Monday effect itself exists (observed as a day-of-the-week effect and not as a Monday bank holiday effect), but we provide evidence that this study's lower volumes on Mondays having at least one regional cross-market holiday are caused by the cross-market holidays. This provides a recommendation for other researchers to take this study further and investigate the Monday effect on the European liquidity.

Based on the precise trading calendar of this study for the European countries and the USA, we observe some strong country susceptibility levels for a few countries (e.g. Belgium, Spain, France, Hungary, Netherlands, Portugal, and South Africa). There are strong cross-market holiday effects originating from large markets (e.g. the USA, the UK, Germany, or Italy), which tend to have a salient effect (in the form of negative coefficients, meaning a subduing effect on the trading volume) across all the other countries, although French holidays exhibit a reverse pan-European cross-market holiday effect, resulting in a regional trading volume increase. Our findings corroborate the results of Casado et al. (2013).

The study presents a number of interesting holidays that exert a strong influence on the liquidity of the European markets. After normalising the US and the pan-European trading calendar, we find that certain periodic (e.g. New Year's Eve, Christmas Day, Boxing Day, May 1st, Easter Monday, Good Friday etc.) and non-periodic (e.g. UK's Golden Jubilee, the Queen's Diamond Jubilee, the Royal Wedding etc.) bank holidays have a blatant effect on volumes. Strong lower volumes are observed on 1st May in Denmark, UK, Ireland, Netherlands and Turkey. The Christmas holidays (i.e. Christmas Day and Boxing Day) have a significant impact on the Turkey market, when the rest of Europe is not trading. Similarly, catholic Easter holidays (i.e. Good Friday and Easter Monday) cause a volume drop in Greece and Turkey. The most noticeable holidays that are subduing the trading volume in the pan-European markets originate from the UK (e.g. Early May Bank Holiday, Spring Bank Holiday, or Summer Bank Holiday), followed by the USA (e.g. Memorial Day, Independence Day, Labor Day, or Presidents Day), whose holidays have a slightly lower intensity on the European volume than the UK. A few other country-specific holidays are reported to affect the volume, e.g. Day of German Unity, Poland Independence Day, and Portugal Day.

The model accounts for a potentially differentiated effect by market capitalisation, but the cross-market holiday effect persists across small-, mid-, and large-cap stocks. We also find a structural break around the financial crisis of 2007-08 for the market capitalisation-based impact of cross-market holidays, where the effect reversed for a couple of indices.

This prediction can also be made in advance of the cross-market holidays using multi-step ahead forecasting, based on a stock's past volumes and the volume levels during the previous cross-market holidays originating from the same country. Having an accurate trading calendar and anticipating a cross-market holiday could predict the trading volume in the run-up to the cross-market holiday. The findings of this study propose a framework for traders and hedge fund managers for planning their portfolios in advance, in order to predict their positions and profits during the cross-market holidays by knowing how much more or less the trading volumes are expected to be.

5. Expiry Day Effects on European Trading Volumes

This study investigates the effect of periodic events, such as the stock index futures expiries and the MSCI quarterly index reviews, on the trading volume in the pan-European equity markets. The motivation of this study stems from the anecdotal evidence of increased trading volume in the equity markets during the run-up to the stock index futures expiries and MSCI rebalances. This study investigates this phenomenon in more detail and analyses the trading volumes of seven European stock indices and the MSCI International Pan-Euro Price Index. The analysis features a multi-step ahead volume forecast, which is important for practitioners in order to plan multi-day trades while looking to minimise the market impact. The results confirm higher trading activity during the four days in the ruan-up to the futures expiry day, lasting two days after the futures expiry, and on the MSCI rebalance day and on its previous day. We report a clear futures expiry effect, which accounts for the Friday effect in terms of larger trading volumes. The MSCI rebalance trading volume is significantly different from the volume of the adjacent months with no MSCI reviews, but they cannot explain the end-of-month effect entirely.

5.1 Introduction

This study investigates the increased trading volume associated with repetitive special events, namely the stock index futures expiries and the MSCI quarterly index reviews; this is named 'the expiry day effect' in the literature. We analyse a number of aspects related to these special dates, such as the existence of an anticipatory/subsequent expiry day effect (i.e. whether the volumes are higher during the days leading up to and following the MSCI rebalances and the stock index futures expiries) and the identification of the principal volume drivers of this phenomenon. The study aims to distinguish between the index expiry day effects being studied and well-established calendar effects, such as the Friday effect or the end-of-month effect. We discriminate between the Friday effect and the stock index futures expiries, and between the end-of-month effect and the MSCI quarterly reviews in order to identify the primary drivers of increased trading activity.

The financial markets are typically in a rather steady state, but they start fluctuating when certain events occur, e.g. company annual reports and announcements, news events, or other (periodic) calendar events, such as the subject of this study (i.e. the index expiries and rebalances). We explore activity surges around the stock index futures expiries and MSCI quarterly reviews.

The contribution of this study is threefold: first, the expiry day effect has been scarcely investigated in the literature, and, out of this small proportion, there is an incredibly small number of papers employing data from the European markets; second, the majority of these studies focus on returns, while the volume dimension is mostly ignored; and third, planning multi-day trades is important to practitioners and we propose a multi-step ahead prediction model for the expiry day effect. As far as we are aware, this is the first pan-European study of index expiries or MSCI rebalances, while employing the most recent 15 years of daily market data.

The aim of this study is to provide a trading volume in-sample analysis while considering aspects such as the futures expiries and the MSCI quarterly index reviews. As with the cross-market holidays, the futures expiries are also an instance of a sparse event and we investigate the futures expiries for 7 of the most liquid European stock indices. The study deals with a set of data analysis challenges and investigates a phenomenon whose scope is new. We propose a novel methodological approach in finance by constructing an accurate expiry calendar data set for the most liquid European indices, retrieving the daily market data for the historical constituents for a 15-year period, rigorously testing the existence of the expiry day effect, and using stepwise regression. We inspect anticipatory and subsequent effects of the index expiry and review dates by analysing the previous and following five business days relative to the expiry days.

The study proceeds as follows: section 2 surveys the relevant literature on calendar effects, including the expiry day effect, the Friday effect, and the end-of-month effect, along with a succinct review of the volume-price relation, the stock index futures expiries and MSCI rebalances; section 3 describes the data sample being investigated, including the stock universe and the calendar data; section 4 introduces our analytical approach, and the methodology of this experiment; section 5 tests the existence of the investigated effects by conducting randomisation tests; based on the results of these tests, section 6 introduces and explains the results of the stepwise regression models fit for the futures expiry and MSCI rebalance analyses, while section 7 concludes the study with a discussion on the main findings.

5.2 Background

This section starts with a survey of some of the relevant calendar effects, in order to understand the periodic/seasonal market dynamics that have been empirically identified as potential drivers of volume or price returns. This is followed by a short review of the relation between trading volume and price returns, and an introduction to stock index futures and MSCI index reviews.

5.2.1 Calendar Effects

The majority of the literature on calendar effects looks at the relation between calendar anomalies and price returns, while the relation with the trading volume is barely covered. We start by reviewing the relevant calendar effects and then we outline the empirical findings on the connection between volume and price, in order to infer the calendar effects and their impact on trading volume.

Calendar effects are essentially anomalies in the financial markets that are associated with the calendar seasonality. The literature on calendar effects (and on behavioural finance, in general) is highly contentious and its empirical findings are usually inconclusive. One of the reasons is that each calendar effect is usually investigated in isolation, while a full universe of calendar effects would diminish the effect size of the calendar anomalies (Sullivan, et al., 2001). It is worth mentioning that the calendar effects have always been identified ex post due to their dependence on empirical evidence from the past time series supporting their existence. The dynamics of some calendar effects is also known to change or reverse over time (Dimson & Marsh, 1999) (Schwert, 2003) (Hansen, et al., 2005) (Pearce, 1996), while other calendar effects tend to persist through time, as reported by Lakonishok and Smidt (1988), Barone (1990), Agrawal and Tandon (1994), and Mills and Coutts (1995). The following review of calendar effects outlines some of the event-driven irregularities markets experience.

5.2.1.1 Weekend Effect and the Friday Effect

The weekend effect consists of a negative weekend return, implying that the Friday returns are greater than the returns on the following Monday. This calendar anomaly has been widely studied in the literature by authors such as French (1980), Gibbons and Hess (1981), Jaffe and Westerfield (1985), Pettengill (2003), or Cross (1973). Based on the correlation between price and volume that will be introduced in this section, we investigate the 'Friday effect' in conjunction with the expiry day effect because the stock index futures expiries typically fall on the third Friday of the expiry months, and both effects are associated with increased trading volume.

5.2.1.2 Expiry Day Effect

The expiry day anomaly consists of higher trading volume and abnormal volatility near the close on expiry days (Stoll & Whaley, 1997) (Sukumar & Cimino, 2012) (Chow, et al., 2003) (Sadath & Kamaiah, 2011). This is particularly of interest to this research, as we investigate the trading volume's relationship to the MSCI rebalances and futures expiries. Pope and Yadav (1992) found an immediate increase in trading volume before the options expiry on London Stock Exchange, followed by an immediate decrease after the expiry. Using Indian financial data, Vipul (2005) observed an abnormally high trading volume, which starts to increase on the previous day of the expiry and continues into the next day for stocks with relatively high volume of derivatives. Chakrabarti et al. (2005) investigate the effects of changes in MSCI indices and find that the trading volume increases significantly and remains high after the change date for the stocks added to the index. Furthermore, Chiang (2009) observed trading volume peaks occurring on the third Friday of each month; this effect is driven by the option expiry, since it appears only among optionable stocks, with options expiring on the third Friday of the month.

5.2.1.3 Turn-of-the-Month and End-of-Month Effects

Another popular effect is the turn-of-the-month effect or end-of-month effect, including other similar effects such as the intra-month, the week-of-the-month and the monthly effects. The intra-month effect consists of positive returns in the first half of the month (and more specifically in the early days of the calendar months) only (Ariel, 1987) (Rosenberg, 2004). The turn-of-the-month effect (Cadsby & Ratner, 1992) has been typically defined as the stock price surge on the last day of one month and the first three days of the next month. The four-day turn-of-month period represents 87% of the average monthly return (Kunkel, et al., 2003). A plausible explanation is the standardisation of payments at the turn of the month (Ogden, 1990). Investigating thinly traded Finnish stocks, Nikkinen et al. (2009) found that the release of major US macroeconomic news is driving the turn-of-the-month and intra-month anomalies. Moreover, the higher returns at the turn-of-the-month are associated with a surge in trading volume, which is potentially caused by the buying pressure at the end of the month (Booth, et al., 2001). Strong effects on volume are found in the last trading week of the month in the Finish stock index futures, options, and cash markets (Martikainen, et al., 1995).

5.2.1.4 Month-of-the-Year Effect and January Effect

The January effect (or turn-of-the-year effect) consists of increased stock prices and trading volume in the last week of December and the first half of January (Thaler, 1987) (Ariel, 1987). This effect is also accompanied by a 'size effect', i.e. the negative relation between abnormal returns in January and firm size (Keim, 1983), with large returns in January and exceptionally large returns in the first few trading days of January for small-capitalisation stocks (Reinganum, 1983) (Ritter, 1987). The main causes of this effect are tax-loss trading (Poterba & Weisbenner, 2001) (Jones, et al., 1987) (D'Mello, et al., 2003) and window-dressing (Ng & Wang, 2004). Many authors investigated and confirmed the January effect in the stock markets (Rozeff & Kinney, 1976), with a steady existence over time (Haugen & Jorion, 1996). Other authors found no statistical support for the January effect, e.g. in the post-1987 market crash period (Mehdian & Perry, 2002) despite the existence of the January effect in the pre-crash period; Ritter and Chopra found no January seasonality in the pre-crash period using NYSE securities (Ritter & Chopra, 1989). The research focusing on the January effect and trading volume found a seasonal volume pattern (Constantinides, 1984), and that the trading volume is increasing towards the end of the year for small companies, while it is decreasing for large companies; the trading volume is lower in early January than in December (Lakonishok & Smidt, 1984). De Bondt and Thaler (2012) reassert the interaction between small (losing) firms and January effects, concluding that losers have large excess returns while winners do not. More generally, the month-of-the-year effect exhibits exceptionally large returns in January in most of the countries and in April in the UK (Gultekin & Gultekin, 1983).

5.2.2 The Volume-Price Relation

The empirical evidence from the literature broadly supports the positive correlation between volume and price changes (Harris & Raviv, 1993) (Hong & Stein, 2007). The articles on the price-volume relation reported two forms of price indicators that are correlated with trading volume: first, the magnitude (or absolute value) of the price change, i.e. $|\Delta p|$ (Assogbavi & Osagie, 2006); second, the price change per se (or the raw price change value), i.e. Δp (Karpoff, 1987). The price change can be either the log-price difference or the percentage price change.

5.2.3 Stock Index Futures Expiry

Stock index futures were introduced in 1982 and are the second most widely traded futures markets by investors, after interest rates (CME Group, 2013). They consist of a prediction of where the underlying index cash market will be and introduced the concept of a cash settlement mechanism in order to address the problem of logistical difficulties regarding the delivery of the actual stocks associated with a particular stock index. Their expiry dates represent the dates when the futures contracts stop trading and when the final price settlement occurs. The expiry dates for the investigated European stock index futures occur on the third Friday of the expiry month or the previous day in case this is a bank holiday. The indices' futures contracts are traded either on a quarterly basis, i.e. March, June, September and December (e.g. FTSE 100 and DAX 30), or monthly (e.g. CAC 40, FTSE MIB, IBEX, Amsterdam Exchange, and OMX Stockholm 30). There are two broad categories of players in the futures market, namely hedgers, who are protecting against price risks, and speculators, who seek profits from the price changes that hedgers are protecting against.

A potential explanation of the larger trading volumes before the stock index futures expiry date consists of the investors who want to roll their futures contracts. They are going to maintain the position beyond the date and they have to exchange their contract for the next contract (i.e. rolling the position) when the contract expires. However, some close their positions in the run-up to the expiry. There are also mechanisms such as ETFs (i.e. exchange-traded funds) that are tracking indices.

When investors engage in a futures contract, they buy the exposure. Entering a futures contract is done synthetically; people are not buying or trading the underlying basket of stocks, however they do something related to the futures. There is increased activity in the run-up volume, which leads to an increase in volatility, which then spills over to the equity markets (temporary departures from the fair value of the index future than the actual basket of stocks).

5.2.4 MSCI Quarterly Index Review

The Morgan Stanley Capital International (MSCI) indices group is an investment decision support provider and its indices have been tracked closely by international fund managers since 1969. Approximately $8 trillion are estimated to be benchmarked to the MSCI indices worldwide (MSCI, 2014) and 97 of the top 100 largest asset managers are served by MSCI (MSCI, 2015). Any stock addition or elimination in any MSCI index attracts significant investor attention across the world. In order to reflect the evolving market, the MSCI indices constituent list changes on a quarterly basis, in February, May, August and November, close to the last trading day of these four rebalancing months. The MSCI national indices' changes are announced two weeks prior to the effective date, allowing the investors to react to the MSCI announcements.

The main objective of index funds is to replicate the performance of a given benchmark. Fund managers need to provide the lowest costs and high transparency to their clients, i.e. equity investors, and are more likely to minimise the benchmark tracking error than to take risks for increasing the returns. MSCI index rebalancing revision schedules are publicly released well before the effective revision date, giving rise to speculations. There are clear abnormal returns around the announcement and implementation dates of the MSCI reviews, with a high concentration in the preceding trading days to implementation. This is followed by reversal after the implementation date. Most importantly, the MSCI abnormal returns were correlated with the trading volume, concluding that the majority of fund managers re-adjust their portfolios at the last minute in order to minimise the tracking error. For the additions and deletions of the MSCI review, the trading volume was on average four times higher on the implementation day than on normal trading days (The Trade, 2007).

5.3 Data Set

The sample data set acquisition and processing is described in this section. The stock universe includes a sample of 495 unique stocks, out of which 401 are members of the indices considered for the stock index futures expiry analysis and 338 are constituents of the MSCI International Pan Euro Price Index. The data spanning over 15 years is complemented by a series of special events, which are potentially associated with non-stationarity, consisting of the stock index futures expiries and MSCI quarterly review dates. The data challenges consisted of the unavailability of a historical 'expiry calendar' for European markets and a public list of leavers and joiners for a given index. Therefore, we constructed an accurate expiry calendar for the pan-European markets covering the futures expiries for seven liquid indices and the MSCI quarterly reviews for the MSCI International Pan Euro Price Index. These were supplemented by historical evidence of additions and eliminations for each index, which allowed us to generate an accurate snapshot of the constituent list for a given index and a given date.

5.3.1 Market Data Acquisition and Processing

The study contains a list of current constituents of the indices in Table 5.1 as of 11th May 2015. We retrieved the historical evidence of additions and eliminations for each of these indices and created a historical log of leavers and joiners, in order to be able to generate an accurate list of constituents for a given index at a given point in time. Table 5.1 includes the RIC (Reuters Identification Codes) of the indices and the total number of constituents as of 11th May 2015 (i.e. 'current constituents') and for the entire study period (i.e. 'historical constituents'). There are 32,408 observations for the stock index futures expiry analysis, and 10,298 observations for the MSCI rebalance analysis.

Based on the union of the current and past constituents, daily market data containing OHLC (open, high, low, close) prices and end-of-day volume is retrieved for each stock. The daily data was extracted from Thomson Reuters using a VBA script in order to automate the market data retrieval. Moreover, we replaced the trading volume of a stock's primary RIC by its consolidated volume, which was computed as the sum of a stock's main exchange trading volume and its volume on MTFs (multilateral trading facilities). The consolidated volume is used throughout this study since it provides a better picture of the real liquidity for a given stock and it is referenced simply as 'volume' hereafter.

Table 5.1
Market data European indices for the futures expiry analysis and MSCI rebalance analysis.

Analysis Type	Index RIC	Index Name	Current Constituents	Historical Constituents	Location
Futures expiry	.AEX	Amsterdam Exchange Index	25	37	Netherlands
	.FCHI	CAC 40 Index	40	54	France
	.FTMIB	FTSE MIB Index	40	51	Italy
	.FTSE	FTSE 100 Index	100	149	United Kingdom
	.GDAXI	Deutsche Boerse DAX Index	30	37	Germany
	.IBEX	IBEX 35 Index	35	44	Spain
	.OMXS30	OMX Stockholm 30 Index	30	33	Sweden
MSCI rebalance	.MSPE	MSCI International Pan Euro Price Index EUR Real Time	204	338	Europe

We dealt with missing data, such as the current constituent list and past leavers and joiners for FTSE MIB Index, which could not be accessed from Thomson Reuters in the first instance. Data pre-processing and cleansing involved filtering stocks with at least 100 days of available daily market data, and appending metadata to each stock, including information such as exchange location, currency, company market capitalisation, economic sector, business sector name, industry/subindustry name, activity name etc.

Table 5.2 shows the country distribution for the MSCI Pan-European Index, where the two-letter country codes are represented using the standard ISO 3166-1 alpha-2. Each stock is associated with a country based on its exchange country; for example, a Spanish stock's country is set as United Kingdom if this stock is trading on the London Stock Exchange.

Table 5.2
MSCI constituents - country breakdown.

Country Code	Country Name	Historical Constituent Count	Historical Constituent Percent	Current Constituent Count	Current Constituent Percent
AT	Austria	6	1.78	2	0.98
BE	Belgium	10	2.96	4	1.96
CH	Switzerland	24	7.10	18	8.82
DE	Germany	39	11.54	33	16.18
DK	Denmark	8	2.37	6	2.94
ES	Spain	17	5.03	12	5.88
FI	Finland	7	2.07	4	1.96
FR	France	53	15.68	35	17.16
GB	United Kingdom	90	26.63	43	21.08
GR	Greece	6	1.78	0	0.00
IE	(Republic of) Ireland	4	1.18	1	0.49
IT	Italy	23	6.80	13	6.37
NL	Netherlands	17	5.03	13	6.37
NO	Norway	8	2.37	4	1.96
PT	Portugal	3	0.89	2	0.98
SE	Sweden	23	6.80	14	6.86

5.3.2 Calendar Data Taxonomy

We constructed an accurate expiry calendar for the stock index futures and a rebalance calendar for the MSCI quarterly reviews, which provide a representative illustration of the main expiry dates in Europe for the most liquid indices. A total number of 1,042 futures expiries, and 49 MSCI rebalance dates are included in the calendar data.

We generate a dynamic calendar of periodic trading events for stock index futures expiries and MSCI rebalances. The futures expiry dates are computed for 7 European indices, which expire on the third Friday of the expiry month, which occurs either monthly or quarterly (i.e. December, March, June and September). When the third Friday is a non-trading day, the stock index futures expiry is substituted by the previous working day. The stock index futures expiries were generated for the 7 European indices below. There was no need to consider the Euro STOXX 50 index since its constituent list overlaps with the blue chip companies contained in the indices below. We retrieved each country's non-trading calendar in order to determine if the expiry for a given index falls on the third Friday of the expiry month or on the previous trading day if the expiry day falls on a bank holiday. The futures contract specifications were retrieved from Euronext (AEX and CAC 40), Eurex Exchange (DAX 30), London Stock Exchange (FTSE 100), Borsa Italiana (FTSE MIB), Bolsas y Mercados Españoles (IBEX 35) and NASDAQ OMX (OMXS30). The stock index futures expiry calendar covers the following seven indices, where only FTSE 100 and DAX have quarterly expiries:

- FTSE 100 Index Futures – quarterly expiry;
- CAC 40 Index Futures – monthly expiry;
- DAX 30 Index Futures – quarterly expiry;
- FTSE MIB Index Futures – monthly expiry;
- IBEX 35 Index Futures – monthly expiry;
- Amsterdam Exchange (AEX) Index Futures – monthly expiry;
- OMX Stockholm 30 (OMXS30) – monthly expiry.

The MSCI rebalances are typically implemented on the last trading day of the following quarterly cycle: February, May, August, and November. However, there are very few exceptions when the MSCI quarterly review date falls a few days before the end-of-month. When the rebalance date falls on a trading holiday in a given market, then the relevant trading date of MSCI rebalance is the closest previous trading day. The review dates were double-checked with the quarterly index review documents from www.msci.com and span from February 2003 until May 2015.

For each of these indices (i.e. the 7 indices for futures expiry and the MSCI Pan-Euro Index), we get a historic log of leavers and joiners covering the entire study period. This facilitated the creation of a snapshot of the constituent list for a given index and a given date. This process starts with the current constituent list (i.e. as of 11th May 2015), and then iterates the historical log of leavers and joiners by going backward in time and reversing each index constituent action (e.g. if a particular stock was added to an index between 11th May 2015 and the given snapshot date, it means it was not part of the index constituent list prior to this intermediate date) until we reach the given snapshot date.

The logarithmic volume histogram in Figure 5.1 exhibits higher volumes on the futures expiry day, whose volumes are more negatively skewed compared to the expiry day's benchmark period volume, which is calculated as the median of the previous 20 trading days.

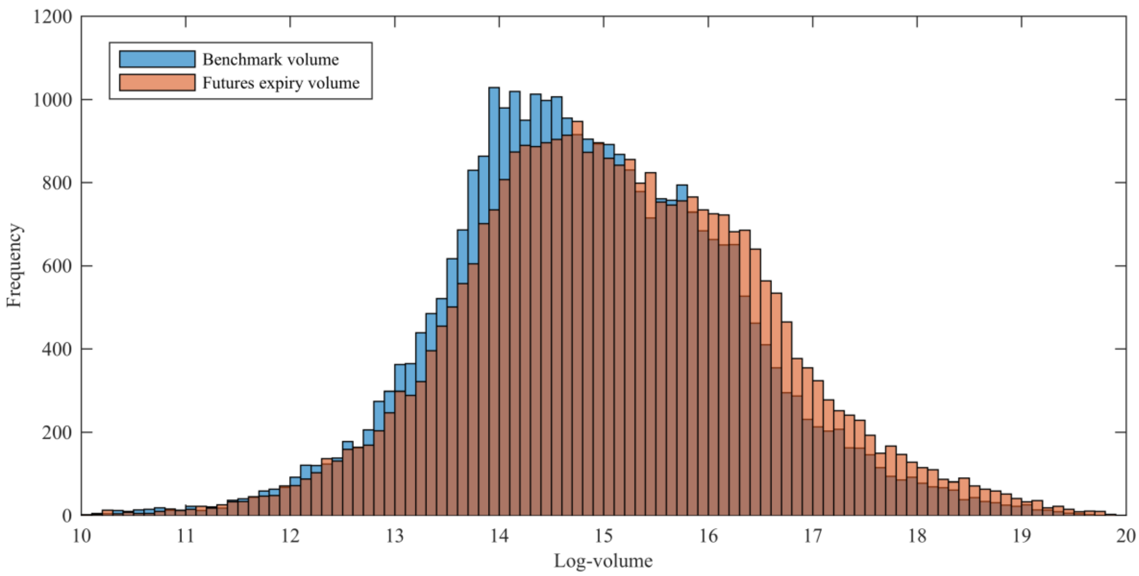

Figure 5.1: Histograms of the logarithmic volume data for the futures expiries and the benchmark periods.

As with the futures expiry effect, the logarithmic volume histogram in Figure 5.2 illustrates a slightly more negative skew for the trading volumes on the MSCI quarterly index review day than the volumes of the MSCI rebalance benchmark period, suggesting higher trading volumes on the MSCI rebalance dates.

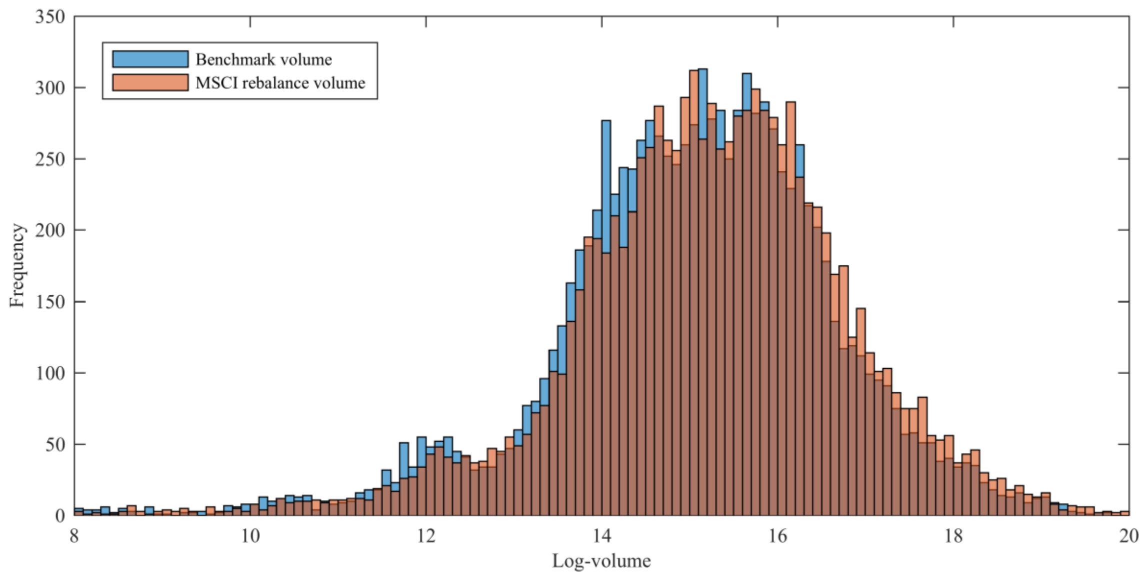

Figure 5.2: Histograms of the logarithmic volume data for the MSCI rebalances and the benchmark periods.

5.4 Analysis Approach

This section describes the analytical approach for the futures expiry and MSCI rebalance models. The study commences by validating the existence of the investigated phenomena (i.e. the futures expiry and the MSCI rebalance, and their relation with an increase in trading activity) by employing randomisation tests. Once their existence is confirmed in the European equity markets as being statistically significant, we build a predictive model, by fitting a number of futures expiry and MSCI rebalance stepwise regressions (i.e. linear regression models, followed by sequential feature selection).

The volume on a special date (also called 'target date' or t_0, i.e. futures expiries or MSCI rebalances) is compared with the volume of a benchmark period, which was defined as the median of the 20 trading days prior to a given future expiry or MSCI rebalance. We chose the median as a measure of central tendency because median is robust to outliers. The study involves data that are periodic, but sparse. There is a number of expiries and rebalances and we normalise the analysis data in order to identify effects that are common to some stocks and a particular target date, either futures expiry or MSCI rebalance.

The study also considers a multi-step ahead prediction, up to a step size of 6 trading days. For instance, a 6-step ahead analysis would compute the benchmark volume for the previous 20 trading days for a given date in order to predict the volume impact in 6 days' time. The default analyses in this study consider one-step ahead forecasting, although the default step size of $n = 1$ day can be lagged and therefore the step ahead lag is defined as $l = n - 1$. Based on this notation, we define the relative volume for a given expiry or rebalance as the log-ratio between the volume on expiry/rebalance day and its benchmark volume, computed as the median of the previous 20 trading days, as shown in Equation (5.1). The target variable in all regression models in this study is the relative volume:

$$V = \log \frac{V_{t0}}{\mathrm{median}(V_{t-l-1}, V_{t-l-2}, \dots, V_{t-l-20})}. \tag{5.1}$$

Figure 5.3 shows the relative volume on futures expiries and MSCI rebalances, using three methods of aggregating the benchmark volumes, i.e. median (Panel A), arithmetic mean (Panel B), and geometric mean (Panel C). These measures of central tendency determined the benchmark volume in order to compute the relative volume using Equation (4.2):

$$V = \log \frac{V_{t0}}{f_{\mathrm{aggr}}(V_{t-1}, V_{t-2}, \dots, V_{t-20})}. \tag{5.2}$$

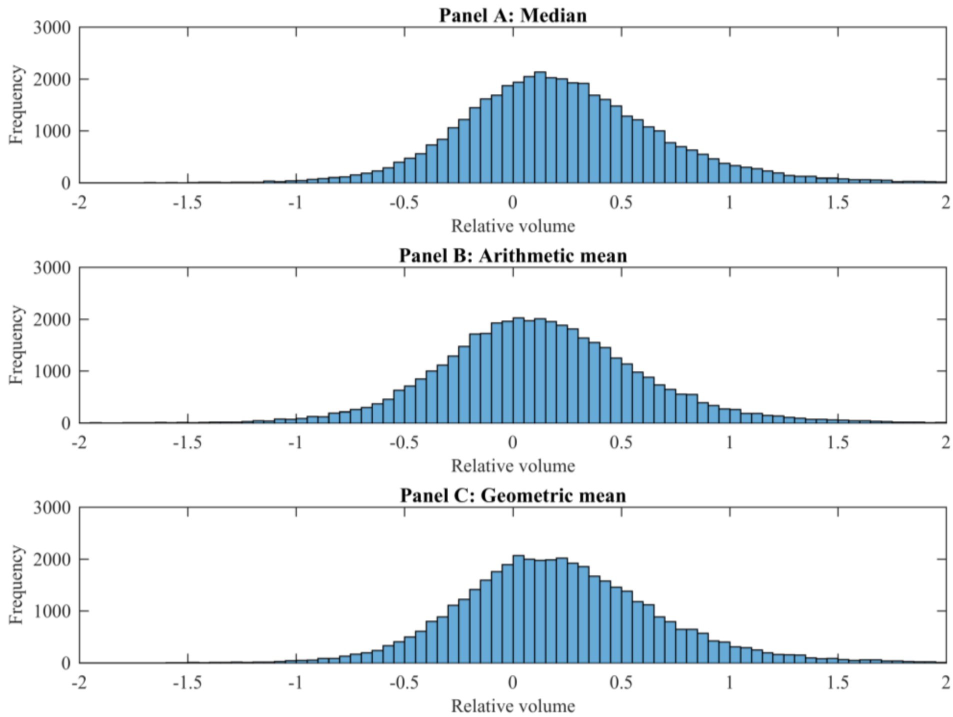

Figure 5.3: Histograms of the relative volume on futures expiries and MSCI rebalances with different methods of benchmark volume aggregation.

The outliers were best handled by the median and therefore we used it throughout the analyses to determine the benchmark volume.

The futures expiry and MSCI rebalance analyses investigated a prior or posterior effect in the trading volumes and therefore it allowed for offsets relative to the target date, ranging from -5 days to +5 days. For example, for an offset of -3 days, we compute the target date by subtracting 3 trading days from the main target date (i.e. the expiry or rebalance day). A zero-offset analysis considers the expiry day or rebalance day itself. Consequently, we could analyse when the trading volume starts increasing and when it returns to the normal level.

The analysis models can be classified as expiry models and rebalance models and are fit on different data sets (i.e. different indices). Since we allow for target date offsets, both model classes are fit with and without indicator variables for the number of days relative to the expiry/rebalance day, resulting in 11 additional predictors (ranging from -5 days to +5 days).

5.5 Randomisation Analysis

The following randomisation tests address the existence of higher trading activity on the expiry and rebalance days and we test the futures expiry effect against the Friday effect, and the MSCI rebalance effect against the end-of-month effect with regard to higher trading activity. Their control dates account for the day-of-the-week effect and maintain the same proportion of days of the week as the target dates.

The existence of potential structural breaks is analysed in the following randomisation tests in order to allow us to assume structural homogeneity. For this reason, the sample period is divided in two halves (i.e. 1st January 2000 – 31st December 2007 and 1st January 2008 – 10th May 2015), and each of these subsamples is analysed, along with the entire sample period. The rationale of dividing the sample on 1st January 2008 is twofold: first, this is an approximate midpoint for our entire sample period; and, second, this coincides with the financial crisis of 2007-08, whose peak was reached when Lehman Brothers collapsed on 15th September 2008.

The randomisation test generally checks whether two data vectors are significantly different. The difference between these vectors' means is the observed statistic. We randomise the two vectors' labels 1,000 times and we compute the newly reshuffled vectors' mean difference. Eventually we test whether the randomised differences are more extreme than the observed difference, resulting in an empirical p-value, which is calculated as the percentage of randomisations where the observed difference is larger (for the right-tailed or two-tailed tests) or smaller (for the left-tailed test) than the randomised differences. The p-value represents the probability of observing a test statistic at least as extreme as the observed value under the null hypothesis, and if it is small then the validity of the null hypothesis is considered uncertain. When the empirical p-value is below the chosen significance level ($\alpha = 5\%$), we reject the null hypothesis.

All of the following randomisation tests are pairwise, and, for each target date, a particular control date is chosen, which is conditioned on the target date. Therefore, the labels are reshuffled on a pairwise basis, flipping a coin for each element in order to decide whether to interchange the target date and the control date.

5.5.1 Futures Expiries vs. Control Dates

The target dates for the randomisation test between futures expiries and control dates consist of all futures expiry dates. For each target date, we choose the closest control date that falls exactly one or two weeks before or after the expiry date. Therefore, the control date falls on the same day of week as the target date. The test is conducted for each target date offset. When the offset is positive, we do not allow the control date to fall one week before the target date, as it would overlap with the critical days around the expiry date. Similarly, when the offset is negative, the control date cannot fall one week after the target date. There is a two-tailed test and a right-tailed test. The null hypothesis of the two-tailed test is that the difference between the relative trading volume on the (offset) futures expiries and the relative trading volume of the control dates comes from a distribution with zero mean, whereas the alternative hypothesis of the right-tailed test is that the mean of the futures expiry relative volumes is less than the mean of the control date relative volumes.

Table 5.3 shows the randomisation test results for the futures expiry day (i.e. no offset), using 1-step ahead modelling. The results include the aggregated indices, along with a breakdown by individual index, and monthly vs. quarterly expiry indices. Table 5.4 shows the results for the aggregated indices for offsets -5 days to +5 days. The randomisation tests reveal that the trading volume on the expiry date of each index is significantly higher. This is also the case for offsets -4, -3, -2, -1, +1 and +2, meaning that the trading volume surges 4 days before the expiry date (i.e. generally on the Monday of the expiry week) and remains at high levels for two days after the expiry day. The results are consistent across the sample period halves and there are no structural breaks for the futures expiry elevated volume. The multi-step ahead modelling rejects the null hypothesis for the same offsets, although the *p*-value varies insignificantly in very few instances, without changing the null hypothesis rejection decision. Figure 5.4 illustrates the relative volume distribution for dates with futures expiries (i.e. target dates) and dates with no futures expiries (i.e. control dates), exhibiting larger trading volumes on futures expiries.

Table 5.3
Randomisation tests between futures expiries and control dates – no target date offset, 1-step ahead modelling.

Sample period/s	Analysis type	Index RIC	Target date offset	Stocks	Target dates	Randomisation tail/s	*p*-value	Reject H_0
2000-2007, 2008-2015, 2000-2015	Individual index	.FTSE	0	149	5,023	both	0	Yes
		.FTSE	0	149	5,023	right	0	Yes
		.GDAXI	0	37	1,724	both	0	Yes
		.GDAXI	0	37	1,724	right	0	Yes
		.FCHI	0	54	6,929	both	0	Yes
		.FCHI	0	54	6,929	right	0	Yes
		.FTMIB	0	51	4,934	both	0	Yes
		.FTMIB	0	51	4,934	right	0	Yes
		.IBEX	0	44	5,307	both	0	Yes
		.IBEX	0	44	5,307	right	0	Yes
		.AEX	0	37	3,633	both	0	Yes
		.AEX	0	37	3,633	right	0	Yes
		.OMXS30	0	33	4,858	both	0	Yes
		.OMXS30	0	33	4,858	right	0	Yes
	Monthly expiry indices	.FCHI, .FTMIB, .IBEX, .AEX, .OMXS30	0	215	25,661	both	0	Yes
		.FCHI, .FTMIB, .IBEX, .AEX, .OMXS30	0	215	25,661	right	0	Yes
	Quarterly expiry indices	.FTSE, .GDAXI	0	186	6,747	both	0	Yes
		.FTSE, .GDAXI	0	186	6,747	right	0	Yes
	All indices	.FTSE, .GDAXI, .FCHI, .FTMIB, .IBEX, .AEX, .OMXS30	0	401	32,408	both	0	Yes
		.FTSE, .GDAXI, .FCHI, .FTMIB, .IBEX, .AEX, .OMXS30	0	401	32,408	right	0	Yes

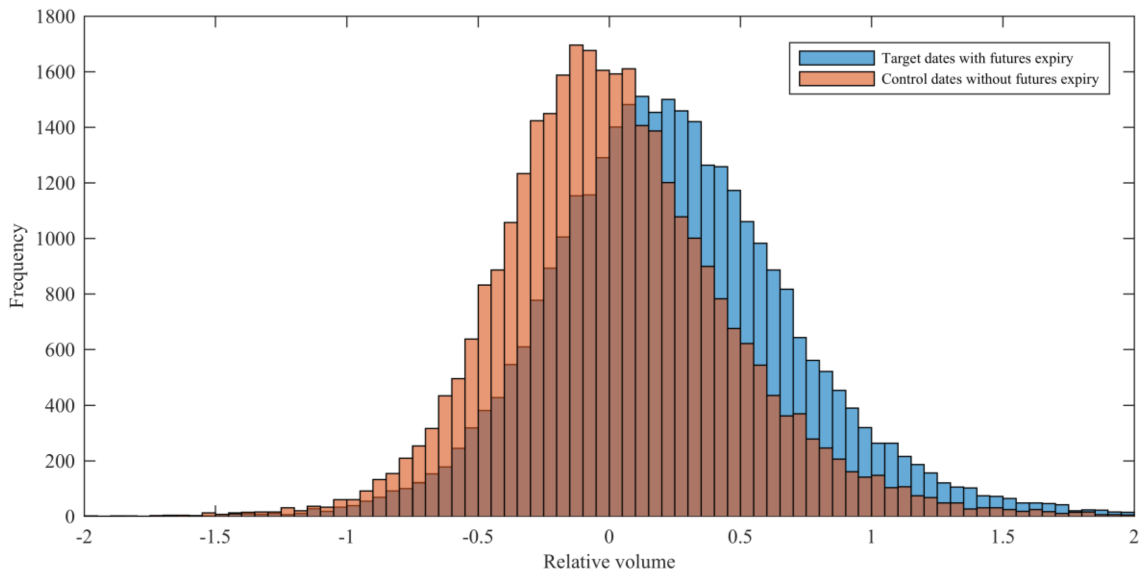

Figure 5.4: Relative volume distribution for dates with futures expiries and dates with no futures expiries.

Table 5.4
Randomisation tests between futures expiries and control dates – all indices, 1-step ahead modelling.

Sample period/s	Index RIC	Target date offset	Stocks	Target dates	Randomisation tail/s	*p*-value	Reject H_0
2000-2007, 2008-2015, 2000-2015	.FTSE, .GDAXI, .FCHI, .FTMIB, .IBEX, .AEX, .OMXS30	0	401	32,408	both	0	Yes
		0	401	32,408	right	0	Yes
		-5	401	32,413	both	0	Yes
		-5	401	32,413	right	1	No
		-4	401	32,406	both	0.008	Yes
		-4	401	32,406	right	0.003	Yes
		-3	401	32,403	both	0.011	Yes
		-3	401	32,403	right	0.002	Yes
		-2	401	32,411	both	0.144	No
		-2	401	32,411	right	0.041	Yes
		-1	401	32,414	both	0	Yes
		-1	401	32,414	right	0	Yes
		1	401	32,415	both	0	Yes
		1	401	32,415	right	0	Yes
		2	401	32,416	both	0	Yes
		2	401	32,416	right	0	Yes
		3	401	32,411	both	0	Yes
		3	401	32,411	right	1	No
		4	401	32,411	both	0	Yes
		4	401	32,411	right	1	No
		5	401	32,407	both	0	Yes
		5	401	32,407	right	1	No

5.5.2 Futures Expiries vs. Fridays

Next, we investigate whether the higher volume associated with the futures expiries is actually caused by the Friday effect or whether it is driven solely by the futures expiry. The target dates consist of all futures expiry dates falling on Fridays. There are 30 instances of index futures expiries falling on the previous day, i.e. on a Thursday. These 30 non-Friday expiries belong to various indices and there are actually 13 unique non-Friday expiries, associated with 912 stocks, which have been discarded for this randomisation test. The control date for each target date is the closest Friday (in terms of the difference in calendar days from the target date) falling one or two weeks from the expiry day (i.e. -2, -1, +1, +2 week/s relative to the expiry day). The alternative hypothesis is that the relative volume on futures expiries is significantly different from (for the two-tailed test) or larger than (for the right-tailed test) the relative volume on non-expiry Fridays. The randomisation tests in Table 5.5 reject the null hypothesis and conclude that the Fridays with futures expiries are the drivers of increased volumes on Fridays. The results are consistent among the two sample halves, i.e. 2000-2007 and 2008-2015. Figure 5.5 contains the overlapping histograms of the relative volume for the Fridays with and without futures expiries and illustrates the larger volumes associated with the expiry Fridays.

Table 5.5
Randomisation tests between futures expiries and Fridays – 1-step ahead modelling, all futures indices.

Sample period/s	Stocks	Target dates	Randomisation tail/s	*p*-value	Reject H_0
2000-2007, 2008-2015, 2000-2015	401	31,496	both	0	Yes
	401	31,496	right	0	Yes

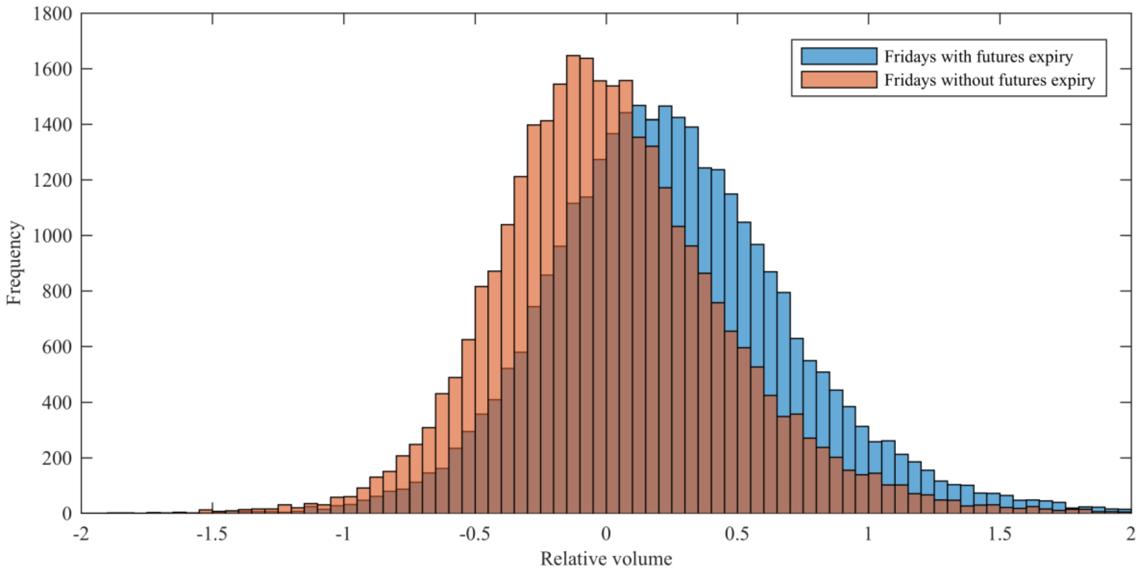

Figure 5.5: Relative volume distribution for Fridays with futures expiries and Fridays with no futures expiries.

5.5.3 MSCI Rebalances vs. Control Dates

We further test whether the relative trading volume on MSCI rebalances is higher than the volume on the last trading day of the previous or following month. The target dates consist of all (offset) MSCI rebalance dates. For each target date, we find the closest control date that is the last trading day of the previous or the following month. If the target date is offset, then the control date is offset as well. We perform a two-tailed test and a right-tailed test. The alternative hypothesis of the two-tailed test is that the relative trading volume of the relative dates is significantly different from the volume on control dates, whereas the alternative hypothesis of the right-tailed test is that the relative volume of the target dates is larger than the relative volume of the control dates. Table 5.6 shows the randomisation test results, which confirm that the relative volume on MSCI rebalances is significantly higher than the relative volume of the last trading days of the months without MSCI rebalances. This is also the case for offset -1 (for 1-step ahead and 2-step ahead analyses only) and for offset +5 (for all step ahead lags). Therefore, the trading volume surges one day before the review date, and then goes back to the normal level after the rebalancing, but picks up again in exactly one week (i.e. first week of the following month). The same results are obtained for 2000-2007. Slightly different results are generated for 2008-2015, where the offsets with larger volume are +1 and +5, instead of -1 and +5 trading days. We conclude that the trading volumes is generally larger on the trading day before the review day and on the effective MSCI rebalance date.

Table 5.6
Randomisation tests between MSCI rebalances and control dates – 1-step ahead modelling.

Sample period/s	Target date offset	Stocks	Target dates	Randomisation tail/s	*p*-value	Reject H_0
2000-2007, 2008-2015, 2000-2015	0	338	10,298	both	0	Yes
	0	338	10,298	right	0	Yes
	-5	338	10,340	both	0	Yes
	-5	338	10,340	right	1	No
	-4	338	10,338	both	0	Yes
	-4	338	10,338	right	1	No
	-3	338	10,341	both	0.181	No
	-3	338	10,341	right	0.069	No
	-2	338	10,341	both	0	Yes
	-2	338	10,341	right	1	No
	-1	338	10,341	both	0.002	Yes
	-1	338	10,341	right	0	Yes
	1	338	10,341	both	0.726	No
	1	338	10,341	right	0.377	No
	2	338	10,337	both	0	Yes
	2	338	10,337	right	1	No
	3	338	10,337	both	0	Yes
	3	338	10,337	right	1	No
	4	338	10,338	both	0	Yes
	4	338	10,338	right	1	No
	5	338	10,340	both	0.001	Yes
	5	338	10,340	right	0	Yes

Figure 5.6 illustrates the relative volume distribution for dates with MSCI rebalances (i.e. target dates) being slightly higher than the relative volumes on dates with no MSCI rebalances (i.e. control dates).

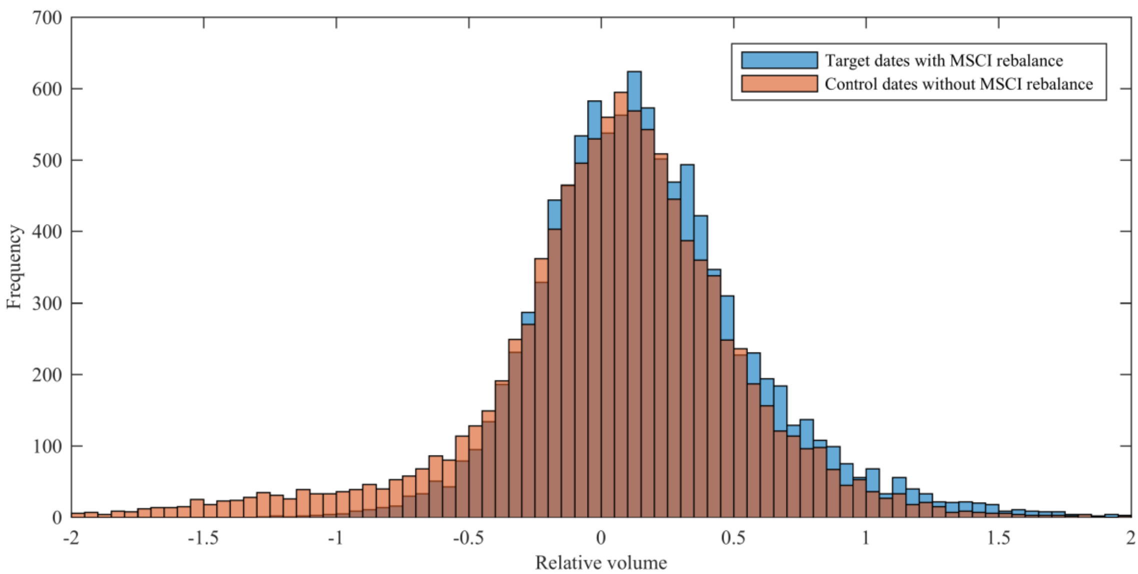

Figure 5.6: Relative volume distribution for dates with MSCI rebalances and dates with no MSCI rebalances.

5.5.4 MSCI Rebalances vs. End-of-Month Effects

The randomisation test between MSCI rebalances and end-of-month effects aims to identify the main driver of larger volumes around the end-of-month. For this test, we define the relative monthly trading volume as the log-ratio between mean volume on the last 5 trading days of that month and the mean volume on the first 10 days of that month, as outlined in Equation (5.3), where V_i represents the daily volume on the i^{th} day of a given month, and n represents the total number of trading days of a given month:

$$V_{\text{month}} = \log \frac{\sum_{i=1}^{5} V_{n-i}}{\sum_{i=1}^{10} V_i}. \tag{5.3}$$

We use the arithmetic mean instead of median (as with the relative volumes for a certain target date) because in this case, we are quantifying the volumes occurring at the beginning of month and at the end-of-month, and the arithmetic mean incorporates better all observations throughout these periods. Certain volume trends occur over multiple dates and therefore such effects would be better accounted for by using the arithmetic mean.

The target dates consist of all MSCI rebalance months. For each MSCI quarterly review month, we consider the previous and following months and ultimately flip a coin in order to choose whether the previous month or the following month is selected as the control date. We perform a two-tailed test and a right-tailed test for the relative monthly volume of the target months and control months. The alternative hypothesis is that the relative monthly volume on MSCI quarterly review months is significantly different from (for the two-tailed test) or significantly larger than (for the right-tailed test) the relative monthly volume on the months with no MSCI rebalance. Based on the results in Table 5.7, we report that volume on MSCI rebalance months is significantly different from the volume on months with no MSCI review, but the large trading activity associated with the MSCI rebalances cannot explain the large volumes around the end-of-month. Figure 5.7 visually supports this conclusion and illustrates the relative volume distribution for months with MSCI rebalances (i.e. target months) and months with no MSCI rebalances (i.e. control months). The monthly volume on MSCI review months has a higher kurtosis than the months with no MSCI rebalances; the monthly volume for control dates has more extreme values on both tails, having a larger dispersion than the target months.

Table 5.7
Randomisation tests between MSCI rebalances and end-of-month effects – 1-step ahead modelling.

Sample period/s	Stocks	Target dates	Randomisation tail/s	p-value	Reject H_0
2000-2007, 2008-2015, 2000-2015	338	10,298	both	0.005	Yes
	338	10,298	right	0.999	No

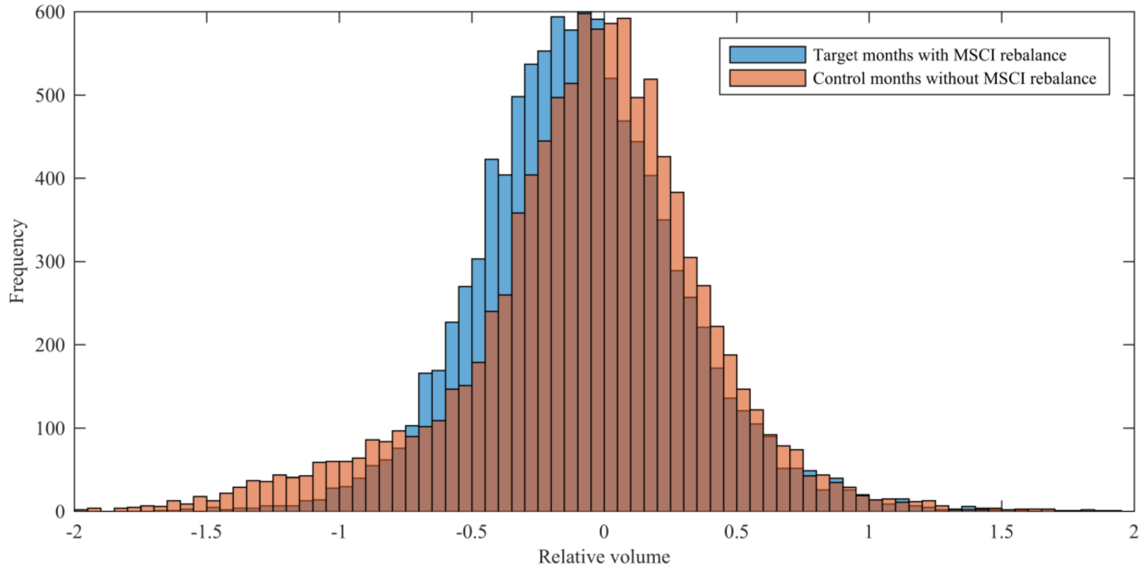

Figure 5.7: Relative volume distribution for months with MSCI rebalances and months with no MSCI rebalances.

5.5.5 Summary

The previous randomisation tests provide a methodological rigour for inferring a conclusion with regard to the existence of the studied phenomena. The tests generally found no structural breaks around the financial crisis of 2007-08, with the exception of a reversing effect for a couple of days adjacent to the MSCI quarterly review dates.

We report significantly higher trading volumes associated with both futures expiries (starting four days before the expiry and lasting two days after the expiry) and MSCI rebalances (starting on the day preceding the rebalance and returning to normal levels the following day after the quarterly review day). We found that the Friday effect does not explain the surge in volumes on futures expiries. Although we present evidence that the trading volumes of the months with MSCI quarterly reviews are statistically significant, we draw the conclusion that the larger volumes of the MSCI rebalances cannot explain the end-of-month effect.

5.6 Predictive Modelling

Given the empirical evidence provided by the randomisation tests, we further investigate the effect size of the futures expiries and MSCI rebalances in connection with trading volume.

5.6.1 Modelling Approach

The models follow a general stepwise regression framework, which starts by collecting the data, depending on the model (i.e. expiry or rebalance model), and aggregates the predictors for each target date in the regression matrix. Then it performs stratified partitioning on the data set, by creating 10 folds of random subsamples with similar proportions of observation classes. Each class is defined for a unique combination of values for the indicator variables (i.e. predictors whose values are only binary, e.g. 'trading *country code*', 'expiry *index RIC*', 'offset +/- *n* days' etc.). The stratified partitioning provides robust results since the classes are evenly distributed across the folds, especially when the data set is unbalanced, and the models are trained and tested based on observations from all classes. Once the 10 folds are defined, the framework proceeds to fitting a multiple linear regression, followed by forward feature selection, where the variable selection objective function minimises the mean squared error (MSE) using 10-fold cross-validation (CV). We did not use backward elimination because the models are defined with a constant term (or intercept) and the regression design matrix contains full categorical variables (i.e. categorical variables with *n* possible values are encoded as *n* predictors, instead of *n-1*, because we are exploring the statistical significance of these predictors and perform feature selection on the *n* possible values) and would lead to multicollinearity issues, where the regression design matrix is rank deficient.

The study also investigates the volume autoregression in the context of special dates (i.e. expiries and rebalances). Hence, we fit the two model classes with and without 20 lagged volumes, which are normalised by dividing them by their benchmark volume (i.e. the median of the 20 lagged volumes). The volume normalisation is performed in order to account for the different magnitude of the trading volume across different stocks. The normalisation is consistent with the relative volume, which also divides the target volume by the benchmark volume.

We fit a linear regression model for the stock index futures effect, and the MSCI rebalance effect, respectively. All of the models contain a constant/intercept term. We reduce the dimensionality of these full models by performing sequential feature selection and retrieving a reduced model with fewer features (or predictor variables), while minimising the predictive error of the fit models using different subsets. When performing feature selection, the intercept is always kept in the reduced model. Similarly, if a given model is defined with 20 lagged volumes, these predictors are kept in the model. The objective function of the sequential feature selection seeks to minimise the criterion, which we chose to be the MSE, throughout the potential feature subsets.

We employed a forward selection sequential search algorithm for feature selection, where features are sequentially added to the starting model (i.e. only the constant/intercept term, and possibly the 20 lagged volumes) until no other features can be added in order to decrease the criterion. It is unfeasible to have an exhaustive approach and fit all the feature subsets of a model with *n* features due to time and processing constraints, and therefore the sequential search algorithm moves only in one direction, always growing the candidate feature set (if using forward selection).

Every time a candidate feature is added to or removed from the model feature set, the candidate model with the new feature set is cross-validated using the objective function, which minimises the MSE criterion. 10-fold stratified cross-validation is applied throughout the analyses of this study, using the same 10 folds that were initially defined in the stratified partitioning of the data set.

5.6.2 Models Outline

There are eight full models that are fit in this study and Table 5.8 outlines their full candidate feature sets. The features whose names are marked in italics on the left-hand side column indicate multiple features. For instance, 'Trading *country*' would substitute *country* by each trading country of the constituents of the MSCI Pan-European Index, e.g. 'Trading GB', 'Trading DE', 'Trading FR' etc. There are also 20 features for the lagged normalised volume corresponding to each trading day.

Table 5.8
Regression models – full candidate features.

	Futures expiry models				MSCI rebalance models			
Intercept	✓	✓	✓	✓	✓	✓	✓	✓
Trading *country code*					✓	✓	✓	✓
Expiry *index RIC*	✓	✓	✓	✓				
Quarterly expiry	✓	✓	✓	✓				
Target date offset (from -5 to 5 days)	✓	✓			✓	✓		
20 lagged normalised volumes	✓		✓		✓		✓	

The study provides two separate models for the expiry day effect, one for the stock index futures expiry and one for the MSCI quarterly index review.

5.6.2.1 Futures Expiry Models

For this part of the study, we use stocks provided that they are members of one of the seven indices allowing for futures. The target date can vary from 5 days prior to the expiry day to 5 days after the expiry day and therefore the benchmark period of 20 days is shifted accordingly, accounting for the chosen step size as well (expressed in days). The left-hand side column in Figure 5.8 (i.e. Panels A – F) shows the relative volume distribution for the negative target date offsets, ranging from 1 to 5 days prior to the futures expiry, whereas the right-hand side column, corresponding to Panels G – L, includes the positive target date offsets, ranging from 1 to 5 days after the expiries. In both columns, the top panel (i.e. Panel A and Panel G) illustrates the volume on the futures expiry.

The stock index futures expiry models include the constant term, seven 'expiry *index RIC*' indicator variables, and an additional indicator variable denoting whether the index has a quarterly expiry. Depending on the model definition, the predictors of some models could include 20 lagged normalised volumes and 11 indicator variables for the target date offset (ranging from -5 days to +5 days).

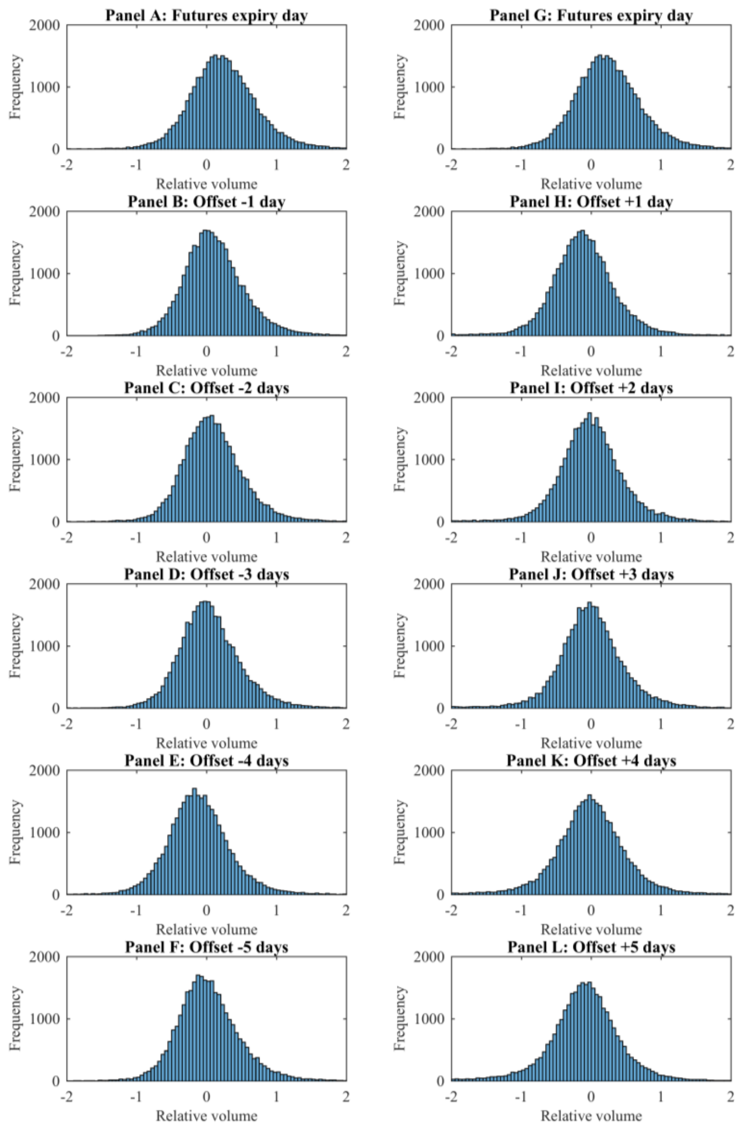

Figure 5.8: Relative volume distribution for positive target date offsets (Panels A – F) and negative target date offsets (Panels G – L) relative to the futures expiries.

5.6.2.2 MSCI Rebalance Models

The MSCI rebalance effect analysis consists of 204 constituents of the MSCI International Pan Euro Price Index from 15 European countries. This is a heuristic approach having a general date for MSCI quarterly index review, which does not account accurately for every country. The target date of the regression model can vary from 5 days prior to the rebalance day to 5 days after the rebalance day. The benchmark volumes are calculated depending on the chosen target date and step size. The model full candidate features include the intercept and ‘trading *country code*’ for each of the unique countries where MSCI constituents trade in. Certain model definitions allow for 11 indicator variables for the target date offset (from -5 days to +5 days) and 20 lagged normalised volumes.

Figure 5.9 contains the relative volumes for the negative target date offsets on the left-hand side column, corresponding to Panels A – F, and the positive target date offsets, corresponding to Panels G – L. The figure illustrates the slight negatively skewed distribution of the relative volume on the MSCI quarterly index review day only.

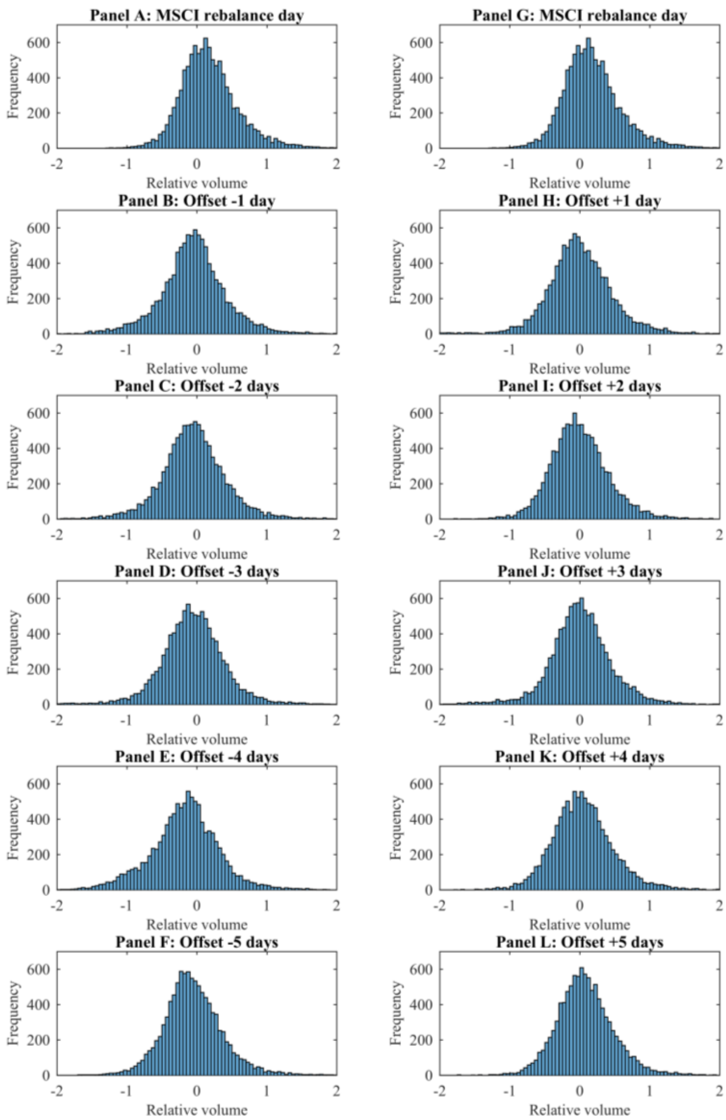

Figure 5.9: Relative volume distribution for positive target date offsets (Panels A – F) and negative target date offsets (Panels G – L) relative to the MSCI rebalances.

Next, we examine the results of the futures expiry and MSCI rebalance models and inspect a series of aspects regarding the coefficients and feature sets of these models.

5.6.3 Volume Autoregression

A constant volume autoregression is reported among the futures expiry and MSCI rebalance models. There is a significantly lower cross-validation MSE associated with the models fit with 20 lagged normalised volumes, as outlined in Table 5.9.

Table 5.9
Comparison of the presence and absence of lagged volumes.

Model	Lagged volumes	Observations	CV MSE
Futures expiry	Yes	32,408	0.17230
	No	32,408	0.21746
MSCI rebalance	Yes	10,298	0.14490
	No	10,298	0.17587

5.6.4 Target Date Offset

Fitting the observations for all the offsets that we considered (i.e. -5 trading days to +5 trading days, relative to the expiry/rebalance day) and including them into a model with 11-indicator variable for the target date offsets significantly increases the cross-validation MSE, which is reported in Table 5.10 for models fit with and without target date offsets.

Table 5.11 outlines the large volume associated with the expiry day and the two days prior to the expiry day in the reduced model for futures expiry; there is a significantly positive correlation between trading volume and the MSCI rebalance day indicator. None of the days prior to or after the MSCI rebalance has any significance in terms of predicting the volumes. These coefficients represent the contribution of each feature to the trading volume and do not reflect the phenomenon documented in the previous randomisation tests, where the futures expiries are associated with high trading volumes from 4 days before the expiry date and until 2 days after the expiry date, and MSCI rebalances cause higher volumes on the day before the rebalance and on the rebalance effective date.

Table 5.10
Comparison of the presence and absence of offsets.

Model	Target date offset	Observations	CV MSE
Futures expiry	Yes	32,408	0.23770
	No	32,408	0.21746
MSCI rebalance	Yes	10,298	0.20140
	No	10,298	0.17587

Table 5.11
Target date offset coefficients.

Model	-5 days	-4 days	-3 days	-2 days	-1 day	Expiry/rebalance day	+1 day	+2 days	+3 days	+4 days	+5 days
Futures expiry	0.14		0.16	0.20	0.22	0.36	-0.01	0.10	0.09	0.08	
MSCI rebalance		-0.12	-0.01			0.23	0.05	0.05	0.07	0.09	0.13

Based on the previous empirical findings, we fit a futures expiry model and an MSCI rebalance model with 20 lagged normalised volumes and without offsets (i.e. considering only the futures expiries and the MSCI rebalances as target dates).

5.6.5 Trading Volume on Stock Index Futures Expiry Dates

The regression coefficients for the reduced and full futures expiry models are summarised in Table 5.12, except for the coefficients for the 20 lagged normalised volumes. There is incredibly high variability among the coefficients from the full model to the coefficients of the reduced model, indicating numerical instability. The 7 'expiry index' predictors consist of indicator variables which are set to 1 if a given stock is the constituent of this index whose expiry day relative volume is the target variable. There is certainly strong multicollinearity, reflected by the zero coefficients of DAX and AEX in the full model. We conclude that we cannot discriminate between the expiry indices of the stocks. There is a moderate feature in terms of quarterly expiry indices against monthly expiry indices, but it is not very salient. Figure 5.10 illustrates the relative volume distribution on the target dates (i.e. the index expiry dates) and on the control dates (i.e. dates with no expiry, falling on the same day-of-the-week as the index expiry, with an offset up to two weeks relative to the expiry date) which were previously generated in the futures expiry randomisation test. We observe strong positive effects driven by the expiry day. The FTSE and DAX exhibit conspicuous expiry day effects. The selected variables in the reduced model and the zero-valued coefficients of DAX and AEX in the full model are most probably caused by multicollinearity among the predictors.

Table 5.12
Futures expiry model coefficients.

Model	Samples	CV MSE	Intercept	Expiry index							Quarterly expiry
				FTSE	DAX	CAC	FTSE MIB	IBEX	AEX	OMXS30	
Full model	32,408	0.172301986	0.17	-0.47	0.00	-0.03	-0.01	-0.01	0.00	-0.14	0.60
Reduced model	32,408	0.172297765	0.16		0.47	-0.02			0.01	-0.13	0.13

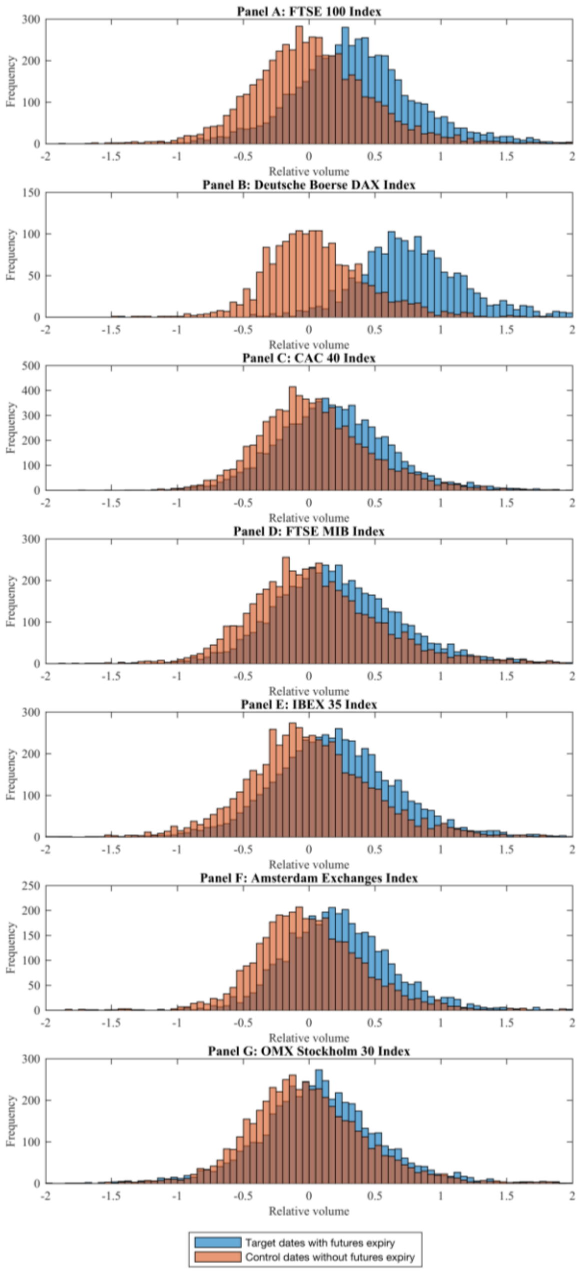

Figure 5.10: Relative volume distribution for the target and control dates for the expiry of each stock index analysed.

5.6.6 Trading Volume on MSCI Rebalance Dates

Table 5.13 outlines the coefficient values for the ‘trading *country code*’ features for the MSCI rebalance reduced and full models. The models are trained with 20 lagged normalised volumes and no target date offsets (i.e. we only consider the MSCI quarterly review dates). The ‘trading country’ predictors are indicator variables denoting the exchange country of each stock that is part of the MSCI index constituent list. We argue that there is no clear discrimination by country of the effect magnitude of MSCI rebalance on the stock volume. The coefficients have high variability between the reduced and full models, which is likely caused by multicollinearity (e.g. Italy’s and Sweden’s coefficients are zero-valued in the full model, while they experience a great increase in the reduced model). The MSCI rebalance randomisation test performed in the Randomisation Analysis section provides evidence of a significantly greater trading volume on MSCI rebalances.

Table 5.13
Pan-European MSCI rebalance model – reduced feature set (selected rebalance-related features exhibited).

Model (Samples)	CV MSE	Intercept	Trading country														
			AT	BE	CH	DE	DK	ES	FI	FR	GB	IE	IT	NL	NO	PT	SE
Full model (10,298)	0.144950458	0.41	-0.06	-0.11	-0.22	-0.24	-0.15	-0.18	-0.21	-0.17	-0.24	0.07	0.00	-0.27	-0.21	-0.20	0.00
Reduced model (10,298)	0.144899932	0.18	0.18	0.13			0.09	0.06		0.07		0.31	0.24	-0.03	0.03	0.03	0.23

5.6.7 Multi-Step Ahead Analysis

Multi-step ahead predictions are proposed besides the standard one-step ahead prediction, in order to allow traders to plan their portfolios by predicting an expiry day effect on a stock’s trading volume. Supposing one wants to predict the impact of the expiry day effect on volume in n days’ time, then one computes the benchmark volume between $(t - n)$ and $(t - 20 - n)$ and compares it against the volume on the expiry/rebalance day (i.e. V_0) in order to train the model. All the n-step ahead expiry/review day models are fit for each step size n, between 1 day and 6 days, and dimensionality reduction is performed on these full models.

The multi-step ahead models perform similarly to the 1-step ahead analysis, for n ranging from 2 to 6. Their reduced models have similar feature sets to the 1-step ahead analysis. The cross-validation MSE is directly proportional with the step size and there is a constant trend of increasing the MSE as the prediction step ahead lag grows, as described in Table 5.14.

Table 5.14
Comparison of the cross-validation MSE between 1-step ahead and multi-step ahead reduced models.

Model name	1-step ahead	Multi-step ahead				
		2	3	4	5	6
Futures expiry	0.17230	0.20178	0.21114	0.21923	0.22700	0.23228
MSCI rebalance	0.14490	0.16417	0.17407	0.17671	0.17879	0.18616

5.7 Discussion

The empirical evidence provided by this study supports a futures expiry effect and an MSCI rebalance effect, corresponding to an increase in trading volume for the constituents of these indices. The study investigates the European equity markets using a comprehensive pan-European stock universe of almost 500 stocks, with 32,408 observations for the stock index futures expiry analysis, and 10,298 observations for the MSCI rebalance analysis, which span almost 16 years. This study complements the existing literature by providing a pan-European empirical study for the expiry day effect on liquidity. We first examine the effect existence and then explore a predictive model. The randomisation tests are an instance of the methodological rigour of this study, while fitting a number of models by applying stepwise regression represents a methodological novelty in finance.

The stock index futures expiries intensify the trading activity four days prior to the futures expiry (i.e. starting on the Monday of the expiry week) and last two days after the futures expiry; the trading volumes start decreasing back on the third day after the expiry day. Similarly, higher trading volumes can be observed on the previous day before the rebalance date and on the MSCI rebalance effective date itself, reaching normal trading levels on the trading day after the quarterly review date. This study confirms that equity markets are in a rather steady state, but the market dynamics differ on some periodic notable events, which have been investigated in this study in order to document the temporal factors driving trading volume. The results are validated by the initial randomisation tests and the robustness of the approximately 500 European stocks data sample, covering the most recent 15 years of daily market data and expiry events (i.e. futures expiries and MSCI quarterly reviews).

We investigate whether it is the Friday effect or the Friday futures expiries that drives the trading volume up and we provide evidence of a strong futures expiry effect. Furthermore, we analyse whether the MSCI rebalances can explain the end-of-month larger volumes, but we conclude that the magnitude of the MSCI quarterly reviews is not sufficient to cause a generalised increase in volumes at the end-of-month throughout the year. There is a potential end-of-month effect itself, which is driven by various factors that are well-documented in the literature, e.g. buying pressure around the end-of-month, standardisation of payments around the turn-of-the-month, or the release of major US macroeconomic news.

Trading volume exhibits a constant and significant autoregression among the futures expiry and MSCI rebalance models. The study comes to an end by proposing a multi-step ahead prediction framework, which could be adapted in the industry such that traders and hedge fund managers could anticipate an expiry day effect by planning their portfolio in advance based on the predicted trading activity.

6. Developing a Volume Forecasting Model

Volume prediction is critical for optimising the market impact of an order. This study builds a series of out-of-sample models to predict trading volume in European markets using different statistical methods. The analysis considers a series of aspects, such as special events (e.g. MSCI rebalances, futures expiries, or cross-market holidays), the day-of-the-week effect, training window types (e.g. different window lengths or growing/moving window approaches), the volume-price relation asymmetry etc., in order to perform contextual one-step ahead prediction. We investigate the prediction error for each calendar circumstance in order to infer a cross-stock switching model for volume prediction. This switching model consists of 42 event-oriented sub-models, which are specifically fit on disjoint data sets, and provides the best performance overall. Having the goal to optimise the error stability, we conclude the study by proposing a stock-specific out-of-sample metamodel that is fit by selecting an initial stock-specific model that has the best performance for the most recent observations.

6.1 Introduction

Measuring trading performance is a challenging research area, but there are certain factors that have a clear influence on the overall trading performance, such as the market impact, which is the effect caused by a market participant who buys or sells shares, consisting in the extent to which the price goes upward for a buy order or downward for a sell order. The market impact cost is defined as the difference between the actual price and the hypothetical price provided that the order was not created (Johnson, 2010). Market impact can move the prices adversely, leading to decreased profits or turning profitable strategies into losing strategies.

The execution style of an order drives the extent of an order's market impact. An example of a trading strategy to decrease the market impact is when an investor needs to break down a large sell order into smaller orders over a longer period in order to trade slowly with a low market impact. Therefore, predicting the trading volume as a measure of liquidity is of vital importance to forecast the expected market impact.

The aim of this study is to propose a switching volume prediction model by fitting a variety of models that employ different machine learning methods and considering endogenous and exogenous variables that may potentially impact the trading volume. This is motivated by the importance of optimally sizing an order for minimising the market impact and ultimately improving the trading performance. Market participants who size their orders incorrectly can either over-participate by producing excessive market impact or under-participate by creating opportunity cost and price uncertainty. Therefore, predicting the trading volume helps better determine the degree of participation in the market.

The primary focus of this study is to fine-tune the models and identify the optimal model given the market context at a certain point in time, in order to achieve optimal prediction accuracy and model stability. We are investigating the error breakdown by different model types and days that matter (e.g. holidays, expiries, days-of-the-week etc.).

Each stock exhibits different levels of trends, volatility, and magnitude in their market data. Consequently, we perform stock-specific predictive modelling throughout this study by independently training a variety of window-based predictive models for seven machine learning techniques: ordinary least squares, stepwise regression (i.e. ordinary least squares with sequential feature selection), ridge regression, lasso regression, k-nearest neighbours with arithmetic average, k-nearest neighbours with inverse distance weighting, and support vector regression. For each statistical method, we iterate every stock in our pan-European stock universe consisting of 2,353 stocks, every training window type (i.e. moving/sliding vs. growing) and every window size (i.e. 1-month, 3-month, 6-month, 1-year, 2-year windows). We also train three models for special events (i.e. cross-market holidays, MSCI rebalances and futures expiries) using the entire stock universe, although they are ultimately used to make stock-specific predictions. We fit these models in isolation and aim to determine a performance metric for each method and window type.

Eventually, we shift from a static process to an adaptive process and construct a switching dynamic model, which switches between these models based on the current context (e.g. regular trading day, cross-market holiday, futures expiry, MSCI rebalance, certain day-of-the-week etc.). The proposed model is a virtually switching model as it does not switch per se. We are post-processing the model performance and investigate the performance metrics by breaking down the errors by: day-of-the-week, cross-market holidays, futures expiries, MSCI rebalances etc. This leads to the metamodel, which is a stock-specific out-of-sample model that selects the best initial stock-specific model on a 1-month and a 3-month rolling window basis, depending on the recent performance of the initial stock-specific models that are trained independently of each other.

The rest of the study is structured as follows: section 2 reviews the key findings that led to our model choice in this study (e.g. the volume-price relation asymmetry, the day-of-the-week effect, the expiry day effect, and the cross-market holidays effect) and outlines the methods employed in this analysis; the market and calendar data sets are introduced in section 3; section 4 provides the analysis approach and briefly describes the high performance computing design of this computationally expensive analysis, followed by a methodological introduction of the cross-stock models and the stock-specific models; this is followed by section 5, which presents the main findings of this study, including a performance breakdown of the models, and introduces the switching model and the out-of-sample stock-specific metamodel; eventually, section 6 provides a conclusion of this analysis and discusses the obtained results.

6.2 Background

In our previous studies, we provided empirical evidence for the volume-price relation and its asymmetry, and the existence of the day-of-the-week effect, the expiry day effect and the so-called 'cross-market holiday' effect in relation with trading volume. These findings are summarised below and are followed by a review of the statistical methods employed in this analysis.

6.2.1 Volume-Price Relation and Asymmetry

The price-volume relation is of great importance for this study as most of the behavioural literature focuses on the impact of certain anomalies on price returns, while trading volume is the main focus of this study. Price changes represent the market response to new information, whereas the trading volume indicates the level of information disagreement among investors (Beaver, 1968). Although the literature on a potential relation between price changes and volume is far from homogenous, there is a large proportion confirming a positive correlation between trading volume and price changes (Harris & Raviv, 1993) (Hong & Stein, 2007). We provided empirical evidence that trading volume is correlated with historical price indicators (i.e. intraday range and intraday return for the previous day, and overnight return for the previous night, which acts as a proxy for the opening auction volume, i.e. more recent information) and that volume exhibits autoregression, where we employed lagged time series volume data (i.e. raw past observations) and also smoothed lagged time series (i.e. moving average of past observations, which acts as a low-pass filter effect in the data). The formulae for the intraday return, intraday range and overnight return are outlined below, where n is the number of intervening nights, t_0 is the day for which we predict the trading volume and t_{-1} is the previous trading day, whose price and volume information is available:

$$\text{intraday return log ratio: } p_{\text{intradayRtn}} = \log \frac{p_{t_{-1}}^{\text{close}}}{p_{t_{-1}}^{\text{open}}} \tag{6.1}$$

$$\text{intraday range log ratio: } p_{\text{intradayRng}} = \log \frac{p_{t_{-1}}^{\text{high}}}{p_{t_{-1}}^{\text{low}}} \tag{6.2}$$

$$\text{overnight return log ratio: } p_{\text{overnightRtn}} = \frac{1}{n} \log \frac{p_{t_0}^{\text{open}}}{p_{t_{-1}}^{\text{close}}}. \tag{6.3}$$

In general, there are two key representations of the volume-price relation, where trading volume is positive correlated either with the magnitude (i.e. absolute value) of the price change (Assogbavi & Osagie, 2006), i.e. $|\Delta p|$, or with the price change per se (i.e. the raw value of the price change), i.e. Δp (Karpoff, 1987) (Ying, 1966). The asymmetric relation in the latter representation exhibits a volume/price change ratio that is different in magnitude for upticks than for downticks. Equation (6.4) shows the levels of volume based on the sign of the price change, compared to the symmetric model in Equation (6.5):

$$\text{asymmetry: } (v_t|\Delta p_t^+) > (v_t|\Delta p_t^-) \text{ or } (v_t|\Delta p_t^+) < (v_t|\Delta p_t^-) \tag{6.4}$$

$$\text{symmetry: } (v_t|\Delta p_t^+) = (v_t|\Delta p_t^-). \tag{6.5}$$

We provided empirical evidence for the price-volume relation asymmetry, which was exhibited in over 70% of the analysed European stocks; there is a moderate overnight asymmetry, which is almost evenly distributed, and a more salient intraday asymmetry (in approximately 60% of the stocks).

6.2.2 The Day-of-the-Week Effect

The day-of-the-week effect consists of certain trends associated with a particular day-of-the-week. The most broadly studied day-of-the-week effect is the weekend effect (French, 1980) (Gibbons & Hess, 1981) (Jaffe & Westerfield, 1985) (Pettengill, 2003) (Cross, 1973) (Dubois & Louvet, 1996) (Harris, 1986) (Abraham & Ikenberry, 1994), or Monday effect, where the closing price on Monday is lower than the closing price of the previous Friday. These results are intriguing as they are opposite to the expectation of higher returns on Monday, as its returns reflect three consecutive days. The weekend effect has been widely documented in conjunction with price changes. There are very few studies investigating the relation between the day-of-the-week effect and trading volume. For example, Berument and Kiymaz (2001) found day-of-the-week anomalies in both returns and volatility, with the highest volatility on Friday and the lowest on Wednesday, while Lakonishok and Maberly (1990) found a relative increase in the trading activity of individuals on Mondays.

We reported a clear improvement of the trading volume prediction model when adding the day-of-the-week features. The indicator variable for Monday improves the model in more than 75% of the cases, having predominantly negative coefficients, despite the fact that we divide the overnight return by the number of intervening nights, which suggests that the negative coefficient for Monday is not a corrective factor and that there is simply less activity on Mondays. Fridays improve the volume model in 45% of the stocks and their coefficients are surprisingly mostly negative, even if the traditional definition of the weekend effect states that the Friday volume and prices are usually higher than those of the following Monday.

6.2.3 The Expiry Day Effect

The expiry day effect exhibits higher trading volume and abnormal volatility around the close on expiry days for futures and options (Stoll & Whaley, 1997) (Sukumar & Cimino, 2012) (Chow, et al., 2003) (Sadath & Kamaiah, 2011) (Pope & Yadav, 1992) (Vipul, 2005) (Chiang, 2009), and for MSCI quarterly reviews (Chakrabarti, et al., 2005).

Following these findings, we further analysed the effect of periodic events on the trading volume, while investigating the stock index futures expiries and MSCI quarterly index reviews in the pan-European markets. The stock index futures expiries occur on the third Friday of each expiry month or on the previous trading day in case that Friday is a bank holiday. The futures contracts are traded either quarterly (i.e. March, June, September and December) or monthly. The indices of Morgan Stanley Capital International (MSCI) are updated quarterly in order to reflect the up-to-date state of the financial markets. The constituent list of these indices changes close to the last trading day of the four rebalancing months: February, May, August and November. We reported the existence of the futures expiry effect and the MSCI rebalance effect, both leading to a surge in trading volume for their index constituents. The trading volume increases significantly during the four days in the run-up to the expiry, lasts two days after the futures expiry, and then returns to normal levels of trading activity starting on the third trading day after the expiry day. The MSCI rebalances exhibit a similar trend, causing surges in the trading volume on the day before the the review day and on the effective rebalance date. We discriminated between these two instances of the expiry day effect and the Friday and end-of-month effects and concluded that the futures expiry effect is essentially causing the so-called Friday effect. However, we could not find enough evidence that the MSCI quarterly reviews could drive the anecdotal end-of-month effect; the trading volumes on the four months with MSCI quarterly reviews are significantly different from those on the adjacent months, but their magnitude is not sufficiently large in order to explain the end-of-month effect throughout the entire year.

6.2.4 The Cross-Market Holiday Effect

In a previous study, we coined the term 'cross-market holiday effect', which refers to the anecdotal evidence of lower volumes in a particular country when one or more external markets are not trading. There are only a couple of studies investigating this effect although they focus mainly on the subduing effect of the US holidays on other markets, such as Canada (Cheung & Kwan, 1992) and Europe (Casado, et al., 2013). We documented a salient cross-market holiday effect when a dominant market is on holiday or when most of the European markets are shut. Since the UK is Europe's largest market, we examined whether it is actually the Monday effect that drives down the volumes, as most of the bank holidays fall on a Monday in the UK. However, we reported strong evidence that the Mondays with at least one cross-market holiday have significantly lower volumes that the other Mondays.

Throughout our previous in-sample analyses on the day-of-week, expiry day and cross-market holiday effects, we observe strong evidence of volume autoregression. Given the results of our previous independent studies, we aim to integrate their findings in an out-of-sample study. Here, we aim to build a virtually adaptive model, which fits a number of models in parallel and switches from one underlying model to another, by taking into account the event dates (e.g. futures expiry, MSCI rebalance, certain day-of-the-week etc.) when we expect the markets to behave significantly different. We also raise additional questions on the optimal training window and the appropriate methodology. We are empirically testing a number of statistical methods in order to understand how the performance of each method is affected and to explore the relationship between trading volume and event dates in a predictive framework.

6.2.5 Methodology Review

In this section, we review the basic principles for the supervised learning models that are employed in this study. There are seven different statistical methods that are fit simultaneously and independently in order to predict the one-step ahead trading volume. We start with the ordinary least squares (OLS); it is the most basic model and estimates the variable coefficients of a linear regression model by minimising the sum of the squared distances between the predicted values and the observed values.

Feature selection can be applied after a model is fit using OLS, by performing stepwise regression. This can be achieved through forward selection, backward elimination, or bidirectional elimination. We chose forward selection for the second method of this study (i.e. stepwise regression), which adds new variables having p-values that are less than a given improvement measure. We start from a reduced model consisting of the intercept, the lagged volumes, and the smoothed lagged volumes, allowing the model to pick the most informative price log-ratio and day-of-the-week features. The rationale for using forward selection is driven by the design of the day-of-the-week categorical variable as five dummy variables. Generally, a categorical variable having n values is encoded as $n - 1$ dummy variables, although in this study, the day-of-the-week dummy variables are mutually exclusive since the aim is to perform feature selection and extract the variables with the highest statistical significance for volume prediction and this is conducted in a feature selection framework. We preferred forward selection to backward elimination because of the potential collinearity problems; adding a collinear variable could make matrix inversion impossible when determining the optimal beta.

The next two techniques employ regression shrinkage methods, namely ridge regression and lasso regression. Linear regression relies on the independence of the model variables and therefore the matrix $(X^{\mathrm{T}}X)^{-1}$ becomes close to singular when the design matrix X has columns that exhibit an approximate linear dependence. As a result, the least squares estimate shown in Equation (6.6) produces a high variance because of its sensitivity to random errors in the observed response variable y:

$$\hat{\beta} = (X^{\mathrm{T}}X)^{-1}X^{\mathrm{T}}y. \tag{6.6}$$

Ridge regression, or L_2 regularisation, addresses the problem of multicollinearity by estimating the regression coefficients using Equation (6.7), where λ is the ridge parameter and I is the identity matrix. This method introduces bias, but reduces the variance of the coefficient estimates, producing a lower mean squared error (MSE) compared to the least squares estimates. We start by identifying the optimal value for λ (i.e. the ridge parameter) that minimises the cross-validation error, by using a two-section search consisting of grid search and followed by the bisection method (also known as binary search). The grid search traverses 21 consecutive values of λ in logarithmic space, from -10 to 10 and cross-validates the data set for each λ. The value with the minimum average MSE across the grid search is then passed to the bisection method, whose initial left and right points are calculated as $\lambda - 1$ and $\lambda + 1$, respectively, which are also expressed in logarithmic space. The bisection method runs until at least one of the following three tolerance criteria is not met anymore: minimum delta (i.e. minimum change in λ) = 0.1%, minimum error change = 10^{-11}%, and maximum number of iterations = 20. The ridge coefficient estimates are restored to the original scale of the data. This transformation also computes the parameter for the constant term (or intercept) and provides a model that is more useful for making predictions, unlike a model with standardised coefficients:

$$\hat{\beta} = (X^{\mathrm{T}}X + \lambda I)^{-1}X^{\mathrm{T}}y. \tag{6.7}$$

Lasso (Tibshirani, 1996), or L_1 regularisation, is another regularisation method that is similar to ridge regression. The main difference is that when the penalty term λ increases, more coefficients are set to zero, whereas ridge regression sets the coefficients close to zero, but not exactly zero. The lasso estimator produces a smaller model with fewer predictors. Based on the resulting model, lasso can be regarded as an alternative to the second methodology described above, i.e. stepwise regression, and other dimensionality reduction techniques. For a nonnegative regularisation parameter λ, lasso solves the regularisation problem in Equation (6.8), where N is the number of observations, y_i represents the response variable for observation i, x_i is the observed data for observation i consisting of a vector of p values that correspond to each predictor, β_0 is a scalar for the intercept coefficient, and β is a p-vector for the other model terms' coefficients:

$$\min_{\beta_0,\beta}\left(\frac{1}{2N}\sum_{i=1}^{N}(y_i - \beta_0 - x_i^{\mathrm{T}}\beta)^2 + \lambda\sum_{j=1}^{p}|\beta_j|\right). \tag{6.8}$$

We implemented lasso regression in a similar manner to ridge regression. The optimal λ is determined through 10-fold cross-validation using a two-section search (i.e. grid search in logarithmic space between -10 and 10, followed by binary search for the same set of tolerance criteria that we defined for ridge regression). MATLAB's implementation of lasso regression fits the regularised regression without a constant term, although its coefficient is returned in the 'FitInfo.Intercept' variable, and is eventually appended to the coefficient vector corresponding to the model's predictors.

The k-nearest neighbours (kNN) technique is a non-parametric method belonging to the instance-based learning family, which can be used for both classification and regression problems, where the function is only approximated locally. It is memory-based and requires no model to be fit, i.e. it memorises all of the observations and predicts the target variable based on the chosen similarity measure, which is typically a distance function. The most common distance metric for continuous variables is the Euclidean distance shown in Equation (6.9), whereas the Hamming distance, represented in Equation (6.10), is typically used for binary/categorical variables and is calculated as the number of instances where two observations are different:

$$d_{\text{Euclidean}} = \sqrt{\sum_{i=1}^{k}(x_i - y_i)^2} \tag{6.9}$$

$$d_{\text{Hamming}} = \sum_{i=1}^{k}|x_i - y_i|. \tag{6.10}$$

The algorithm retrieves the k memorised examples that are the most similar to the one that is used for the current prediction using an appropriate distance function. The kNN method does not have any costs associated with the learning process as there is no model inferred and, because of this, it is also known as 'lazy learning', as the entire cost of this technique consists of the prediction computation; there are no assumptions about the characteristics of the data, although the lack of any learning costs makes kNN impossible to be interpreted as there is no description of the learnt concepts. Moreover, the accuracy of kNN can be significantly impacted by the presence of noisy or irrelevant features. The basic version of kNN is the 1-nearest neighbour estimate, whose bias if often low, but the variance is high. An interesting property of the nearest-neighbour is that its error rate is never more than twice the minimum achievable error rate of an optimal classifier (Hastie, et al., 2011) (Bishop, 2007).

In order to identify the optimal value of k, we perform 10-fold cross-validation and we pick the value of k that minimises the cross-validation average error. A small value of k means that noise will have a higher impact on the results, whereas a large value of k is computationally expensive and signals a highly non-linear and noisy structure. The number of neighbours can be regarded as a measure of noisiness; for example, 1NN is an indication of clear data. In general, a larger value for k is more precise, although the boundaries within the feature space become blurred. A few authors (Duda, et al., 2000) (Hassanat, et al., 2014) suggest an empirical rule-of-thumb, and setting k equal to the square root of the number of instances, $k = n^{1/2}$, as a starting point. We also attempted to apply PCA on the standardised variables in order to remove the correlations before running kNN, but it mainly dealt with the intercept only and did not improve the resulting model.

We begin by standardising each feature of the data set to have mean zero and variance 1, because the variables have different measurement scales and there is also a mixture of continuous and categorical/binary variables (Hastie, et al., 2011). This allows us to use the Euclidean distance as the nearest neighbours' similarity measure.

The following two methods represent slightly different implementations of kNN, which vary in their approach of aggregating the contribution of the identified neighbours. The first approach is kNN with arithmetic mean, which treats all of a point's neighbours equally and computes the prediction as the average of the target variable of the k nearest neighbours, as shown in Equation (6.11). The second approach is kNN with inverse distance weighting, where the neighbours are assigned weights based on their distance from the prediction point, such that the nearer neighbours contribute more to the average compared to the further neighbours. This method assigns a weight to each neighbour, which is equal to the inverse of its distance to the prediction point; this weighted average is illustrated in Equation (6.12). The algorithm finds the k nearest observations using the Euclidean distance metric, then calculates the inverse distance weight of each neighbour and normalises the inverse distances such that their sum is equal to one. Finally, the method computes the weighted average of the k neighbours using their inverse distance weights. We implemented both methods in order to better understand the data structure. Using the inverse distance weighting could potentially lead to a large number of neighbours being identified, where most of them could have extremely small weights that would not influence the prediction significantly and would simply introduce more noise to the model. If this is the case, a parsimonious model could be identified by using the arithmetic mean and implicitly assigning equal weights to the neighbours:

$$\hat{y} = \frac{1}{k}\sum_{i=1}^{k} y_i \tag{6.11}$$

$$\hat{y} = \sum_{i=1}^{k} d'^{-1}_{i} \cdot y_i. \tag{6.12}$$

Both techniques begin by identifying the optimal value for k. This is accomplished by performing 10-fold cross-validation grid search for k ranging between 1 and 49, with a step size of 1, and between 50 and 100, with a step size of 5. Once the optimal value of k is found, the algorithm retrieves the closest k neighbours by standardising the training data set (or computing the z-score, such that each variable has unit variance and zero mean), and then standardising the test set using the mean and variance obtained from the training set. Then the predicted value is computed either by the mean of the target variable of the k neighbours for the kNN with arithmetic mean method, or by weighing the neighbours taking into account their normalised inverse distances.

The last method is based on support vector machine (SVM) analysis (Cortes & Vapnik, 1995) (Vapnik, 1999), which is a popular method that was traditionally employed for classification; a version of SVM for regression was introduced later (Drucker, et al., 1997) and is called support vector regression (SVR). SVM is non-parametric as it relies on kernel functions. SVM produces non-linear boundaries by creating a linear boundary in a transformed representation of the feature space (Hastie, et al., 2011). SVM maximises the margin around the separating hyperplane and defines the solution in terms of a small subset of training samples, which are called the support vectors, i.e. the training data points that are closest to the decision hyperplane and that are most difficult to classify. SVR produces a model that depends only on a subset of the training data, since the model's cost function ignores the training data points that are close to the model prediction. We implemented the sequential minimal optimisation (SMO) algorithm (Platt, 1998), which does not require a numerical optimisation algorithm or matrix computation and storage, because it divides a very large quadratic programming (QP) optimisation problem into a series of smallest possible QP problems that are solved analytically; this eliminates the need for a time-consuming numerical optimisation as an inner loop. The SMO algorithm is fast, easy to implement, and provides better scaling properties. The algorithm also flags the day-of-the-week features as categorical predictors. The SVR models in this study use the Gaussian kernel function. Keerthi and Lin (2003) proved that the linear kernel is a degenerate version of the Gaussian kernel, also called radial basis function (RBF), and therefore the linear kernel would never have a better accuracy than the Gaussian kernel.

Our SVR method implements the linear epsilon-insensitive SVM (ε-SVM) regression, which is also called the L_1 loss. By using the predictor variables and the observed response variables, the goal of ε-SVM is to identify a function $f(x)$ such that its deviation from y_n is no greater than ε for each training point and is as flat as possible (MathWorks, 2016). There are two main formulations for the optimisation problem: the primal formula and the dual formula. The primal formula consists of a convex optimisation problem, where it is possible that there is no function that satisfies the constraints for all points. This issue is overcome by introducing slack variables, which help deal with infeasible constraints and lead to the objective function, also known as the primal formula. The primal formula includes the box constant, which acts as a regularisation method in order to prevent overfitting; this imposes a penalty on all of the observations lying outside the ε margin and determines the trade-off between the flatness of the function $f(x)$ and its tolerance. The SVR loss is calculated based on the distance between the observed target variable y and the ε boundary. The dual formula provides a computationally simpler solution to the primal formula; it employs Lagrange multipliers in order to transform the optimisation problem into a form that can be solved analytically. The optimal values of these two problem formulations are not necessarily equal and their difference is known as the duality gap. The solution of the dual problem is used exclusively when the problem is convex and meets a constraint qualification condition.

6.3 Data Set

We compile one of the most comprehensive pan-European data sets, ranging from 1st January 2000 to 10th May 2015. It consists of over 7 million observations of daily market data for 2,353 stocks, 3,039 bank holidays for 22 countries, 1,042 stock index futures expiries for 7 indices, and 49 MSCI quarterly review dates, along with a historical log of 1,420 leavers and joiners for the investigated futures and MSCI indices. A great effort has been put into collecting, cleansing and processing the calendar data set due to the lack of a comprehensive database of bank holidays for financial markets.

6.3.1 Market Data

The market data contains daily observations consisting of the opening, closing, low and high prices and the trading volume for the constituents of the 31 most important European indices. The data set was retrieved from Thomson Reuters and was further processed. We compute the consolidated trading volume for each stock by retrieving the corresponding trading volume across the main European multilateral trading facilities (MTFs), i.e. BATS, CHI-X and Turquoise, and adding the MTF volume to the trading volume of the primary exchange. The resulting consolidated volume is used across this study in order to better reflect the true liquidity of a stock. The analysis discards the stocks with less than 100 trading days. South Africa was included in the analysis due to its close ties with the European financial markets. The processed market data covers 21 European countries and Table 6.1 outlines the number of stocks and their daily observations for each country.

Table 6.1
Stock universe – Breakdown by country.

Country Name	Country Code	Number of Stocks	Number of Observations
Austria	AT	32	98,179
Belgium	BE	62	205,414
Czech Republic	CZ	5	14,491
Denmark	DK	43	144,352
Finland	FI	130	390,209
France	FR	346	1,117,220
Germany	DE	176	539,142
Greece	GR	61	209,103
Hungary	HU	4	15,311
(Republic of) Ireland	IE	43	100,910
Italy	IT	111	330,609
Netherlands	NL	46	157,156
Norway	NO	69	172,562
Poland	PL	65	162,509
Portugal	PT	18	53,449
South Africa	ZA	42	139,568
Spain	ES	61	179,410
Sweden	SE	158	462,935
Switzerland	CH	104	339,998
Turkey	TR	130	412,273
United Kingdom	GB	647	1,952,265

6.3.2 Calendar Data

The market data is augmented by a comprehensive list of event dates, which can be classified as bank holidays and expiry days (i.e. stock index futures expiries and MSCI rebalances). These special events are expected to impact on the normal state of the financial markets and cause non-stationarity in trading volume.

6.3.2.1 Bank Holidays

The data set for bank holidays is customised specifically for the financial markets and can be different in certain instances from the official national public holidays for a given country: when an exchange venue is owned by a company which is based in another country (e.g. Euronext) and enforces a different trading calendar, when a trading venue is located in a region with additional holidays, or when unforeseeable events occur (e.g. Hurricane Sandy, 11th September Terrorist Attacks etc.). This calendar is an accurate reflection of the trading state of the US and the pan-European exchanges, consisting of 22 countries, whose distribution of non-trading days is outlined in Table 6.2. The United States of America was included in the data set since it is a dominant financial market, whose magnitude could potentially influence the European liquidity. The non-trading calendar was meticulously compiled from scratch and multiple sources (e.g. the trading calendar on the exchanges' websites and public holidays from www.timeanddate.com) were used to make decisions on the final outcome. These were double-checked with the empirical trading calendar resulting from the market data, which truly proved whether an exchange has been trading on a particular day. The accuracy of this calendar was vital to perform a cross-market holiday model and had to be manually constructed because there was no such trading calendar available; there are very few such calendars, although their information is either incomplete or they contain conflicting information.

Table 6.2
Frequency table of non-trading days per country.

Country Name	Country code	Number of Non-Trading Days
Austria	AT	202
Belgium	BE	84
Czech Republic	CZ	152
Denmark	DK	166
Finland	FI	153
France	FR	82
Germany	DE	111
Greece	GR	179
Hungary	HU	171
Ireland; Republic of	IE	121
Italy	IT	112
Netherlands	NL	81
Norway	NO	154
Poland	PL	158
Portugal	PT	100
South Africa	ZA	170
Spain	ES	116
Sweden	SE	153
Switzerland	CH	146
Turkey	TR	157
United Kingdom	GB	127
United States	US	144

There are also country-specific characteristics for generating the public holidays calendar. For example, if a public holiday falls on a weekend, different countries substitute it with the previous trading day (e.g. New Year's Eve in Austria and Belgium), with the following day, or do not substitute it at all. Additional 'bridge' holidays can be observed in particular countries (e.g. Hungary and Poland), when a holiday falls on a Tuesday or on a Thursday, resulting in four-day weekends.

An illustrative example of the difference between the official public holidays and the non-trading calendar is on 1st May in the Netherlands, where the financial markets are shut despite the fact that 1st May is not a bank holiday. This is observed after the Amsterdam stock exchanged merged with the Brussels and Paris stock exchanges, in order to form the Euronext group. Similarly, the Belgian, Portuguese and French trading calendars changed after their main trading exchanges joined Euronext and therefore the public holidays between 1st May and Christmas Eve became regular trading days.

6.3.2.2 Expiry Days

The expiry day calendar incorporates periodic trading events which could be positively correlated with the trading volume, and consists of the futures expiries for seven liquid indices and the MSCI quarterly review for the MSCI International Pan Euro Price Index. By using the most liquid indices in Europe, this expiry calendar is an accurate representation of the main expiry dates in the European markets.

We retrieved the up-to-date constituents for these indices as of 11th May 2015, which represent the 'current constituents'. In order to create an accurate representation of the expiring indices at a given point in our analysis timeframe, we constructed a historical list of additions and eliminations for each index, which allowed the generation of a snapshot of a stock's constituent stocks. Table 6.3 outlines the number of constituents for each index, for both futures expiries and MSCI rebalances, where the 'historical constituents' column represents the number of previous stocks that were part of the constituent list of a given index before 11th May 2015, but which were subsequently eliminated, such that they are not a constituent anymore on 11th May 2015.

Table 6.3
Market data European indices for the futures expiry analysis and MSCI rebalance analysis.

Analysis Type	Index Name	Current Constituents	Historical Constituents	Location
Futures expiry	Amsterdam Exchanges Index	25	37	Netherlands
	CAC 40 Index	40	54	France
	FTSE MIB Index	40	51	Italy
	FTSE 100 Index	100	149	United Kingdom
	Deutsche Boerse DAX Index	30	37	Germany
	IBEX 35 Index	35	44	Spain
	OMX Stockholm 30 Index	30	33	Sweden
MSCI rebalance	MSCI International Pan Euro Price Index EUR Real Time	204	338	Europe

Stock Index Futures Expiries

There are 32,408 observations of stock index futures expiries for seven indices, whose expiries occur either monthly or quarterly (i.e. December, March, June and September) as follows:

- Monthly: CAC 40 Index Futures, FTSE MIB Index Futures, IBEX 35 Index Futures, Amsterdam Exchanges (AEX) Index Futures, and OMX Stockholm 30 (OMXS30);
- Quarterly: FTSE 100 Index Futures, and DAX 30 Index Futures.

The expiry occurs on the third Friday of the expiry month, or on the previous trading day when the third Friday is a non-trading day. The futures contract specifications were retrieved from Euronext (AEX and CAC 40), Eurex Exchange (DAX 30), London Stock Exchange (FTSE 100), Borsa Italiana (FTSE MIB), Bolsas y Mercados Españoles (IBEX 35) and NASDAQ OMX (OMXS30), in order to verify the expiry specifications for each index.

MSCI Quarterly Reviews

The MSCI rebalances have 10,298 observations across 16 countries: Austria, Belgium, Switzerland, Germany, Denmark, Spain, Finland, France, United Kingdom, Greece, (Republic of) Ireland, Italy, Netherlands, Norway, Portugal, and Sweden. Each stock's country represents the country where that stock is trading, e.g. the United Kingdom is defined as a Spanish stock's country if this stock is trading on the London Stock Exchange.

In general, the MSCI quarterly reviews are implemented on the last trading day of the February, May, August, and November quarterly cycle, although there are a few exceptions when the MSCI rebalance falls a few days before the end of the month. The MSCI quarterly review dates were double-checked with the quarterly index review documents from www.msci.com.

6.4 Predictive Modelling

We build a 1-step ahead out-of-sample model for predicting the trading volume, while fitting different supervised learning methods and examining event dates.

6.4.1 Analysis Approach

The methodological approach for constructing the predictive model is described in this section. For a given stock, these models predict the volume of the next day (i.e. the target date) based on past observations, employing a variety of machine learning methods and training window types. All of the models are fit with a constant term.

There are three cross-stock models for event dates (i.e. cross-market holidays, stock index futures expiries, and MSCI quarterly reviews), which are fit using normalised data from all of the relevant stocks. In the case of special events, very few training observations would be available for an individual stock, hence the necessity of aggregating the training points for multiple stocks. However, after learning the model on the normalised data set, the volume for each stock is predicted individually, by using the stock-specific benchmark volume that was used for normalising the stock's past volumes. The feature set of these cross-stock models includes 20 lagged volumes for each stock's observation, which are normalised by dividing them by their median, in order to remove any differences in magnitude across our stock universe.

Besides these three cross-stock models, there are seven stock-specific models, which are fit using different types of supervised learning methods: OLS, stepwise regression, ridge regression, lasso regression, kNN with arithmetic mean, kNN with inverse distance weighting, and SVR. We define an iteration as a fit model for every combination of stock, target date, learning method, and window type. For each iteration, these stock-specific models follow a similar training routine, starting by defining the 10-fold stratified cross-validation (CV) partitions from the beginning of the analysis, in order to conduct the entire iteration analysis on the same data partitions (e.g. when cross-validating potential values for method-specific parameters such as λ or k). The CV splits the data into 10 equally-sized partitions, while ensuring these are stratified by the binary indicator variables (i.e. the day-of-the-week binary features), such that these features are evenly distributed across the folds; its aim is to minimise the average mean squared error (MSE) throughout the 10 folds. The models can potentially contain 15 raw lagged volumes (i.e. autoregressive past observations) and 14 smoothed lagged volumes (i.e. moving average past observations), in order to explain the trading volume using recent time series. The iteration analysis identifies the optimal orders for the raw lagged volumes (or 'volume lags') and the smoothed lagged volumes (or 'volume windows'). It starts by fitting a linear regression for the lowest orders (i.e. volume lag 1, or volume window 2), then it increments the order by one, fits the second model, and compares the CV average MSE for these two models. If the higher order model performs better, the process is incrementally repeated for the next pair of orders (up to order 15), until the optimal order has been found, either when the higher order model has a larger error (and therefore the current model pair's lower order becomes the optimal one), or when the order reaches the maximum limit of 15. This incremental comparison of nested models is conducted independently for the volume lags and the volume windows.

When kNN, ridge regression or lasso regression are employed, the model proceeds to parameter calibration and runs grid search for k and performs a two-section search (i.e. grid search and binary search) for λ. These searches perform 10-fold cross-validation for each value. Then, all of the models proceed to feature construction and model training. The iteration analysis ends by testing the learnt model, i.e. computing the 1-step ahead prediction for the target date.

6.4.1.1 Training Windows

Each model is trained using two approaches: moving window and growing window. This helps understand whether a model relies only on recent data or whether it improves when more and more data points are used for training the model. These two approaches differ in the size of past observations when learning a model. When iterating the target dates of a stock for which predictions are made, the moving window approach trains the model using a fixed number n of past observations, starting from the most recent data point (i.e. the observation occurring right before the prediction 'unseen' data point) and going backward until n points are accumulated. Throughout the next iterations, the moving window gradually adds a newer observation and drops the oldest observation, whereas the growing window approach adds a newer observation without discarding any other observations. Therefore, the number of observations on the k^{th} iteration of a model is n for the moving window and $n + k - 1$ for the growing window.

There is a discrete number of sliding window sizes, whose representations are marked in brackets and are used when outlining the model results for this study: 1 month ('MW_1M'), 3 months ('MW_3M'), 6 months ('MW_6M'), 1 year ('MW_1Y'), and 2 years ('MW_2Y').

The growing window starts with a training size that is equal to the largest moving window size, i.e. 2 years, and is represented by 'GW'.

Each stock-specific learning method is trained using the five moving window types and the growing window, whereas the cross-stock models are fit using only the two largest moving window sizes (i.e. 'MW_1Y' and 'MW_2Y') and the growing window. The rationale of using only the 1-year and 2-year moving windows is driven by the significantly lower number of observations in the case of event dates (i.e. cross-market holidays, futures expiries, and MSCI rebalances).

For each training window iteration, the models are re-trained based only on the data available in that particular training window. The window sizes have been translated into a certain number of trading days, such that a constant number of observations are used to train the models throughout the different window iterations and stocks. There are 2,937 holidays for 22 countries whose market data is investigated, throughout 15 full years, between 1st January 2000 and 31st December 2014. This period covers exactly 15 years, or 5,479 days including weekends, or 3,913 days excluding weekends. On average, there are 252 trading days per year for each country, which are derived from the difference between the total number of business days and the average number of holidays per country, which is then divided by the number of years: $(3913 - 2937/22)/15 = 251.97$. Therefore, the fixed-length moving windows are defined in trading days as follows: 21 days for 1 month, 63 days for 3 months, 126 days for 6 months, 252 days for 1 year, and 504 days for 2 years. The year 2015 was excluded from this averaging because our data set includes observations until 10th May 2015 and therefore this year has incomplete data.

Out of the 2,353 pan-European stocks, there are 163 stocks (or 6.93%) whose number of observations is less than 504 (corresponding to the 2-year window). As for the remaining 2,190 stocks with available data spreading on over 2 years, there are 26 stocks with less than 100 days outside the 2-year period, 150 stocks with more than 100 days and less than 1,000 days, and 2,014 stocks with over 1,000 days of observations outside the 2-year period.

6.4.1.2 Cluster Job Management

Given the tremendous number of iterations and runtime required by the stock-specific models, we ran these models on two distinct computer clusters for high performance computing, which operate on the Sun Grid Engine grid computing system.

The stock-specific total runtime was 11,878 days (or 33 years), excluding the queuing times associated with each job, which tended to reach even several days during peak times. The stock-specific models have been split into jobs of maximum 1,000 iterations (i.e. 1,000 consecutive target dates for a given stock). For example, a stock with 3,683 observations running a 2-year moving (or growing) window, needs $3683 - 504 = 3179$ iterations to traverse all of the target dates for 1-step ahead volume prediction; therefore, there are 4 jobs for this stock (broken down into 3 jobs of 1,000 iterations and another job of 179 iterations), for a particular learning method. Table 6.4 outlines the total runtime for each method (across all of the stocks and window types) and for each window type (across all of the stocks and learning methods), along with the corresponding number of jobs and target dates.

Table 6.4
The distribution of runtime and number of iterations/target dates by method and window type.

Breakdown Item	Item Name	Runtime (Days)	Jobs	Target Dates
Method	OLS	554.64	45,990	39,604,267
	Stepwise regression	819.76	45,990	39,604,267
	Ridge regression	767.18	45,990	39,604,267
	Lasso regression	6,014.76	45,990	39,604,267
	kNN (arithmetic mean)	1,791.22	45,990	39,604,267
	kNN (inverse distance)	1,361.45	45,990	39,604,267
	SVR	569.41	45,990	39,604,267
Window type	Moving window, 1 month	938.53	55,622	49,802,970
	Moving window, 3 months	993.51	55,391	49,111,188
	Moving window, 6 months	1,142.35	54,978	48,074,677
	Moving window, 1 year	1,196.28	53,571	46,029,606
	Moving window, 2 years	1,494.24	51,184	42,105,714
	Growing window	6,113.52	51,184	42,105,714

Lasso and kNN are computationally expensive, mainly because lasso performs a two-section search (i.e. grid search and bisection method) and deals with a large number of features (up to 40 variables) when regularising their coefficients and sets some to zero, while kNN is memory-based and requires heavy resources when finding the k nearest neighbours for a test point. The runtime for the various window sizes is larger when the window grows in size and is significantly larger for the growing window approach.

6.4.2 Cross-Stock Models

We investigated the effect of event dates on trading volume, focusing on cross-market holidays, stock index futures expiries, and MSCI quarterly reviews. The sparsity of these observations determined the models to be trained on cross-stock data. Since stocks exhibit different volume and price magnitudes, we normalised the past observations of trading volume and aggregated the data for the entire stock universe. Even after aggregating, the number of observations was significantly less than in the case of stock-specific models; there are 2,904 target dates and predicting their volume had a runtime of 15 days. Each model corresponds to only one learning method, whose definition of predictors and methodology was dictated by a previous study. The cross-market holiday model employs ridge regression, whereas the futures expiries and the MSCI rebalances are fit using OLS.

Unlike the stock-specific models where the target variable consists of the logarithmic consolidated volume, the cross-stock models employ the 'relative volume' as the target variable. Equation (6.13) shows the formula for the relative volume, which is determined by the log-ratio between the consolidated volume on the target date (also called 'event date' or 'special date') and the stock-specific benchmark volume. This benchmark is computed as the median of the trading volumes of the 20 trading days prior to the target date (i.e. the futures expiry, MSCI rebalance, or cross-market holiday). The median was selected among other measures of central tendency (e.g. geometric mean or arithmetic mean) because it was the most robust to the outliers in our data set. By dividing a stock's target date volume by the benchmark volume, we get a normalised value for the trading volume, which works well across our stock universe. This normalisation, consisting in the identification of observations from multiple stocks that have a common target date, was necessary as these event dates are periodic, but sparse:

$$V_{\mathrm{rel}} = \hat{y} = \log \frac{V_{t0}}{V_{\mathrm{benchmark}}} = \log \frac{V_{t0}}{\mathrm{median}(V_{t-1}, V_{t-2}, \dots, V_{t-20})}. \tag{6.13}$$

When performing 1-step ahead prediction, these models estimate the relative volume. In order to be able to make stock-specific predictions, the relative volume needs to be converted to a particular stock's logarithmic volume. Essentially, we train the model on the entire stock universe sharing a common event date, but we make stock-specific predictions by transforming the target variable from being stock-agnostic to being stock-specific. Equation (6.14) shows how to calculate the stock-specific volume estimate $\widehat{y'}$ based on the relative volume. We add the benchmark volume to the relative volume, as this is the stock-specific term that customises the volume prediction for a given stock:

$$V_{\mathrm{stock}} = \widehat{y'} = V_{\mathrm{rel}} + \log V_{\mathrm{benchmark}}. \tag{6.14}$$

6.4.2.1 Cross-Market Holidays

The cross-market holiday model implements ridge regression, which performs a two-section search for each iteration. Ridge regression was appropriate for the cross-market holidays as it addresses the problem of multicollinearity and reduces the coefficient variance. Its predictors consist of the constant term, 20 lagged normalised volumes (i.e. a stock's most recent 20 volumes divided by their median), 21 indicator variables for the trading country, and 22 indicator variables for the holiday country, adding the US on top of the 21 trading countries. The regression line is outlined in Equation (6.15), where β_0 is the constant term, T_i is the indicator variable signalling whether the i^{th} country is trading, and H_i indicates whether it is on holiday:

$$\hat{y} = \beta_0 + \sum_{i=1}^{20} \beta_i^{\mathrm{lag}} \frac{V_{t-i}}{V_{benchmark}} + \sum_{i=1}^{21} \beta_i^{\mathrm{country}} T_i + \sum_{i=1}^{22} \beta_i^{\mathrm{country}} H_i. \tag{6.15}$$

6.4.2.2 Stock Index Futures Expiries

The futures expiry model is fit using OLS. Stepwise regression, or more generally feature selection, was not performed based on the previous findings, where the OLS provided a more stable model across the analysis. The feature set consists of the constant term, 20 lagged normalised volumes and 7 indicator variables corresponding to the futures indices included in this pan-European analysis (i.e. Amsterdam Exchange, CAC 40, FTSE MIB, FTSE, Deutsche Boerse DAX, IBEX 35, and OMX Stockholm 30), showing which expiring index a particular observation is a member of. The model is summarised in Equation (6.16), where E_i indicates whether a particular stock is the constituent of the i^{th} index:

$$\hat{y} = \beta_0 + \sum_{i=1}^{20} \beta_i^{\mathrm{lag}} \frac{V_{t-i}}{V_{\mathrm{benchmark}}} + \sum_{i=1}^{7} \beta_i^{\mathrm{index}} E_i. \tag{6.16}$$

6.4.2.3 MSCI Quarterly Reviews

The MSCI rebalance model is similar to the futures expiry model and is fit using OLS due to the same considerations. It is modelled for the MSCI International Pan Euro Price Index, which covers 204 stocks from 16 countries: Austria, Belgium, Switzerland, Germany, Denmark, Spain, Finland, France, United Kingdom, Greece, (Republic of) Ireland, Italy, Netherlands, Norway, Portugal, and Sweden. The model terms include the intercept, 20 lagged normalised volumes and 16 indicator variables for the trading country of each stock, i.e. the exchange country where the stock is trading; these are outlined in (6.17), where C_i represents the indicator variable for the i^{th} country included on the MSCI pan-European index:

$$\hat{y} = \beta_0 + \sum_{i=1}^{20} \beta_i^{\text{lag}} \frac{V_{t-i}}{V_{\text{benchmark}}} + \sum_{i=1}^{16} \beta_i^{\text{country}} C_i. \tag{6.17}$$

6.4.3 Stock-Specific Models

There are seven stock-specific models employing different supervised learning techniques and they all begin from the model function in Equation (6.18). The initial feature set includes the constant term, 15 volume (autoregressive) lags, 14 volume (moving average) windows, 5 price metrics P_i that are trained using the opening, closing, low and high prices of the previous trading day (i.e. $t-1$) and the opening price of the target day (i.e. t_0) in the case of the overnight return, and finally five indicator variables corresponding to each business day, denoted by D_i, where i ranges from 1 to 5 (i.e. Monday to Friday). The target variable of these models is a particular stock's estimated logarithmic volume for the next trading day (i.e. t_0):

$$\hat{y} = \beta_0 + \sum_{i=1}^{15} \beta_i^{\text{lag}} V_{t-i}^{\text{lag}} + \sum_{i=2}^{15} \beta_i^{\text{lag}} V_{t-i}^{\text{win}} + \sum_{i=1}^{5} \beta_i^{\text{p}} P_{i,t-1} + \sum_{i=1}^{5} \beta_i^{\text{d}} D_i. \tag{6.18}$$

We use three main price log-ratios: intraday range, asymmetric intraday return, and asymmetric overnight return; their formulae are shown in Equation (3.8), Equation (3.9), and Equation (3.10), respectively. The overnight return is divided by the number of intervening nights in order to correct for the additional non-trading day observed throughout weekends and bank holidays. Our previous empirical evidence found that better performance is achieved when splitting the intraday return and overnight return log-ratios at zero, into positive absolute values (denoted by 'absPos', representing the absolute value of positive returns only), and negative absolute values (denoted by 'absNeg', corresponding to the absolute value of negative returns). Consequently, we include the following 5 price metrics in the initial model: intraday range, 'absPos' intraday return, 'absNeg' intraday return, 'absPos' overnight return, and 'absNeg' overnight return.

Given this initial model, the analysis follows the framework described in the Analysis Approach section for each individual iteration (i.e. for each target date, given a particular learning method, a particular stock and a particular window type): partitioning the data for the subsequent stratified 10-fold cross-validation applications and determining the optimal orders for the volume lags and the volume windows, producing a model with potentially less features than the initial model, where 15 volume lags and 14 volume windows were included. Then, if the method is a shrinkage method (i.e. ridge regression or lasso regression) or kNN, the method optimal parameter is identified using cross-validation on that iteration's training set. Eventually, each of the following methods is applied to this model definition, using the methodology described in the Methodology Review section: OLS, stepwise regression, ridge regression, lasso regression, kNN with arithmetic mean, kNN with inverse distance weighting, and SVR. A particular constraint is applied to stepwise regression, where we force the constant term, the volume lags and volume windows to be kept into the reduced model when performing sequential feature selection.

6.5 Results

The results of this study are outlined in this section, along with an interpretation of their meaning and implication. We start by investigating the distribution of the volume lags and windows, and then explore the method-specific results, such as the model parameter distribution and feature selection. Next, we provide a performance benchmark for the various models employed in this model, leading to an interpretation of the optimal learning method and training window, and to a breakdown of the model performance by event dates. Based on this performance breakdown by special events, we propose a switching model that virtually adapts from one underlying model to another, based on the current state of the market, which is driven by event dates (e.g. futures expiries, MSCI rebalances, and cross-market holidays) or the current day-of-the-week.

6.5.1 Contribution of Recent Data: Volume Lags and Windows

For each unique combination of stock, learning method and window type, a certain optimal order for the volume lag and volume window is determined for every target date. Since a stock has different cross-validation partitions across its various models consisting of different learning methods and window types, we report very minor fluctuations in the order distribution of volume lags and volume windows for the same window type across the seven training methods. Therefore, we aggregated the order values across the seven models, grouped by window type.

Table 6.5 outlines the descriptive statistics for each window for volume lags and volume windows. We observe a correlation between the size of the training window and the mean and median of the volume lag/window orders. This suggests that the larger the training set, the more relevant past volumes tend to become in fitting an accurate prediction model. This confirms that trading volume is autoregressive and that past observations are meaningful if a substantial training set is available to learn the model.

Table 6.5
Descriptive statistics for the orders of the volume lag and the volume window, grouped by window type.

Past Volume Type	Window Type	Min	Max	Mean	Median	Standard Deviation
Volume lag	MW_1M	1	15	1.22	1	0.51
	MW_3M	1	13	1.33	1	0.63
	MW_6M	1	11	1.65	1	0.89
	MW_1Y	1	13	2.35	2	1.23
	MW_2Y	1	15	3.56	3	1.57
	GW	1	15	7.67	7	3.22
Volume window	MW_1M	2	15	2.24	2	0.52
	MW_3M	2	10	2.24	2	0.54
	MW_6M	2	12	2.34	2	0.67
	MW_1Y	2	15	2.65	2	0.98
	MW_2Y	2	15	3.50	3	1.50
	GW	2	15	7.70	7	3.28

In Figure 6.1, we can visualise how the highly positive skewness from Panel A, which corresponds to the smallest training window (i.e. 1-month moving window) gradually transforms into a relatively symmetrical distribution in Panel F, where the growing window approach trains the model using a variety of lag orders, including the high orders towards 15.

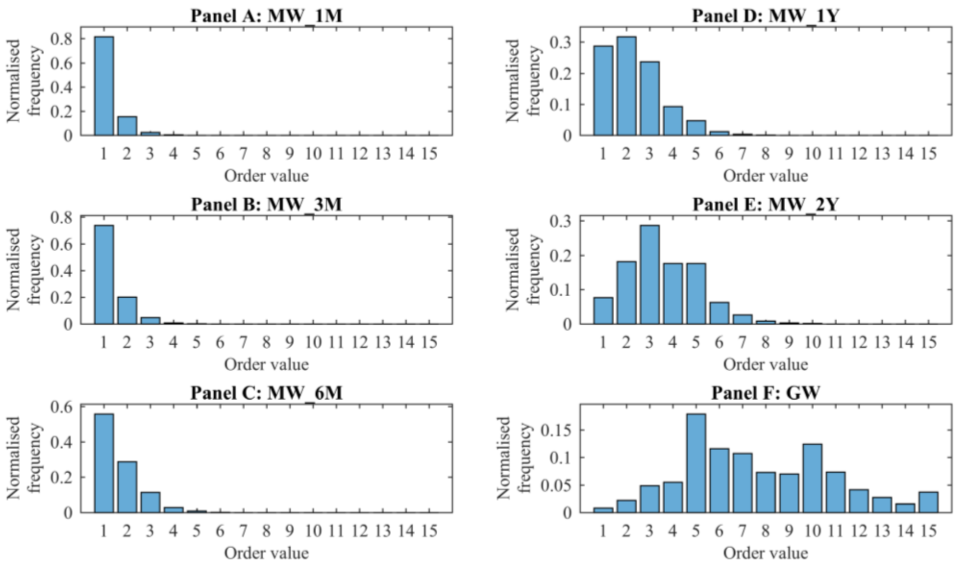

Figure 6.1: Distribution of the volume lag orders across the six different window types.

The moving average-based volume windows in Figure 6.2 exhibit a similar pattern and the volume window orders are positively skewed for the smallest training window (i.e. 1-month moving window in Panel A). The positive skewness decreases once the training window is extended, and becomes rather symmetrical for the largest training window in Panel F (i.e. the growing window). The growing window starts from an initial window size of 2 years, whose distribution is outlined in Panel E. However, the larger the window becomes, the less the order distribution is positively skewed, exhibiting a negatively skewed distribution for the largest window sizes of the growing window, which ultimately yields the relatively symmetrical distribution in Panel F.

The shift in the order distribution from smaller to larger training windows provides evidence that recent volume data contributes to the prediction accuracy and that the amount of meaningful recent data (in the form of lag and window orders) increases with the number of observations in the training window.

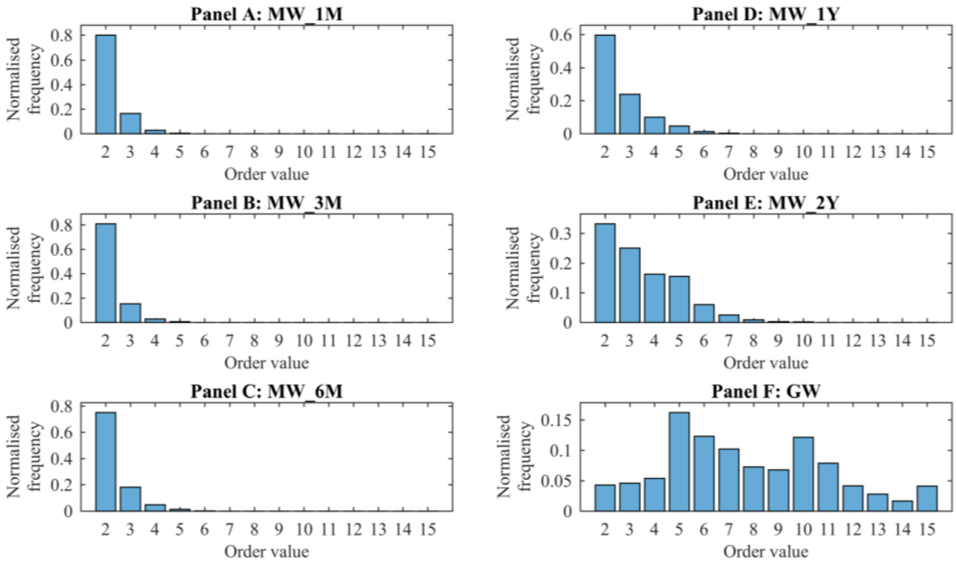

Figure 6.2: Distribution of the volume window orders for different window types.

6.5.2 Method-Specific Parameters

We performed grid search between 1 and 50, with a unit-sized step, and between 50 and 100, with a step size of 5, in order to identify the optimal value of k for the two kNN models, while conducting a two-section search for identifying the optimal value of λ for the two regularisation methods in this analysis. At each step, 10-fold stratified cross-validation was performed to validate the model performance. Below, we outline the distributions and patterns of these two method-specific parameters, i.e. k and λ.

6.5.2.1 *k*-Nearest Neighbours

There are two models implementing kNN, one that treats neighbours equally and uses the arithmetic mean to determine the target variable, and one that penalises the distance between the test point and its neighbours through inverse distance weighting. Table 6.6 includes the descriptive statistics for the values of k for each window type of the two kNN models, i.e. kNN with arithmetic mean and kNN with inverse distance weighting, for the entire stock universe. We observe similar results for the distribution of k across these two models, although the inverse distance weighting approach tends to have slightly higher values of central tendency, having the mean and median with almost 3 neighbours more than the arithmetic mean approach. We report that the mean and median increase with the window size, especially when comparing the 2-year moving window with the growing window, as their initial iteration is identical, confirming that the market data has a highly noisy structure. The value of k for the 1-month moving window reaches is less than 18 as it contains 21 trading days and similarly the 3-month moving window has less than 56 neighbours as it contains 63 trading days in total. However, k reaches the maximum number of 100 neighbours once the training window is at least 6 months long; we did not allow for more than 100 neighbours in order to avoid over-smoothing and eliminating important properties in the data distribution. Although we expected the inverse distance weighting approach to have a significantly higher number of neighbours on average potentially because it could assign very low weights to a high number of neighbours with a possible blurring effect, the difference is not very conspicuous. We conclude that the kNN with arithmetic average approach produces a model that is slightly more parsimonious that the one yielded by the inverse distance weighting, although their overall parameter distribution is rather similar. Their performance is discussed in a subsequent section.

Table 6.6
Descriptive statistics for the values of k for the 6 different window types of the two kNN models.

kNN Approach	Window Type	Observations	Min	Max	Mean	Median	Standard Deviation
Arithmetic Mean	MW_1M	7,111,213	1	18	10.82	11	5.25
	MW_3M	7,012,429	1	55	23.26	20	13.92
	MW_6M	6,864,419	1	100	32.43	29	22.29
	MW_1Y	6,572,392	1	100	38.17	32	27.68
	MW_2Y	6,012,088	2	100	39.42	27	29.51
	GW	6,012,088	3	100	48.23	41	26.47
Inverse Distance Weighting	MW_1M	7,111,213	1	18	11.74	12	5.23
	MW_3M	7,012,429	1	55	26.20	24	14.70
	MW_6M	6,864,419	1	100	36.44	32	24.23
	MW_1Y	6,572,392	1	100	41.72	37	28.99
	MW_2Y	6,012,088	1	100	42.37	30	30.13
	GW	6,012,088	1	100	51.02	44	26.99

The distribution of k for every window type is illustrated in Figure 6.3 for the arithmetic mean approach, and in Figure 6.4 for the inverse distance weighting approach. The corresponding distributions for each window type are very similar for the two approaches and both models exhibit a gradual increase in the value of k once the window length grows.

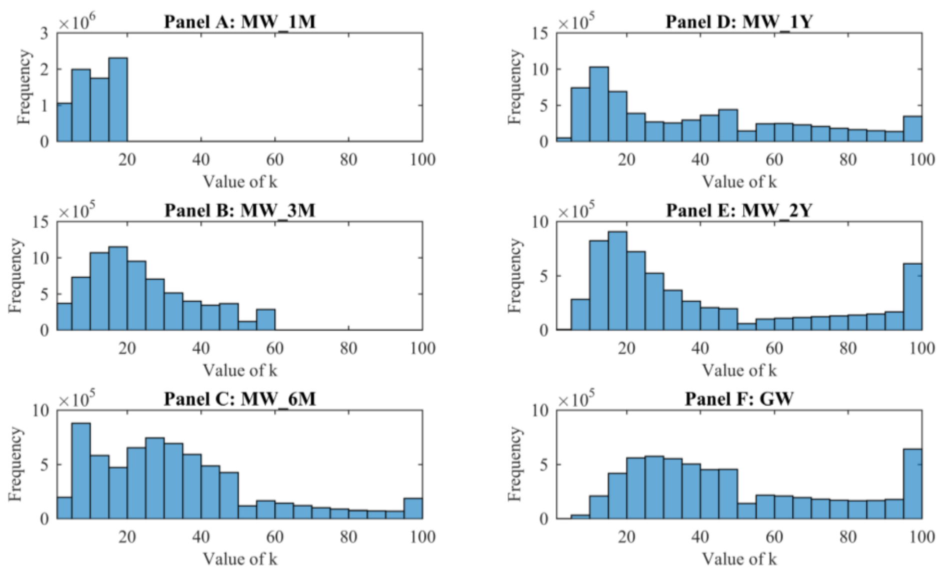

Figure 6.3: Distribution of k in the kNN with arithmetic mean model.

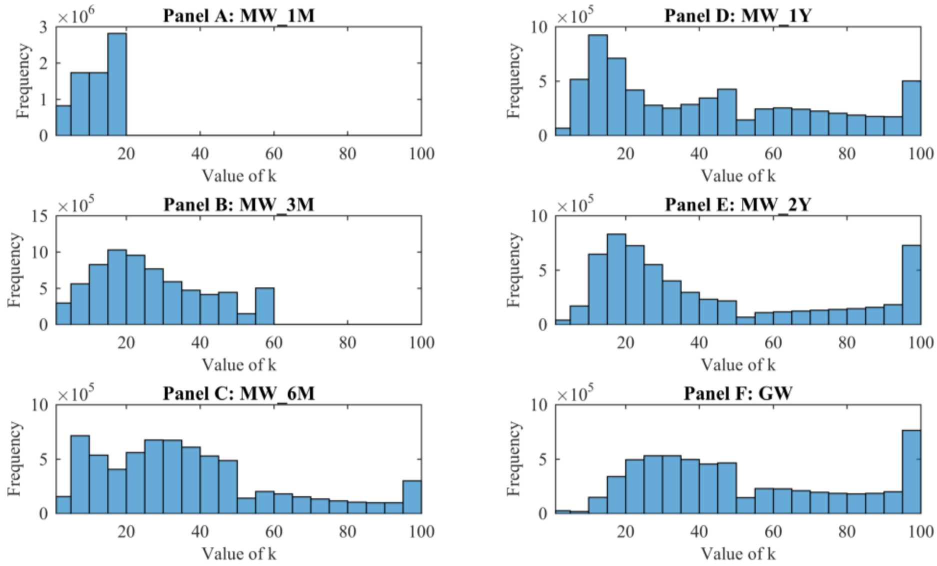

Figure 6.4: Distribution of k in the kNN with inverse distance weighting model.

The empirical cumulative distribution function (CDF) plot in Figure 6.5 is for the growing window models; Panel A represents the kNN with arithmetic mean model, while Panel B represents the kNN with inverse distance weighting. The minor difference in central tendency is noticeable, e.g. in Panel A 65% of the values of k are less than or equal to 50, whereas the proportion in Panel B is 60%.

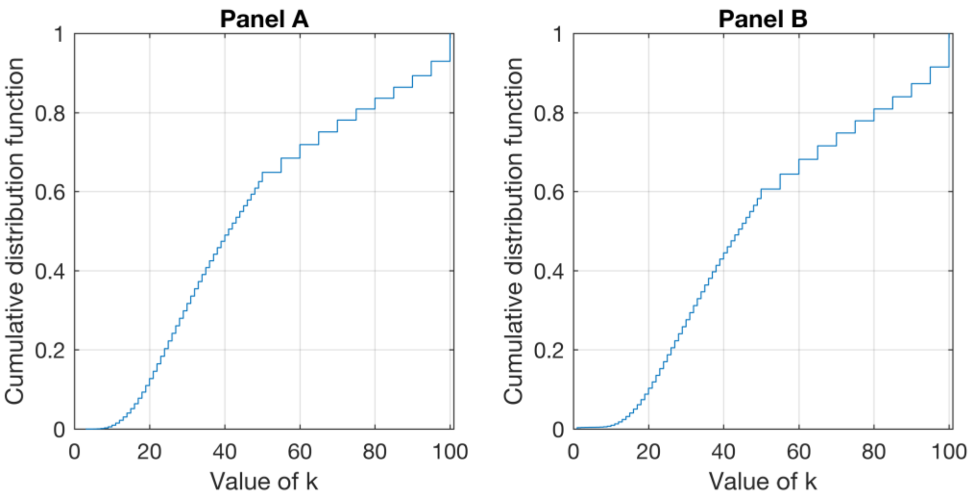

Figure 6.5: Empirical CDF of k for the growing window model for kNN with arithmetic mean in Panel A and kNN with inverse distance weighting in Panel B.

6.5.2.2 Regularisation Methods

The shrinkage methods in this analysis consist of an initial identification step for λ, employing a two-section search, where we first perform grid search to locate the optimal λ in the (base 10) common logarithm interval $[-10,10]$, using a unit-sized step in the $\log_{10}$ space, and then use the spotted value to perform bisection method for determining a more precise value for λ. The most extreme values that λ can take are -11 and 11; this happens when the optimal value for λ in the grid search section is either -10 or 10 and this value is then used as the initial midpoint of the bisection method, with potential extreme values lying one unit away from this midpoint, allowing for values between the interval $[-11,11]$, expressed in base 10 logarithm space. This leads to 24,596 unique values for λ

across the six window types in the ridge regression model, and 18,862 unique values in the lasso regression model. Based on the descriptive statistics for λ in Table 6.7, which are reported for the entire stock universe in the base 10 logarithmic space, we observe significant differences in the distribution of λ between ridge regression and lasso regression. The values of λ are more dispersed throughout the interval $[-11,11]$ in the case of ridge regression (Figure 6.6), whereas lasso regression (Figure 6.7) exhibits a positively skewed distribution, with mostly negative values, where the maximum is either 0 or 1. While the median for λ is around 2 for the ridge regression model, it is -2 for lasso regression.

Table 6.7
Descriptive statistics for the values of λ.

Shrinkage Model	Window Type	Observation	Min	Max	Mean	Median	Standard Deviation
Ridge Regression	MW_1M	7,111,213	-11	11	5.13	2.15	4.59
	MW_3M	7,012,429	-11	11	3.05	2.00	3.15
	MW_6M	6,864,419	-11	11	2.26	1.96	1.98
	MW_1Y	6,572,392	-11	11	1.94	2.00	1.17
	MW_2Y	6,012,088	-11	11	1.86	2.00	0.87
	GW	6,012,088	-11	11	2.04	2.00	0.80
Lasso Regression	MW_1M	7,111,213	-11	1	-0.36	0.00	0.87
	MW_3M	7,012,429	-11	1	-0.93	-1.00	1.08
	MW_6M	6,864,419	-11	1	-1.42	-1.25	1.22
	MW_1Y	6,572,392	-11	1	-1.82	-2.00	1.28
	MW_2Y	6,012,088	-11	0	-2.12	-2.00	1.27
	GW	6,012,088	-11	0	-2.43	-2.01	1.16

The difference in the distribution of λ stems from the different ways in which the two penalties work when dealing with two variables that are highly correlated: the L_1 regulariser (i.e. lasso regression) picks only one of the two correlated predictors, whereas the L_2 regulariser (i.e. ridge regression) keeps both of them in the model and jointly shrinks their coefficients. Therefore, L_2 penalties can minimise the prediction error better than L_1 penalties, although L_1 penalties can reduce overfitting and produce a more parsimonious model. In this section, we only examine the patterns observed in the distribution of these method-specific parameters. The predictive power of these models is compared in a subsequent section of this study.

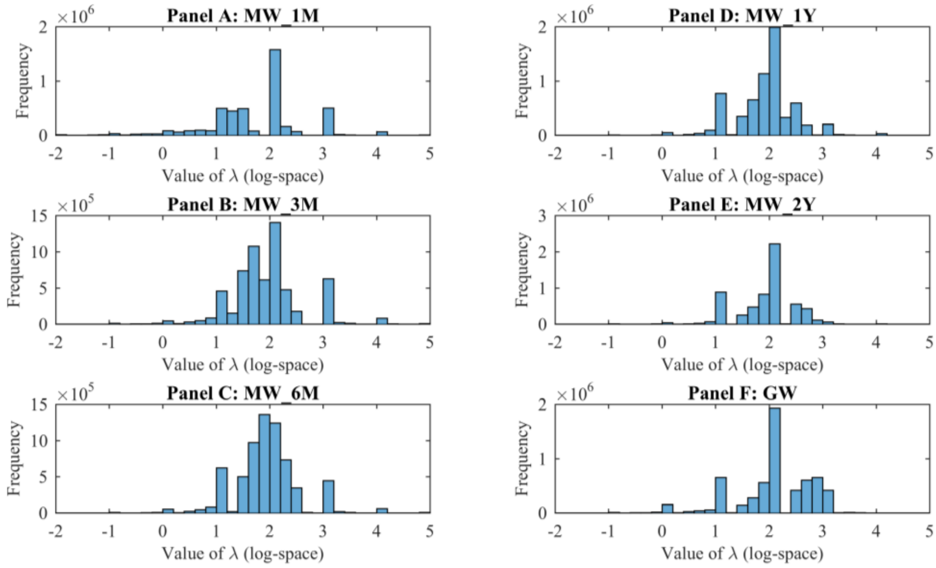

Figure 6.6: Distribution of λ in the ridge regression model.

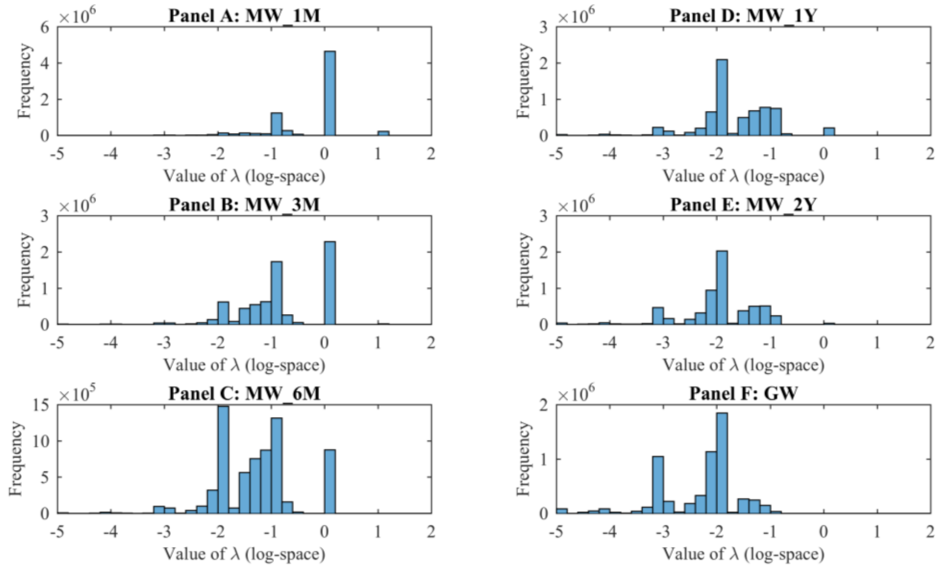

Figure 6.7: Distribution of λ in the lasso regression model.

6.5.3 Feature Selection

We analyse the results of feature selection, which was conducted by two methods: stepwise regression and lasso regression. The stepwise regression models were enforced to keep the intercept and the volume features (i.e. volume lags and volume windows), performing feature selection on the price features (i.e. intraday range, intraday return absPos/absNeg, and overnight return absPos/absNeg) and the five day-of-the-week features, whereas the lasso models could eliminate any feature from the full model. Because of this methodology difference, we start by investigating the selection of the volume features in the reduced model produced by lasso regression. Since every model starts by identifying the optimal order of the volume lags and volume windows, Table 6.8 outlines the proportion of each volume order (ranging from 1 to 15 for the volume lags and from 2 to 15 for the volume windows) in the full models throughout all of the window types of lasso regression. We observe that the volume lag and window orders below 7 are initially included in over

10% of the samples, out of a total of 39,584,629 model iterations. Once these full models are fit with the optimal orders, lasso regression performs variable selection.

Table 6.8
The proportion of volume lag and volume window orders in the full models of lasso regression, averaged over the six window types.

	Order														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Volume lags	100	56.52	36.80	24.50	18.69	12.30	9.34	7.26	6.00	4.89	2.98	1.86	1.23	0.81	0.57
Volume windows	-	100.00	42.65	25.27	18.46	12.42	9.34	7.35	6.09	5.01	3.15	1.94	1.30	0.88	0.63

In order to compute the proportion of each volume feature in the reduced lasso model, we take into account that the volume lag and window orders are mostly less than the maximum value (i.e. 15) and, for each stock, we count the number of occurrences of each feature in the reduced model and also the number of models where a particular volume feature could not be possibly part of the reduced model, because the initially identified optimal order of the full model is lower than this particular volume order. Lasso regression selected the intercept in 100% of the model iterations. Figure 6.8 illustrates the selection proportions of the volume features (i.e. both volume lags and volume windows) in the reduced models across the six window types of lasso regression. The volume lags up to order 11 are selected in more than half of the model iterations, whereas the proportions for the volume window features are significantly lower, ranging from 10% to 35%.

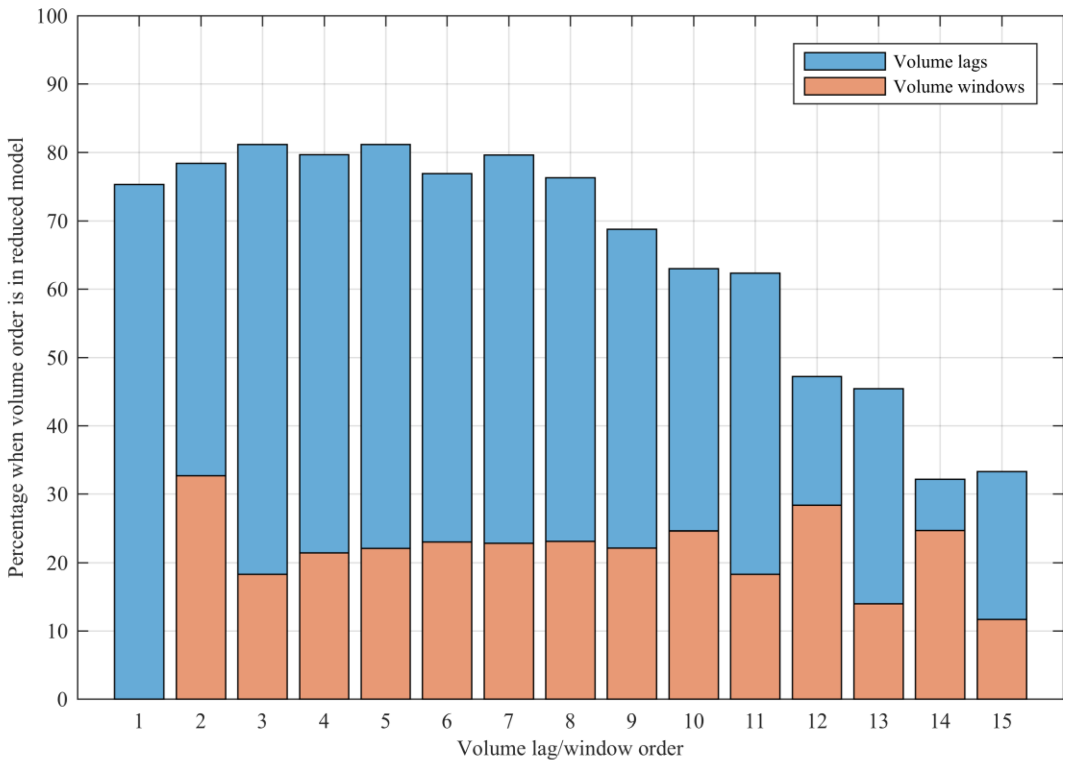

Figure 6.8: The proportion of volume lag and volume window orders in the reduced model produced by lasso regression.

The distribution of the selection proportions of the volume features for each window type is outlined in Table 6.9, for the purpose of spotting potential trends in the volume autoregressive nature. We observe a very high retention of the volume lag features for the models that are trained on at least 6 months of data (i.e. MW_6M, MW_1Y, MW_2Y, and GW). This trend is not followed by the selection of the volume windows, but we can conclude that past observations of the trading volume become more relevant when the learnt model is trained on more than 6 months of observations, and that volume becomes more autoregressive in this context.

Table 6.9
Selection proportion for the volume lag and volume window orders for each window type.

Order	Window type for volume lags						Window type for volume windows					
	MW_1M	MW_3M	MW_6M	MW_1Y	MW_2Y	GW	MW_1M	MW_3M	MW_6M	MW_1Y	MW_2Y	GW
1	17.70	56.56	82.94	95.47	99.28	99.94	0	0	0	0	0	0
2	23.68	65.29	88.02	95.89	98.12	99.52	5.74	18.10	32.64	43.14	48.83	47.80
3	29.75	71.37	91.92	97.07	97.83	99.24	7.60	20.23	23.94	16.51	14.25	27.20
4	30.43	67.71	90.10	96.30	96.47	97.27	11.94	27.23	31.64	19.56	12.71	25.57
5	36.85	71.30	90.46	96.15	96.27	96.05	15.30	28.71	33.85	20.43	12.09	22.30
6	40.33	60.84	79.05	91.04	95.13	95.23	19.04	26.46	33.56	22.16	14.42	22.60
7	41.57	64.21	84.99	94.46	97.57	95.03	19.62	23.95	31.74	24.32	15.86	21.45
8	34.09	60.38	79.70	92.01	97.63	94.07	20.73	24.49	29.78	25.01	17.96	20.65
9	5.56	45.45	77.14	92.23	97.66	94.61	20.45	23.91	26.90	23.27	18.61	19.56
10	0	0	90.32	95.59	97.02	95.08	0	33.33	28.57	39.39	26.99	19.87
11	0	0	90.91	92.65	93.68	96.84	0	0	41.67	21.57	26.95	19.59
12	0	0	0	91.30	95.28	96.77	0	0	100.00	27.27	25.22	17.76
13	0	0	0	75.00	99.29	98.41	0	0	0	33.33	32.28	18.31
14	0	0	0	0	95.45	97.72	0	0	0	100.00	28.95	19.41
15	0	0	0	0	100.00	99.80	0	0	0	0	35.71	34.42

Next, we discuss the feature selection of the price variables and the day-of-the-week indicator variables for both stepwise regression and lasso regression. Table 6.10 shows the selection percentage for each feature after averaging the results across the six window types. There is a notable difference in the selection proportion of the price features with significantly higher values for the stepwise regression implementation, whereas the price features are selected in approximately 3% of the lasso regression models. The day-of-the-week variables are similarly selected in both learning methods. Mondays have the highest percentage (approximately 42%), proving their great significance, either as a weekend effect or as an impact of the Monday bank holidays. Friday is the second most frequent day-of-the-week in the reduced model (being selected in approximately 25% of the models), possibly because of the weekend effect or due to the expiry day effect (e.g. stock index futures expiries or MSCI quarterly reviews).

Table 6.10
The features selected by stepwise regression and lasso regression, averaged across the six window types.

Variable	% in model	
	Stepwise regression	Lasso regression
Intraday range	27.62	3.71
Asymmetric intraday return (absPos)	26.44	3.34
Asymmetric intraday return (absNeg)	26.91	3.03
Asymmetric overnight return (absPos)	42.32	3.31
Asymmetric overnight return (absNeg)	38.60	3.19
Day-of-week: 1	41.81	42.84
Day-of-week: 2	20.60	26.18
Day-of-week: 3	18.88	25.36
Day-of-week: 4	17.63	21.43
Day-of-week: 5	23.52	26.99

For a better visual comparison of the feature selection results among the two learning methods, Figure 6.9 illustrates the selection proportion for stepwise regression (Panel A) and lasso regression

(Panel B), where we observe a significant difference in the frequency of the five price features and rather similar results for the day-of-the-week features. The prediction power of these two techniques is discussed in the next section.

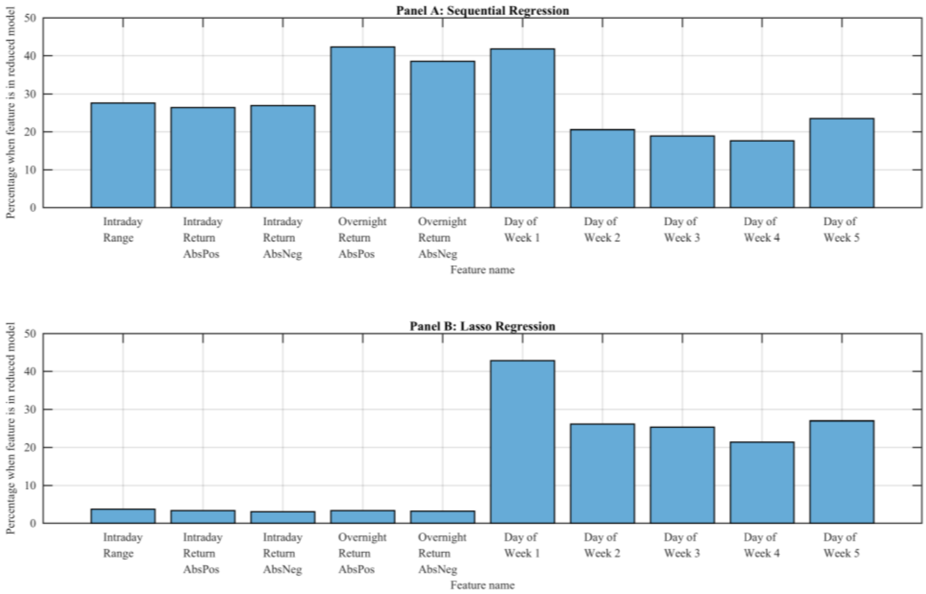

Figure 6.9: The proportion of features selected by stepwise regression (Panel A) and lasso regression (Panel B).

6.5.4 Methodology Performance

For the comparison of the various methods, we need to bear in mind that different stocks have different error magnitudes. Employing the commonly used residual-based evaluation (i.e. including the cross-stock MSE of each method) would not be informative as we are looking to obtain stock-specific model stability. We need to look at some type of error normalisation and a simple way of doing this is to rank the different methods/models for each stock, and then look at the overall (average) ranks when comparing across stocks. Here, we ask the question “What proportion of the time was one method better than the other” and look at the relative performance for each stock. We perform the rankings for each stock and then answer how often each method was the best.

This error-based ranking approach is common in statistics (Rosset, et al., 2007), and can be used overall, as well as on the specific event days. Ranking-based evaluation measures for regression models are interpretable and they are robust against extreme outliers. We used the prediction data, containing the predicted and the observed trading volume for each trading day. Then, we computed the MSE for every stock and then ranked the methods based on the MSE. We used dense ranking (or "1223" ranking), where the models with equal predictions get the same ranking number and the next model receives the following ranking number.

The rank averages for each method and window type in Table 6.11 show that ridge regression trained on a 2-year moving window is the best method for all of the trading dates, including special events. The optimal length of the sliding window approach is 2 years, both for the cross-stock models (i.e. futures expiries, MSCI rebalances, and cross-market holidays) and the stock-specific models (i.e. OLS, stepwise regression, ridge regression, lasso regression, and kNN), with the exception of SVR, whose best error is achieved by the 6-month moving window. The 2-year moving window and the growing window tend to have the best performance across all methods. The 2-year moving window is better in 5 models (i.e. futures expiries, stepwise regression, ridge regression, lasso regression, and SVR), whereas the growing window is better in the other 5 models (i.e. MSCI rebalances, cross-market holidays, OLS, kNN with arithmetic average, and kNN with inverse distance weighting). The average rank for the stock-specific models are: 37.85 for MW_1M, 28.99 for MW_3M, 23.21 for MW_6M, 18.80 (18.61 including the special event models) for MW_1Y, 17.30 (16.55 including the special event models) for MW_2Y, and 17.52 (16.61 including the special event models) for GW. For the moving windows, the rank improves once the window length increases. However, the growing window has a slightly worse rank than MW_2Y, suggesting that recent data might be more relevant as there could be structural breaks across the years. This pattern is not applicable to the OLS method, where the lowest rank across all the models is achieved by OLS GW.

Table 6.11
The mean of the rank of each method and window type for all of the target dates.

Window size	Futures expiries	MSCI rebalances	Cross-market holidays	OLS	Stepwise regression	Ridge regression	Lasso regression	kNN (Arithmetic mean)	kNN (Inverse distance)	SVR
MW_1M	-	-	-	45.82	44.32	38.93	29.93	34.79	33.50	37.65
MW_3M	-	-	-	36.33	29.81	19.56	21.01	30.07	28.37	37.74
MW_6M	-	-	-	22.54	18.73	13.71	15.63	28.43	26.57	36.88
MW_1Y	13.93	16.85	23.74	13.01	11.12	8.78	11.16	26.10	24.52	36.92
MW_2Y	12.11	11.76	20.55	8.27	7.78	6.96	9.85	25.22	23.87	39.14
GW	13.93	10.57	19.03	7.94	8.20	8.16	10.09	23.62	22.27	42.33

The performance of the two kNN methods improves when the window size is larger, as more similar data points are found among the past observations. Throughout the stock-specific learning methods, ridge regression is the best one for 4 window types (MW_3M, MW_6M, MW_1Y, and MW_2Y). OLS has the best average rank for the growing window approach, although the rank of ridge regression growing window is the second best. When using fewer points to train the model, lasso regression achieves the best error. SVR with Gaussian kernel has the poorest performance; this could be further improved by implementing SVR with feature selection.

The standard deviations in Table 6.12 show the performance volatility of the three cross-stock models compared to the stock-specific models.

Table 6.12
The standard deviation of each method and window type for all of the target dates.

Window size	Futures expiries	MSCI rebalances	Cross-market holidays	OLS	Stepwise regression	Ridge regression	Lasso regression	kNN (Arithmetic mean)	kNN (Inverse distance)	SVR
MW_1M	-	-	-	2.06	2.55	5.37	8.03	5.68	5.48	5.51
MW_3M	-	-	-	6.93	7.12	5.38	6.58	6.43	6.33	3.79
MW_6M	-	-	-	7.32	5.88	4.84	4.97	6.61	6.59	3.95
MW_1Y	15.39	17.75	14.49	6.34	5.09	4.55	4.67	5.94	6.21	4.77
MW_2Y	14.99	16.13	13.88	5.72	4.66	4.92	4.88	5.62	6.05	4.74
GW	16.91	15.73	13.91	6.39	6.02	6.68	6.27	8.39	8.37	4.45

We further break the performance of these models by day-of-the-week, event dates (e.g. cross-market holidays, futures expiries, MSCI rebalances), futures index, and non-event dates. The best method and window size, along with their rank's average and standard deviation, are outlined in Table 6.13.

Table 6.13
The best models for various temporal circumstances.

Event Type	Samples	Method	Window	Rank Average	Rank Standard Deviation
All dates	6,012,088	Ridge regression	MW_2Y	6.96	4.92
All Mondays	1,172,481	Ridge regression	MW_2Y	7.78	5.84
All Tuesdays	1,218,480	Ridge regression	MW_2Y	8.47	5.89
All Wednesdays	1,218,790	Ridge regression	MW_2Y	8.65	5.56
All Thursdays	1,209,699	Ridge regression	MW_2Y	8.79	5.56
All Fridays	1,192,638	Ridge regression	MW_2Y	8.83	5.34
All event dates	1,142,613	OLS	GW	9.11	6.46
Cross-market holidays	1,104,456	Cross-market holidays	GW	5.92	6.89
Futures expiries	21,309	Futures expiry	GW	5.55	7.84
MSCI rebalances	8,484	MSCI rebalances	GW	10.34	11.65
Cross-market holidays and futures expiries	7,239	Futures expiry	GW	8.75	11.67
Cross-market holidays and MSCI rebalances	1,125	Cross-market holidays	MW_1Y	16.68	12.47
Amsterdam Exchanges futures expiries	3,202	Futures expiry	GW	3.57	3.48
CAC 40 futures expiries	5,830	Futures expiry	GW	4.57	4.75
FTSE MIB futures expiries	4,690	Futures expiry	MW_2Y	5.77	6.09
FTSE 100 futures expiries	4,447	Futures expiry	GW	3.73	4.86
DAX 30 futures expiries	1,523	Futures expiry	MW_1Y	2.51	2.64
IBEX 35 futures expiries	4,567	Futures expiry	GW	6.64	8.63
OMX Stockholm 30 futures expiries	4,289	Ridge regression	MW_2Y	8.97	4.54
All non-event dates	4,869,475	Ridge regression	MW_2Y	5.71	4.64
All non-event Mondays	813,433	Ridge regression	MW_2Y	7.45	5.80
All non-event Tuesdays	1,063,098	Ridge regression	MW_2Y	7.45	5.63
All non-event Wednesdays	1,075,948	Ridge regression	MW_2Y	7.23	5.33
All non-event Thursdays	986,720	Ridge regression	MW_2Y	7.41	5.49
All non-event Fridays	930,276	Ridge regression	MW_2Y	7.25	5.22

The temporal breakdown above considers the special events exclusively, where the event type is 'cross-market holidays', 'futures expiries' or 'MSCI rebalances', meaning that there are no other special events on the same date. Such overlapping events are covered by 'Cross-market holidays and futures expiries' and 'Cross-market holidays and MSCI rebalances'; futures expiries and MSCI rebalances are mutually exclusive.

6.5.5 The Switching Model

From the methodology performance ranks, we infer an adaptive switching model. This is a cross-stock in-sample analysis that aims to better understand the performance of the various models on specific dates of interest. The 6,012,088 samples are drilled down to the lowest possible granularity by various temporal characteristics, such as: non-event dates (i.e. dates without any special event such as cross-market holidays, futures expiries, or MSCI rebalances), futures expiry index, MSCI rebalances, cross-market holidays, and day-of-the-week. This breakdown incorporates all combinations of these temporal aspects in order to find the best local model. Table 6.14 provides a dissection of the switching model for all the temporal combinations (i.e. every combination of event dates) and outlines each of the 42 sub-models, along with their best models, window sizes and average ranks. This is an extension of the less granular results in Table 6.13, where the best models for certain dates are introduced. For a given trading day, the switching model chooses between these sub-models and picks the locally optimal model. Non-event dates and (special) event dates are mutually exclusive. Moreover, futures expiries and MSCI rebalances also have no overlapping days. The switching model is fit based on the 42 time intervals and their associated best models.

We make a specific comparison of errors on the various event days. The performance comparison between two methods for a given temporal circumstance is computed by getting the intersection of the trading dates that match the current temporal circumstance (e.g. non-event date, special event, certain day-of-the-week etc.) and comparing their ranks. In the situation of a clash between two special events, we choose the model preference by investigating the performance of these models using the intersection of the trading days for these special events, and then computing the MSE per stock and ranking each method for this reduced data set.

Table 6.14
Switching model drilldown based on granular temporal circumstances.

Event type	Samples	Method	Window	Rank Average	Rank Standard Deviation
Non-event Mondays	813,433	Ridge regression	MW_2Y	7.45	5.80
Non-event Tuesdays	1,063,098	Ridge regression	MW_2Y	7.45	5.63
Non-event Wednesdays	1,075,948	Ridge regression	MW_2Y	7.23	5.33
Non-event Thursdays	986,720	Ridge regression	MW_2Y	7.41	5.49
Non-event Fridays	930,276	Ridge regression	MW_2Y	7.25	5.22
FTSE MIB Futures expiry (Thursdays)	94	OLS	MW_2Y	16.73	10.92
IBEX 35 Futures expiry (Thursdays)	25	Lasso regression	MW_6M	18.32	12.41
OMX Stockholm 30 Futures expiry (Thursdays)	105	OLS	MW_2Y	13.30	9.90
Amsterdam Exchanges Futures expiry (Fridays)	2,410	Futures expiry	GW	3.06	3.90
CAC 40 Futures expiry (Fridays)	4,380	Futures expiry	GW	4.53	5.56
FTSE MIB Futures expiry (Fridays)	3,571	Futures expiry	GW	5.96	7.19
FTSE 100 Futures expiry (Fridays)	3,035	Futures expiry	GW	3.72	6.90
DAX Futures expiry (Fridays)	1,037	Futures expiry	MW_1Y	3.76	8.13
IBEX 35 Futures expiry (Fridays)	3,432	Futures expiry	GW	5.55	8.58
OMX Stockholm 30 Futures expiry (Fridays)	3,220	Ridge regression	MW_1Y	8.22	6.15
MSCI rebalance Mondays	417	Lasso regression	MW_3M	17.96	11.51
MSCI rebalance Tuesdays	2,091	MSCI rebalances	GW	11.95	12.90
MSCI rebalance Wednesdays	1,556	MSCI rebalances	MW_2Y	9.20	9.40
MSCI rebalance Thursdays	1,152	MSCI rebalances	GW	16.61	13.90
MSCI rebalance Fridays	3,268	MSCI rebalances	GW	15.04	14.57
Cross-market holiday Mondays	358,289	Cross-market holidays	GW	7.96	9.37
Cross-market holiday Tuesdays	153,291	Cross-market holidays	GW	10.54	9.82
Cross-market holiday Wednesdays	141,087	Cross-market holidays	GW	8.91	10.15
Cross-market holiday Thursdays	221,034	Cross-market holidays	GW	8.57	9.58
Cross-market holiday Fridays	230,755	Cross-market holidays	GW	8.31	8.63
Cross-market holiday and Amsterdam Exchanges Futures expiry (Thursdays)	59	OLS	GW	17.31	12.42
Cross-market holiday and Amsterdam Exchanges Futures expiry (Fridays)	733	Futures expiry	GW	8.94	11.58
Cross-market holiday and CAC 40 Futures expiry (Thursdays)	108	OLS	MW_3M	20.23	16.55
Cross-market holiday and CAC 40 Futures expiry (Fridays)	1,342	Futures expiry	MW_2Y	6.47	9.86
Cross-market holiday and FTSE MIB Futures expiry (Thursdays)	70	Stepwise regression	MW_1Y	12.14	9.63
Cross-market holiday and FTSE MIB Futures expiry (Fridays)	955	Futures expiry	MW_2Y	9.69	11.07
Cross-market holiday and FTSE 100 Futures expiry (Thursdays)	86	Futures expiry	GW	13.67	12.41
Cross-market holiday and FTSE 100 Futures expiry (Fridays)	1,326	Futures expiry	GW	7.19	10.91
Cross-market holiday and DAX Futures expiry (Thursdays)	29	Futures expiry	MW_2Y	14.03	17.95
Cross-market holiday and DAX Futures expiry (Fridays)	457	Futures expiry	MW_1Y	2.62	3.18
Cross-market holiday and IBEX 35 Futures expiry (Thursdays)	84	Stepwise regression	MW_2Y	15.12	10.86
Cross-market holiday and IBEX 35 Futures expiry (Fridays)	1,026	Futures expiry	MW_2Y	7.35	10.78
Cross-market holiday and OMX Stockholm 30 Futures expiry (Thursdays)	133	Cross-market holidays	GW	6.80	7.79
Cross-market holiday and OMX Stockholm 30 Futures expiry (Fridays)	831	Futures expiry	GW	11.91	7.10
Cross-market holiday and MSCI rebalance Mondays	342	Cross-market holidays	GW	15.51	12.88
Cross-market holiday and MSCI rebalance Wednesdays	199	Cross-market holidays	GW	14.53	11.80
Cross-market holiday and MSCI rebalance Fridays	584	MSCI rebalances	GW	14.67	14.42

Next, we compare the switching and the non-switching models, using the average ranks of these methods. The switching model does not have a certain window size enforced as it adapts to the right window size depending on the temporal circumstance. The average ranks in Table 6.15, along with the standard deviations in Table 6.16, show the impressive performance of the switching model, which strongly suggests that markets switch to different states on special events. The switching model has the lowest average rank (5.64); the next best rank is achieved by ridge regression MW_2Y (7.73) and the worst by OLS MW_1M (46.82). The switching model was ranked first in 26.32% of the 2,181 stocks, whereas ridge regression MW_2Y is the best in only 1.65% of the cases. Throughout 76.98% of the stocks, the switching model outperforms the second best model, i.e. ridge regression MW_2Y. Moreover, the switching model is better than the least performing model for every stock in our universe.

Table 6.15
The average rank for every method and window type, along with the switching model, for all of the target dates.

Window size	Futures expiries	MSCI rebalances	Cross-market holidays	OLS	Stepwise regression	Ridge regression	Lasso regression	kNN (Arithmetic mean)	kNN (Inverse distance)	SVR	Switching model
MW_1M	-	-	-	46.82	45.32	39.92	30.91	35.78	34.49	38.64	-
MW_3M	-	-	-	37.33	30.81	20.52	21.96	31.05	29.34	38.74	-
MW_6M	-	-	-	23.52	19.70	14.62	16.55	29.40	27.54	37.87	-
MW_1Y	14.41	17.36	24.53	13.90	11.98	9.59	12.01	27.07	25.50	37.92	-
MW_2Y	12.56	12.12	21.31	9.07	8.58	7.73	10.71	26.21	24.86	40.14	-
GW	14.39	10.90	19.77	8.57	8.84	8.79	10.80	24.58	23.23	43.33	-
-	-	-	-	-	-	-	-	-	-	-	5.64

Table 6.16
The standard deviation of the rank of each method and window type, along with the switching model, for all of the target dates.

Window size	Futures expiries	MSCI rebalances	Cross-market holidays	OLS	Stepwise regression	Ridge regression	Lasso regression	kNN (Arithmetic mean)	kNN (Inverse distance)	SVR	Switching model
MW_1M	-	-	-	2.06	2.56	5.40	8.08	5.71	5.50	5.54	-
MW_3M	-	-	-	6.95	7.14	5.48	6.69	6.48	6.38	3.82	-
MW_6M	-	-	-	7.35	5.93	4.97	5.11	6.68	6.66	3.97	-
MW_1Y	15.77	18.14	14.79	6.44	5.21	4.65	4.82	5.99	6.26	4.79	-
MW_2Y	15.34	16.52	14.20	5.81	4.73	4.98	4.96	5.65	6.07	4.75	-
GW	17.29	16.11	14.22	6.62	6.27	6.93	6.50	8.45	8.43	4.46	-
-	-	-	-	-	-	-	-	-	-	-	4.82

Figure 6.10 illustrates one of the best switching models (i.e. the lowest MSE for a particular stock) for Telefonica SA (TEF.MC), whose MSE for the entire period is 0.078. The plot shows the observed volume and the predicted volume of the switching model for a cropped period of 1 year, due to clarity considerations, between 02/01/2009 and 30/12/2009, where the 1-year MSE is 0.063 for 254 observations.

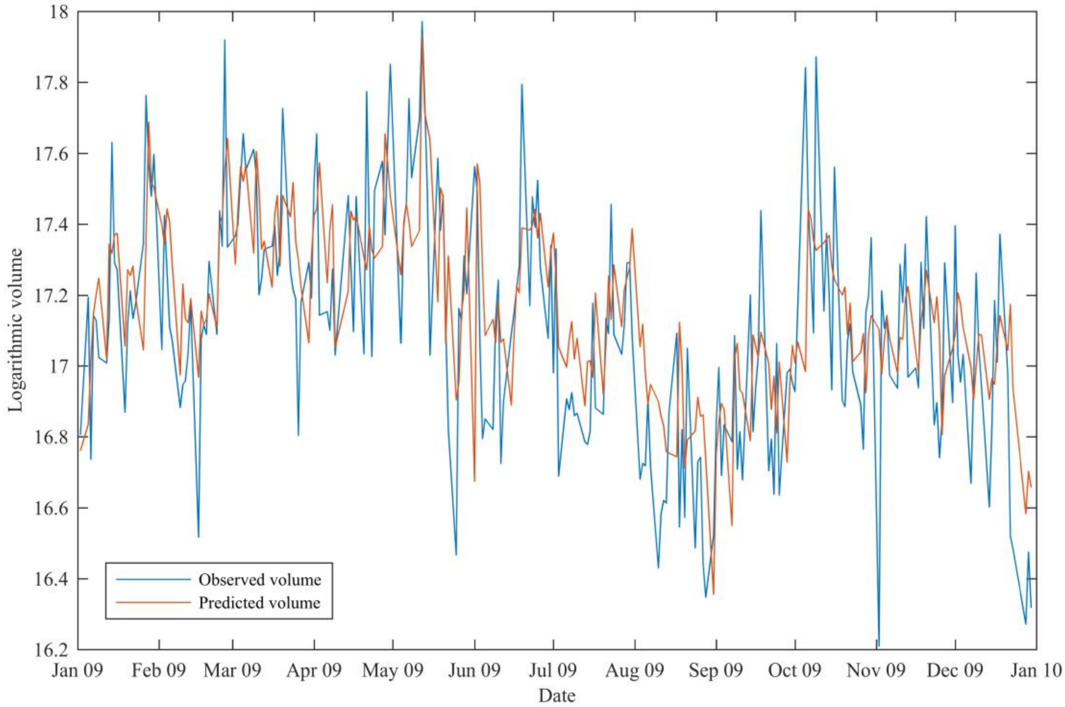

Figure 6.10: Volume prediction using the switching model over one year for Telefonica SA.

We further provide a zoomed version of the time series of Telefonica SA and we plot the volume prediction of the switching model over a period of 6 months, between 02/01/2003 and 30/06/2003, in Figure 6.11. The MSE of the 124 observations is 0.051.

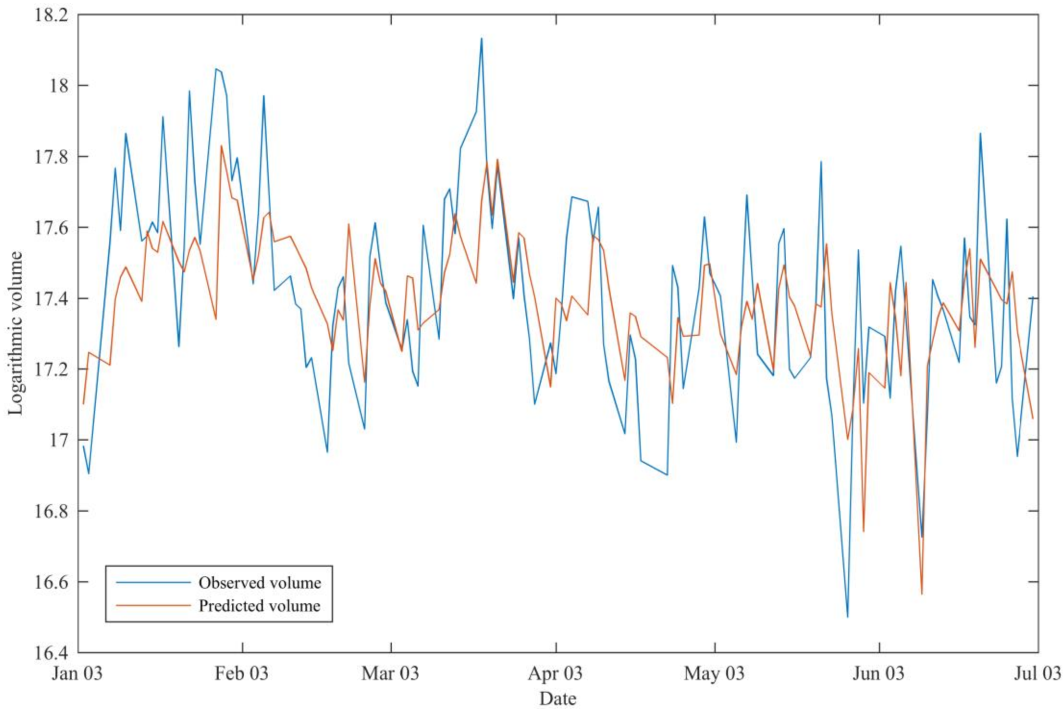

Figure 6.11: The volume prediction of the switching model for Telefonica SA throughout a six-month period.

The best performance improvement achieved by the switching model, compared to the best initial stock-specific models, is 17.41% for Total SA (TOTF.PA). This is computed using the relative change in MSE from the best initial stock-specific model to the switching model. For clarity purposes, we cropped its timeline in Figure 6.12 to 1 year, between 02/01/2013 and 31/12/2013. For these 255 observations, the improvement percentage is 33.40%, the MSE of the best initial model (i.e. ridge regression MW_2Y) is 0.095884, while the MSE of the switching model is 0.063856.

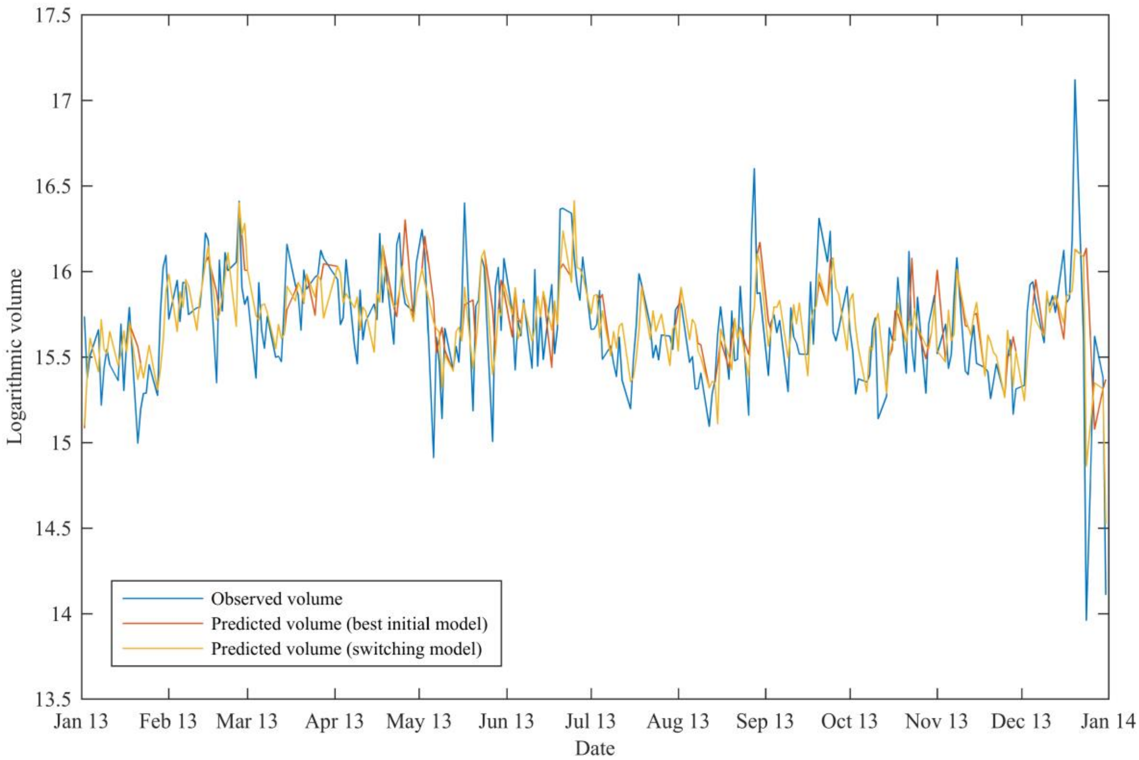

Figure 6.12: The performance improvement of Total SA from the best initial stock-specific model to the switching model.

The largest performance improvement of the switching model when compared to the worst performing initial model is 99.99% improvement and is achieved for 5 stocks. As an illustrative example, Avenir Finance SA (AVEF.PA) has 3,220 observations and the MSE of the worst initial model (i.e. OLS MW_1M) is 4640865.802, whereas the switching model MSE is 4.448. Across all stocks, the performance of the worst initial models is improved by the switching model by 74.595% on average.

6.5.6 Stock-Specific Metamodel

Since the switching model provides an in-sample analysis suggesting the various states markets shift between, we further pose the question whether we can improve the switching model better and provide an out-of-sample model for a given stock. Therefore, we use the ranking-based evaluation measures in order to build an out-of-sample stock-specific metamodel (or surrogate model). For a given stock, we employ a fixed size window of past observations, where the various methods are ranked and the best method is picked to make the next one-step ahead volume forecast. We train two metamodels, using a 1-month and a 3-month moving window. At each step, we evaluate the previous month (corresponding to 21 trading days) or 3 months (corresponding to 63 trading days) and we pick the current best performing method at a given time to make the next day volume prediction. We must note that these metamodel moving windows are different from the concept of moving windows applied to the stock-specific initial models. Here, we still train the initial models using the various training windows (ranging from the one-month moving window to the growing window), and then we investigate the prediction error over the past month or 3 months in order to select the best model throughout the recent time series.

We compute the squared errors for all of the stocks. Then, for each stock, we perform a moving average over one month (21 days) and three months (63 days). We discard 20 stocks having less than 100 test dates as these would not provide enough data for this out-of-sample analysis. This results in 6,011,125 samples, which are further processed by discarding the initial *n* days for each stock, where *n* is the lag number (i.e. 21 trading days or 63 trading days), yielding 5,965,744 samples for the 1-month metamodel and 5,874,982 samples for the 3-month metamodel.

6.5.6.1 One-Month Metamodel

The one-month metamodel is the 27th best model based on the average rank (23.42) in Table 6.17, having a standard deviation of 6.40. Throughout the initial models, there are 8 cross-stock models and 19 stock-specific models that outperform the one-month metamodel. The best rank is achieved by ridge regression MW_2Y and the worst one by OLS MW_1M. The one-month metamodel was the best model for 2 stocks (0.09%), whereas ridge regression MW_2Y was the best in 8.33% of the stocks. The metamodel is better than ridge regression MW_2Y for 43 stocks (1.99%) and it is better than the least performing model, i.e. OLS MW_1M, for all of the 2,161 analysed stocks.

Table 6.17
The mean of each method and window type, along with the one-month metamodel, for all of the target dates.

Window size	Futures expiries	MSCI rebalances	Cross-market holidays	OLS	Stepwise regression	Ridge regression	Lasso regression	kNN (Arithmetic mean)	kNN (Inverse distance)	SVR	1-month meta model
MW_1M	-	-	-	46.84	45.37	40.06	30.79	35.80	34.48	38.68	-
MW_3M	-	-	-	37.34	30.77	19.93	21.55	30.83	29.06	38.78	-
MW_6M	-	-	-	22.94	18.92	13.76	15.79	29.23	27.32	37.85	-
MW_1Y	14.28	17.31	24.40	12.98	11.10	8.73	11.15	26.82	25.19	37.90	-
MW_2Y	12.36	12.05	21.09	8.22	7.71	6.91	9.82	25.85	24.47	40.13	-
GW	14.27	10.86	19.58	7.87	8.18	8.13	10.05	24.11	22.72	43.37	-
-	-	-	-	-	-	-	-	-	-	-	23.42

The largest improvement from the best initial stock-specific model to the one-month metamodel is 3.88% and it is achieved for DBV Technologies SA (DBV.PA). Figure 6.13 illustrates the predictions of the one-month metamodel compared to the best initial model (i.e. stepwise regression) for 247 observations of DBV Technologies SA, between 14/05/2014 and 30/04/2015, where the metamodel performance improvement is 4.5458%. The best initial model MSE is 0.51099, whereas the metamodel MSE is 0.48776.

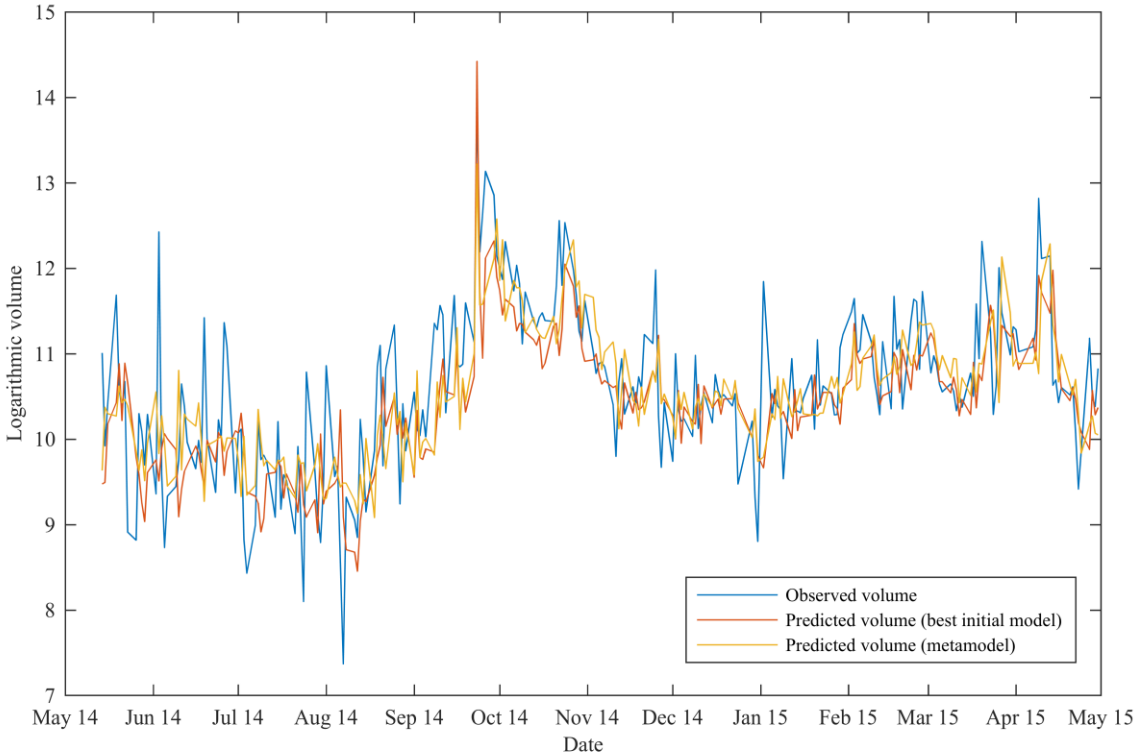

Figure 6.13: The volume prediction of the best initial model and the one-month metamodel for DBV Technologies SA.

6.5.6.2 Three-Month Metamodel

The three-month metamodel model is the 19th best model based on the average rank (14.93) outlined in Table 6.18, having a standard deviation of 5.31. There are 13 cross-stock models and 5 initial stock-specific models that are better than the three-month metamodel. Again, the best rank is achieved by ridge regression MW_2Y and the worst by OLS MW_1M. The three-month metamodel was the best model in only 0.42% of the stocks, i.e. 9 out of 2,161 stocks, whereas ridge regression MW_2Y is the top ranked model in 8.28% of the stocks. The metamodel is better than the ridge regression MW 2Y model in 9.58% of the stock universe (i.e. 207 stocks) and it outperforms the least performing initial model across all of the stocks.

Table 6.18
The mean of each method and window type, including the three-month metamodel, for all of the target dates.

Window size	Futures expiries	MSCI rebalances	Cross-market holidays	OLS	Stepwise regression	Ridge regression	Lasso regression	kNN (Arithmetic mean)	kNN (Inverse distance)	SVR	3-month meta model
MW_1M	-	-	-	46.83	45.35	40.06	30.97	35.83	34.52	38.75	-
MW_3M	-	-	-	37.25	30.77	20.39	21.89	31.07	29.27	38.79	-
MW_6M	-	-	-	23.41	19.55	14.18	16.39	29.42	27.53	37.83	-
MW_1Y	14.44	17.57	24.61	13.28	11.31	8.87	11.44	27.09	25.40	37.82	-
MW_2Y	12.53	12.18	21.42	8.32	7.78	6.96	10.01	26.11	24.75	40.03	-
GW	14.36	11.11	19.87	7.99	8.34	8.28	10.31	24.43	23.04	43.36	-
-	-	-	-	-	-	-	-	-	-	-	14.93

The largest improvement from the best initial model is achieved by the three-month metamodel in the case of Sponda Oyj (SDA1V.HE), with a performance improvement of 3.24%. Figure 6.14 illustrates the volume predictions made by the best initial stock-specific model and the three-month metamodel for Sponda Oyj. There are 253 observations in the one-year cropped timeline, between

03/01/2005 and 30/12/2005. The metamodel, whose MSE is 0.8728, improves the performance of the best initial model (i.e. lasso MW_1Y), whose MSE is 1.0219, by 14.5936%.

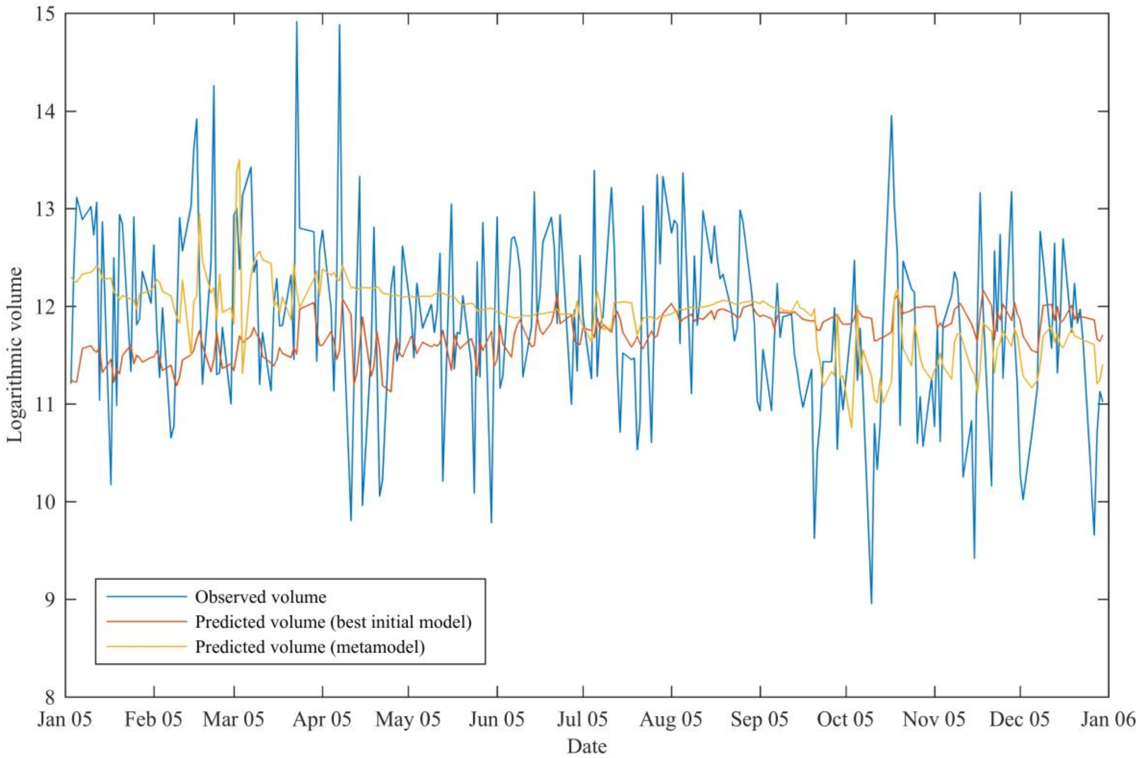

Figure 6.14: The best initial model vs. the 3-month metamodel volume prediction for Sponda Oyj.

The three-month metamodel has a significantly better performance than the one-month metamodel and provides improved model stability, by exhibiting a lower standard deviation.

6.6 Discussion

Volume prediction is critically important for optimal order allocation in order to minimise the market impact. Traders and portfolio managers aim to model the market liquidity by predicting the trading volume such that they do not over-participate, by incurring excessive market impact, or under-participate, by incurring opportunity cost. The study employs an enormous data set, comprising the daily market data for 2,353 European stocks from 21 countries, along with a precisely constructed trading calendar covering more than 15 years for these 21 European countries and the United States.

The aim of this study is to train a variety of learning methods and window types in order to better understand how they perform in certain circumstances, by specifically investigating event dates, such as cross-market holidays, futures expiries, or MSCI quarterly reviews, along with other aspects, e.g. day-of-the-week effect, price-volume relation asymmetry etc. Considering the difference in the volume and price magnitudes among our European stock universe, we independently train 42 stock-specific models, by fitting seven learning methods (i.e. OLS, stepwise regression, ridge regression, lasso regression, kNN with arithmetic mean, kNN with inverse distance weighting, and SVR) for each window type (i.e. the 1-month, 3-month, 6-month, 1-year, and 2-year moving windows, and the growing window). These independently fit stock-specific models had a remarkable runtime of 33 years on the high performance computing clusters. Three additional models are trained using cross-stock normalised observations for the special events (i.e. cross-market holidays, futures expiries, and MSCI rebalances), which are eventually used to make stock-specific predictions. These cross-stock models are learnt using the 1-year and 2-year moving windows and the growing window, producing 9 cross-stock models in total.

Our results corroborate previous findings and provide empirical evidence that the trading volume is autoregressive and this property becomes stronger (i.e. the autoregressive order increases) once the size of the training set is large enough (i.e. in excess of 6 months of training data). For example, the median order of the volume lags and volume windows for the growing window approach is 7. The volume observations from the previous one and a half weeks provide relevant trends, given that the model is trained on a substantial number of data points. The number of neighbours selected by kNN increases gradually once the window length becomes larger. Both kNN with arithmetic mean and kNN with inverse distance weighting reach the maximum number of 100 neighbours that we imposed in our analysis only when the size of the training window is at least 6 months long.

While investigating the effects on volume of the days of the week, we provide consistent results with our previous findings. Mondays are retained by the feature selection methods in 42% of the models, followed by Fridays, whose indicator variable is kept in almost 25% of the models.

Using a ranking-based evaluation, we report that the best model is trained using ridge regression on a two-year moving window. The results indicate that OLS, i.e. the study's most rudimentary method, trained on a growing window has a marginally worse performance than ridge regression, which deals with the multicollinearity problems. The rank of the moving windows improves once the window length increases and the optimal size of the moving window approach is 2 years, whose performance is similar to that of the growing window, although the 2-year moving window has a better rank average across all of the seven learning methods. This could be explained by possible structural breaks across the 15 years analysed by this study, potentially worsening the performance of the growing window when the window reaches a very large size.

Based on a thorough dissection of the temporal circumstances for all of the stocks, we infer a cross-stock switching model that employs the best initial stock-specific model for a given date characteristic. There are 42 disjoint temporal circumstances that are described by different models, which best apply to a particular state of the financial markets. This cross-stock in-sample analysis drills down the 6 million samples into high granularity circumstances identified based on a variety of temporal factors, such as non-event dates, futures expiries, MSCI rebalances, cross-market holidays, day-of-the-week etc. The excellent performance achieved by the switching model confirms our hypothesis that markets are event-driven and shift to different states based on special events.

Ultimately, the goal of this research is to improve model stability and we propose an out-of-sample stock-specific metamodel that evaluates the initial independent stock-specific models on a time window of one month or three months, and picks the model whose performance rank is the best throughout the chosen time window, in order to predict the following day's trading volume. The average performance rank of the one-month metamodel is 23rd, whereas the three-month metamodel performs significantly better and its rank decreases to the 15th position. These metamodels provide an out-of-sample dynamic framework, which aims to improve error stability and forecasts the expected volume to mitigate market impact.

This page intentionally left blank.

7. Conclusion and Future Work

This chapter concludes the thesis and revisits the key findings of its studies, highlights the achievements and contributions of this thesis, and provides recommendations for future work to be carried on by other researchers.

7.1 Summary

The importance of this research is encapsulated by the title of this thesis, namely "Analysis of Key Drivers of Trading Performance", which is emblematic of the broad range of topics covered by the studies of this research. This research examines the European equity markets and is concerned with trading performance and liquidity modelling, by investigating the drivers of trading volume in the context of algorithmic order execution. Trading performance can be severely affected by sizing the orders incorrectly. Traders and portfolio managers can either under-participate and pay the opportunity cost, or they might over-participate by incurring market impact. The market impact is the main topic addressed by this thesis and refers to the adverse effect in the price change caused by a trade or order such that it drives the price against the transaction itself. This is the rationale behind this thesis aiming to accurately predict the trading volume, in order to attempt to model the market impact of an order, since the market impact cost is defined as a function of trading volume. This has practical implications in the financial services industry and it can reduce profits and potentially transform a profitable strategy into a losing strategy by adversely moving the prices. The market impact increases with the trading size and, therefore, practitioners need to disguise their participation on the financial markets and determine the optimal order size allocation given the predicted trading volume for the following time window.

The four studies conducted in this thesis sequentially provided empirical results in order to solve the puzzle of identifying the key drivers of trading volume. These studies were critically important in the way they were interrelated for the overall success of this work. The research starts with an understanding of the market dynamics and it further delves into various calendar effects in order to explore the potential drivers of trading volume. In this regard, the present chapter concludes this thesis by summarising the motivation and the achievements of this research and recommending future work in this area.

This research provides one of the largest European data sets in the literature, covering a pan-European stock universe comprising 2,353 stocks from 21 countries. The daily market data from Thomson Reuters spreads over 15 years, from 1st January 2000 to 10th May 2015. This is complemented by a thoroughly constructed calendar data set, including the trading holidays for 21 European countries and for the United States, along with stock index futures expiries and MSCI quarterly reviews. Having an accurate trading calendar, which is often different from the public holidays announced by national governments, facilitated a better understanding of the calendar effects.

The thesis consists of a series of exploratory studies and comes to an end by proposing a liquidity extraction model, which aims to mitigate the market impact and to ultimately improve trading performance.

Examining Drivers of Trading Volume

The thesis starts by exploring the market dynamics, aiming to identify possible drivers of trading volume. Some of the determined aspects include lagged volume time series, price indicators (e.g. intraday range, intraday return, or overnight return) and the day-of-the-week, where notable negative effects are found on Mondays and on Fridays. The price metrics of the previous trading day (i.e. intraday return and intraday range) improve the volume model in 87% of the stocks. The results of this in-sample analysis indicate an overall price asymmetry in more than 70% of the stocks and we conclude that the intraday return and the overnight return need to be represented asymmetrically by splitting them at zero. We also analyse potential structural breaks around the financial crisis of 2007-08 by splitting the data set on 1st January 2008 into two halves. We do not report any significant structural changes, except for the overnight return, which is rather salient in the pre-crisis period, but becomes neutral after the financial crisis. Also, the Friday feature proportion and its coefficient distribution differ among the two data sets, but the Friday effect does not reverse or vanish. The day-of-the-week features and the overnight return, which is used as a proxy indicator for the opening auction volume (i.e. the most recent information that is publicly available once the continuous trading phase starts), constantly improve the volume prediction performance. The day-of-the-week effect is stronger and provides better performance when it is used in conjunction with the endogenous variables we mentioned above, compared to when it is fit in isolation. This suggests that the traditional calendar effect methodology of fitting a dummy variable model that is isolated from other important variables might not be meaningful enough. The day-of-the-week findings provide a discussion point that leads to the next studies, where the statistical significance of the Friday effect and the Monday effect is analysed in conjunction with the effects of futures expiries and Monday bank holidays.

European Trading Volumes on Cross-Market Holidays

Then, we shift our attention to special event dates, such as cross-market holidays, stock index futures expiries and MSCI quarterly reviews. There is anecdotal evidence of reduced trading volumes on cross-market holidays and surges in volume on futures expiries and MSCI rebalances, but these have not been investigated in detail in the literature on European financial markets. We provide empirical evidence for the existence of a so-called 'cross-market holiday effect' among the European markets, where the volume is 8.5% lower when there is at least one bank holiday in another European market or in the US. Salient effects are observed when there is a bank holiday in a dominant market (e.g. the USA, the UK, Germany and Italy) or when many countries have their stock exchanges shut simultaneously. A few countries tend to have a strong susceptibility to cross-market holidays, e.g. Belgium, Spain, France, Hungary, Netherlands, Portugal, and South Africa. We further test whether it is the Monday bank holidays (mainly originating in the UK, where the majority of the bank holidays fall on a Monday) or the weekend effect that accounts for the lower volumes on Mondays. The results indicate a clear effect that is significantly driven by the Monday bank holidays, although we do not generalise this to claim that the Monday effect does not exist in general, especially on Mondays without bank holidays. The models on cross-market holidays address the practical problem of multi-day trade planning and propose a multi-step ahead forecast model, whose performance decreases with the size of the step and provides best accuracy for 1-step ahead predictions. The empirical evidence suggests that the cross-market holiday effect persists across small-, mid-, and large-cap stocks, and that the market capitalisation does not seem to cause differentiated effect magnitudes on the trading volume. However, we report a structural break around the financial crisis, where the market capitalisation-based impact of two indices has reversed.

Expiry Day Effects on European Trading Volumes

We further delve into the effects of special dates on trading volume and we focus on periodic events, namely the stock index futures expiries and the MSCI rebalances. We analyse the most liquid seven European indices and the MSCI International Pan-Euro Price Index. We also perform randomisation tests in order to determine the real phenomenon that drives the larger volumes, and we discriminate between the Friday effect and the futures expiries, and between the end-of-month effect and the MSCI quarterly reviews. The results indicate a strong futures expiry effect, which increases the volume four days before the expiry date (i.e. on the Monday of the expiry week) and lasts two days after the futures expiry, returning to normal levels three days after the futures expiry date. There is no subsequent effect for the MSCI rebalances, which affect the volumes positively on the day before the review date and on the effective rebalance day. The futures expiries are significantly different from the Friday effect and the expiries account for the larger volumes on Fridays. The months with MSCI rebalances are also significantly different from the adjacent months when there are no MSCI rebalances, but they are not strong enough to explain the overall end-of-month effect. We also propose a practical application of the futures expiry model and the MSCI rebalance model, enabling multi-step ahead forecasts. There is a constant trend of having a larger cross-validation MSE as the step size grows. Throughout the cross-market holiday, futures expiry, and MSCI rebalance models, the trading volume exhibits a constant autoregression property.

Developing a Volume Forecasting Model

Considering the findings of these studies, which support the hypothesis that markets are event-driven and that they are transitioning to different states based on certain events (e.g. bank holidays, futures expiries, MSCI rebalances, days of the week etc.), we incorporate the results of the previous in-sample analyses and propose an out-of-sample volume model that exploits this event-driven nature of the markets. 42 stock-specific models are trained for different statistical methods (OLS, stepwise regression, ridge regression, lasso regression, kNN with arithmetic mean, kNN with inverse distance weighting and SVR), along with other 9 cross-stock models for special events (i.e. cross-market holidays, futures expiries, and MSCI rebalances). These are fit using different window types (e.g. the 1-month, 3-month, 6-month, 1-year, and 2-year moving windows and the growing window). Having a better understanding of how markets behave in certain circumstances and specifically investigating the event dates (e.g. cross-market holidays, futures expiries, and MSCI rebalances), the aim of this study is to improve the prediction accuracy and improve the error stability. The out-of-sample results corroborate the previous findings and show that the trading volume is autoregressive and this property intensifies when the training window has at least 6 months of observations. The most selected day-of-the-week features are Monday (42%) and Friday (25%), suggesting that volumes tend to follow a significantly different trend on these two days. With regard to the various statistical methods that are fit, we report that the 2-year moving window ridge regression model has the best performance, mainly because shrinkage methods cope well with multicollinearity problems in the data set. Next, we shift from a static process to an adaptive process and construct a switching model, by providing a high granularity drilldown of the temporal circumstances depending on the day-of-the-week and whether that particular date has a cross-market holiday, an MSCI rebalance or a futures expiry, and whether the stock is the constituent of a certain index. This dissection results in 42 sub-models and each of these is associated with the best performing method for a particular description of the temporal context. The switching model yields an excellent performance and reinforces the theory that markets are event-driven and that they shift to different states based on the calendar circumstance, especially on special events. Finally, the research concludes with a stock-specific out-of-sample metamodel, which employs the stock-specific method that achieved the best performance throughout a recent time window. This model is proposed as an alternative to the initial stock-specific models, which have a fixed underlying methodology, and allows for more flexibility as it chooses a new model (i.e. different statistical method and different window type) for each test instance.

7.2 Contributions

This thesis contributes to the existing literature in a number of ways. First, it provides a detailed exploration of endogenous and exogenous drivers of trading volume, along with an investigation of potential structural breaks around the financial crisis of 2007-08. Second, the research provides an understanding of the impact of calendar anomalies on trading volume, unlike the majority of the literature where the central focus is on price returns. Third, this is probably the first European study employing a stock universe on such a large scale, where there are daily market data observations from 2,353 stocks spreading over 15 years. Fourth, the calendar data set constructed in this work is of high accuracy and covers 21 European countries and the USA, providing the most complex trading calendar used in the behavioural finance literature to the best of our knowledge. Moreover, this thesis contributes through the novelty of the application of statistical methods such as ridge regression, lasso regression, stepwise regression, SVR, and kNN, to the analysis of calendar effects and trading volumes; these are established statistical methods, but they are not traditionally employed in this application domain. We provide further insights into the impact of calendar effects on the European trading volumes, including the day-of-the-week effect, the end-of-month effect, the holiday effect and the expiry day effect. Finally, the thesis presents a variety of out-of-sample volume prediction models and the implementation of an adaptive model selection approach, whose goal is to improve the volume prediction accuracy, leading to minimal market impact and optimal trading performance. The validation of this research is driven by the collaboration with a leading investment bank. Deutsche Bank provided guidance on practical problems having real-world impact, such as the liquidity extraction model, the anecdotal evidence on special events, or the multi-day trade planning. The rigour of the results is emphasised by the use of (pairwise) randomisation tests, where controlled rearrangements are performed throughout 1,000 repetitions, producing an empirical p-value. The research presented in this thesis has come a long way throughout its studies, starting from a series of in-sample explorations and concluding with an out-of-sample volume prediction framework, whose application is of critical importance to the minimisation of market impact.

7.3 Further Work

While investigating the outcome of this research within its predetermined scope, this work pursued its objectives with satisfactory results. There are always various methods to improve the existing research. The studies in this thesis suggest several possible extensions, and future work can be carried out to answer additional open questions. Below, we define some possible improvements such that further development of this research could be pursued.

From an application viewpoint, this study calls for further work to analyse the trading systems in depth and provide an integrated framework for trading performance analysis. Other implicit transaction costs besides market impact could be considered, such as price trending, timing risk, spread cost, or delay cost. The performance of certain algorithms should also be benchmarked against alternative algorithms in order to determine the hypothetical performance. This work could be challenging because transaction cost analysis data is extremely hard to obtain for academic purposes, due to the intellectual property constraints imposed by the financial companies.

Examining Drivers of Trading Volume

Firstly, our licence with Thomson Reuters did not allow for the retrieval of opening auction volume for our stock universe. We used the overnight returns as a proxy for the opening auction volume, but it would be interesting to incorporate the opening auction volume as a direct indication of recent data. Another area of future research involves intraday volume prediction in order to further analyse the price-volume relation based on tick data, while revisiting the asymmetry of the price-volume relation. A replication of this research investigating the intraday volume prediction could be potentially vital for the improvement of intraday strategies, especially in the context of the continuous growth of high frequency trading.

European Trading Volumes on Cross-Market Holidays

The pre-holiday and post-holiday effects could be investigated, although this analysis would be conducted 'within' markets and not 'across' markets, because the effects are likely to be more relevant for the respective market (e.g. studying the pre-UK bank holiday effect on the UK trading volumes). Moreover, despite the fact that this research is conducted on a massive data set, it could be further extended to construct a global model, including American and Asian stocks, besides the European stocks. Certain clusters and regional differences could be identified in this global study.

Expiry Day Effects on European Trading Volumes

The regression models for the futures expiries and MSCI rebalances suggested potential issues with numerical instability, probably caused by the multicollinearity among the predictors. A further analysis of the data structure is required for a better model definition. Our results indicate that the magnitude of the MSCI quarterly reviews is not strong enough to account for a general end-of-month volume surge. We recommend further testing based on a different quantification of the end-of-month volume. Further investigating the weekend effect in the European equity markets could be a promising area of research. Our empirical results indicate that the Monday bank holidays and the futures expiries account for the different trading volumes on Mondays and Fridays, but we did not investigate the so-called weekend effect in general.

Developing a Volume Forecasting Model

Another practical problem could be addressed by implementing multi-step ahead predictions for the out-of-sample analysis. Since the total runtime of the models in this out-of-sample study added up to 33 years, this runtime should be expected for every analysis re-run, for each different step size. Moreover, the analysis could be further enhanced by connecting it to a machine-readable news feed (e.g. Thomson Reuters News Analytics) in order to get company-specific news and their associated metrics (e.g. relevance score, sentiment, topic code etc.). Financial markets are event-driven and the news stories, company announcements, and quarterly earnings reports could improve the performance of a stock-specific volume prediction model. Additional natural language processing can be modelled on the headlines or contents of the news stories to predict the extent of the impact of that particular news event on the associated stock.

Bibliography

Abhyankar, A., Ghosh, D., Levin, E. & Limmack, R. J., 1997. Bid-Ask Spreads, Trading Volume and Volatility: Intra-Day Evidence from the London Stock Exchange. *Journal of Business Finance & Accounting,* 24(3), pp. 343-362.

Abraham, A. & Ikenberry, D. L., 1994. The Individual Investor and the Weekend Effect. *Journal of Financial and Quantitative Analysis,* 29(2), pp. 263-277.

Agrawal, A. & Tandon, K., 1994. Anomalies or Illusions? Evidence from Stock Markets in Eighteen Countries. *Journal of International Money and Finance,* 13(1), pp. 83-106.

Aitken, M. & Comerton-Forde, C., 2003. How Should Liquidity Be Measured?. *Pacific-Basin Finance Journal,* Volume 11, pp. 45-59.

Aitken, M., Comerton-Forde, C. & Frino, A., 2005. *Closing Call Auctions and Liquidity,* s.l.: Blackwell Publishing.

Al-Deehani, T. M., 2007. Modeling Asymmetry in the Price-Volume Relation: Evidence from Nine Stock Markets. *Investment Management and Financial Innovations,* 4(4), pp. 8-15.

Al-Ississ, M., 2010. *The Impact of Religious Experience on Financial Markets,* Cambridge, MA: Harvard University.

Almgren, R., Thum, C., Hauptmann, E. & Li, H., 2005. Direct Estimation of Equity Market Impact. *Risk,* 18(7), pp. 58-62.

Ariel, R. A., 1987. A Monthly Effect in Stock Returns. *Journal of Financial Economics,* 18(1), pp. 161-184.

Ariel, R. A., 1990. High Stock Returns before Holidays: Existence and Evidence on Possible Causes. *Journal of Finance,* 45(5), pp. 1611-1626.

Arsad, Z. & Coutts, A. J., 1997. Security Price Anomalies in the London International Stock Exchange: A 60 Year Perspective. *Applied Financial Economics,* 7(5), pp. 455-464.

Assogbavi, T., Khoury, N. & Yourougou, P., 1995. Short Interest and the Asymmetry of the Price-Volume Relationship in the Canadian Stock Market. *Journal of Banking & Finance,* 19(8), pp. 1341-1358.

Assogbavi, T. & Osagie, J. E., 2006. Equity Valuation Process and Price-Volume Relationship on Emerging Markets. *International Business & Economics Research Journal,* 5(9), pp. 7-18.

Assogbavi, T., Schell, J. & Fagnissè, S., 2007. Equity Price-Volume Relationship on the Russian Stock Exchange. *International Business & Economics Research Journal,* 6(9), pp. 107-116.

Barber, B. M., Odean, T. & Zhu, N., 2009. Systematic Noise. *Journal of Financial Markets,* 12(4), pp. 547-569.

Barberis, N., Shleifer, A. & Wurgler, J., 2005. Comovement. *Journal of Financial Economics,* 75(1), pp. 283-317.

Barone, E., 1990. The Italian Stock Market: Efficiency and Calendar Anomalies. *Journal of Banking and Finance,* Volume 14, pp. 483-510.

Beaver, W. H., 1968. The Information Content of Annual Earnings Announcements. *Journal of Accounting Research,* Volume 6, pp. 67-92.

Becker, K. G., Finnerty, J. E. & Gupta, M., 1990. The Intertemporal Relation Between the U. S. and Japanese Stock Markets. *Journal of Finance,* 45(4), pp. 1297-1306.

Benjamini, Y. & Hochberg, Y., 1995. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B (Methodological),* 57(1), pp. 289-300.

Bertie, A. J. & Cran, G. W., 1985. Estimation of the Constant Term When Using Ridge Regression. *International Journal of Mathematical Education in Science and Technology,* 16(1), pp. 63-65.

Berument, H. & Kiymaz, H., 2001. The Day of the Week Effect on Stock Market Volatility. *Journal of Economics and Finance,* 25(2), pp. 181-193.

Berument, H. & Kiymaz, H., 2003. The Day of the Week Effect on Stock Market Volatility and Volume: International Evidence. *Review of Financial Economics,* 12(4), pp. 363-380.

Bialkowski, J., Etebari, A. & Wisniewski, T. P., 2010. *Piety and Profits: Stock Market Anomaly During the Muslim Holy Month,* Christchurch, New Zealand: University of Canterbury.

Bishop, C. M., 2007. *Pattern Recognition and Machine Learning.* 1st ed. New York: Springer-Verlag New York.

BlackRock, 2014. *US Equity Market Structure: An Investor Perspective,* London: BlackRock ViewPoint.

Bloomberg Finance, 2010. *Costs in Context - How to Analyze Trading Performance with TCA,* s.l.: Bloomberg Finance.

Booth, G. G., Kallunki, J.-P. & Martikainen, T., 2001. Liquidity and the Turn-of-the-Month Effect: Evidence from Finland. *Journal of International Financial Markets, Institutions and Money,* 11(2), pp. 137-146.

Bordino, I. et al., 2012. Web Search Queries Can Predict Stock Market Volumes. *PLoS ONE,* 7(7), pp. 1-17.

Bouman, S. & Jacobsen, B., 2002. The Halloween Indicator, "Sell in May and Go Away": Another Puzzle. *American Economic Review,* 92(5), pp. 1618-1635.

Brockman, P. & Michayluk, D., 1998. The Persistent Holiday Effect: Additional Evidence. *Applied Economic Letters,* 5(4), pp. 205-209.

Brusa, J., Liu, P. & Schulman, C., 2000. The Weekend Effect, 'Reverse' Weekend Effect, and Firm Size. *Journal of Business Finance & Accounting,* 27(1), pp. 555-574.

Cadsby, C. B., 1989. Canadian Calendar Anomalies and the Capital Asset Pricing Model. In: R. M. C. Guimarães, B. G. Kingsman & S. J. Taylor, eds. *A Reappraisal of the Efficiency of Financial Markets.* Berlin: Springer Berlin Heidelberg, pp. 199-226.

Cadsby, C. B. & Ratner, M., 1992. Turn-Of-Month and Pre-Holiday Effects on Stock Returns: Some International Evidence. *Journal of Banking & Finance,* 16(3), pp. 497-509.

Cao, M. & Wei, J., 2005. Stock Market Returns: A Note on Temperature Anomaly. *Journal of Banking & Finance,* 29(6), pp. 1559-1573.

Casado, J., Muga, L. & Santamaria, R., 2013. The Effect of US Holidays on the European Markets: When the Cat's Away.... *Accounting and Finance,* Volume 53, pp. 111-136.

Chakrabarti, R., Huang, W., Jayaraman, N. & Lee, J., 2005. Price and Volume Effects of Changes in MSCI Indices–Nature and Causes. *Journal of Banking & Finance,* 29(5), pp. 1237-1264.

Chan, M., Khanthavit, A. & Hugh, T., 1996. Seasonality and Cultural Influences on Four Asian Stock Markets. *Asia Pacific Journal of Management,* 13(2), pp. 1-24.

Chen, G., Firth, M. & Xin, Y., 2005. The Price-Volume Relationship in China's Commodity Futures Markets. *The Chinese Economy,* 37(3), pp. 87-122.

Che, S. et al., 2008. Accelerating Compute-Intensive Application with GPUs and FPGAs. *Symposium on Application Specific Processors,* pp. 101-107.

Cheung, C. S. & Kwan, C. C. Y., 1992. A Note on the Transmission of Public Information Across International Stock Markets. *Journal of Banking & Finance,* 16(4), pp. 831-837.

Chiang, C.-H., 2009. *Trading Volume, Returns and Option Expiration Date,* New York: Columbia University.

Chong, R., Hudson, R., Keasey, K. & Littler, K., 2005. Pre-Holiday Effects: International Evidence on the Decline and Reversal of a Stock Market Anomaly. *Journal of International Money and Finance,* 24(8), pp. 1226-1236.

Chow, Y., Yung, H. H. & Zhang, H., 2003. Expiration Day Effects: The Case of Hong Kong. *Journal of Futures Markets,* 23(1), pp. 67-86.

Clark, P. K., 1973. A Subordinated Stochastic Process Model with Finite Variance for Speculative Prices. *Econometrica,* 41(1), pp. 135-155.

CME Group, 2013. *Understanding Stock Index Futures,* Chicago: Financial Research & Product Development.

Cohen, L. & Frazzini, A., 2008. Economic Links and Predictable Returns. *Journal of Finance,* 63(4), pp. 1977-2011.

Connolly, R. A. & Wang, F. A., 2003. International Equity Market Comovements: Economic Fundamentals or Contagion?. *Pacific-Basin Finance Journal,* 11(1), pp. 23-43.

Constantinides, G. M., 1984. Optimal Stock Trading with Personal Taxes: Implications for Prices and the Abnormal January Returns. *Journal of Financial Economics,* 13(1), pp. 65-89.

Cortes, C. & Vapnik, V., 1995. Support-Vector Networks. *Machine Learning,* 20(3), pp. 273-297.

Cross, F., 1973. The Behavior of Stock Prices on Fridays and Mondays. *Financial Analysts Journal,* 29(6), pp. 67-69.

Crouch, R. L., 1970. A Nonlinear Test of the Random-Walk Hypothesis. *American Economic Review,* 60(1), pp. 199-202.

Crouch, R. L., 1970. The Volume of Transactions and Price Changes on the New York Stock Exchange. *Financial Analysts Journal,* 26(4), pp. 104-109.

Darbar, S. M. & Deb, P., 2002. Cross-Market Correlations and Transmission of Information. *Journal of Futures Markets,* 22(11), pp. 1059-1082.

De Bondt, W. F. M. & Thaler, R., 2012. Does the Stock Market Overreact?. *Journal of Finance,* 40(3), pp. 793-805.

De Bondt, W. F. M. & Thaler, R. H., 1994. *Financial Decision-Making in Markets and Firms: A Behavioral Perspective,* Cambridge, Massachusetts, USA: National Bureau of Economic Research.

DellaVigna, S. & Pollet, J. M., 2009. Investor Inattention and Friday Earnings Announcements. *Journal of Finance,* 64(2), pp. 709-749.

Deutsche Bank, 2008. *Equity Market Impact Models. Mathematics at the Interface between Business and Research,* Berlin: Deutsche Bank.

Dimson, E. & Marsh, P., 1999. Murphy's Law and Market Anomalies. *Journal of Portfolio Management,* 25(2), pp. 53-69.

D'Mello, R., Ferris, S. P. & Hwang, C. Y., 2003. The Tax-Loss Selling Hypothesis, Market Liquidity, and Price Pressure Around the Turn-of-the-Year. *Journal of Financial Markets,* 6(1), pp. 73-98.

Dodd, O. & Gakhovich, A., 2011. The Holiday Effect in Central and Eastern European Financial Markets. *Investment Management and Financial Innovations,* 8(4), pp. 29-35.

Drucker, H. et al., 1997. Support Vector Regression Machines. *Advances in Neural Information Processing Systems,* Volume 9, pp. 155-161.

Dubois, M. & Louvet, P., 1996. The Day-of-the-Week Effect: The International Evidence. *Journal of Banking & Finance,* 20(9), pp. 1463-1484.

Duda, R. O., Hart, P. E. & Stork, D. G., 2000. *Pattern Classification.* 2nd Edition ed. New York: John Wiley & Sons, Inc..

Dumitriu, R., Stefanescu, R. & Nistor, C., 2011. Holiday Effect on the Romanian Stock Market. *Vanguard Scientific Instruments in Management 2011,* 1(4), pp. 35-40.

Edgington, E. S., 1964. Randomization Tests. *Journal of Psychology,* 57(2), pp. 445-449.

Efron, B., Hastie, T., Johnstone, I. & Tibshirani, R., 2004. Least Angle Regression. *The Annals of Statistics,* 32(2), pp. 407-499.

Epps, T. W., 1975. Security Price Changes and Transaction Volumes: Theory and Evidence. *American Economic Review,* 65(4), pp. 586-597.

Epps, T. W., 1977. Security Price Changes and Transaction Volumes: Some Additional Evidence. *Journal of Financial and Quantitative Analysis,* 12(1), pp. 141-146.

Epps, T. W. & Epps, M. L., 1976. The Stochastic Dependence of Security Price Changes and Transaction Volumes: Implications for the Mixture-of-Distributions Hypothesis. *Econometrica,* 44(2), pp. 305-321.

Eun, C. S. & Shim, S., 1989. International Transmission of Stock Market Movements. *Journal of Financial and Quantitative Analysis,* 24(2), pp. 241-256.

European Parliament, Council of the European Union, 2006. *Directive 2004/39/EC of the European Parliament and of the Council of 21 April 2004 on Markets in Financial Instruments.* [Online]
Available at: http://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX:02004L0039-20060428
[Accessed 12 06 2014].

Fabozzi, F. J., Focardi, S. M. & Kolm, P. N., 2010. *Quantitative Equity Investing: Techniques and Strategies.* 1st ed. Hoboken, New Jersey: John Wiley & Sons, Inc..

Fabozzi, F. J., Ma, C. K. & Briley, J. E., 1994. Holiday Trading in Futures Markets. *Journal of Finance,* 49(1), pp. 307-324.

Fama, E. F., 1965. The Behavior of Stock-Market Prices. *Journal of Business,* 38(1), pp. 34-105.

Fama, E. F., 1969. Efficient Capital Markets: A Review of Theory and Empirical Work. *Journal of Finance,* 25(2), pp. 383-417.

Fisher, R. A., 1935. *The Design of Experiments.* 1st ed. Edinburgh: Oliver and Boyd.

Flannery, M. J. & Protopadakis, A. A., 1988. From T-Bills to Common Stocks: Investigating the Generality of Intra-Week Return Seasonality. *Journal of Finance,* 43(2), pp. 431-450.

Folger, J., 2013. *Live Trading Performance Issues - How To Start Trading/Live Trading Performance.* [Online]
Available at: http://www.investopedia.com/university/how-start-trading/how-start-trading-live-trading-performance.asp
[Accessed 17 03 2014].

French, K. R., 1980. Stock Returns and the Weekend Effect. *Journal of Financial Economics,* 8(1), pp. 55-69.

Gerety, M. S. & Mulherin, J. H., 1994. Price Formation on Stock Exchanges: The Evolution of Trading within the Day. *Review of Financial Studies,* 7(3), pp. 609-629.

Gibbons, M. R. & Hess, P., 1981. Day of the Week Effects and Asset Returns. *Journal of Business,* 54(4), pp. 579-596.

Godfrey, M. D., Granger, C. W. J. & Morgenstern, O., 1964. The Random Walk Hypothesis of Stock Market Behavior. *Kyklos,* Volume 17, pp. 1-30.

Goetzmann, W. N. & Zhu, N., 2005. Rain or Shine: Where Is the Weather Effect?. *European Financial Management,* 11(5), pp. 559-578.

Goldman Sachs Global Investment Research, 2013. The Unforced Errors of Behavioural Biases. *Fortnightly Thoughts*, 14 02, Issue 49, p. 11.

Greenwich Associates, 2014. *2014 Greenwich Leaders: Global Foreign Exchange Services,* Stamford, Connecticut, USA: Greenwich Associates.

Gultekin, M. N. & Gultekin, B. N., 1983. Stock Market Seasonality: International Evidence. *Journal of Financial Economics,* 12(4), pp. 469-481.

Guyon, I. & Elisseeff, A., 2003. An Introduction to Variable and Feature Selection. *Journal of Machine Learning Research,* Volume 3, pp. 1157-1182.

Hall, C., 2013. *A Fresh Approach to Measuring Market Impact.* [Online]
Available at:
http://thetradenews.com/People_in_The_Trade/A_fresh_approach_to_measuring_market_impact.aspx
[Accessed 06 05 2014].

Hamao, Y., Masulis, R. W. & Ng, V., 1990. Correlations in Price Changes and Volatility across International Stock Markets. *Review of Financial Studies,* 3(2), pp. 281-307.

Hanna, M., 1978. Security Price Changes and Transaction Volumes: Additional Evidence. *American Economic Review,* 68(4), pp. 692-695.

Hansen, P. R., Lunde, A. & Nason, J. M., 2005. Testing the Significance of Calendar Effects. *Working Paper (Federal Reserve Bank of Atlanta)*.

Harris, F. H. d. et al., 2011. Evidence-Based Regulatory Policy Making for Financial Markets: A Panel Discussion of a Proposed Framework for Assessing Market Quality. *Journal of Trading,* Volume 1, pp. 69-89.

Harris, L., 1986. A Transaction Data Study of Weekly and Intradaily Patterns in Stock Returns. *Journal of Financial Economics,* 16(1), pp. 99-117.

Harris, L., 1987. Transaction Data Tests of the Mixture of Distributions Hypothesis. *Journal of Financial and Quantitative Analysis,* 22(2), pp. 127-141.

Harris, L., 1989. A Day-End Transaction Price Anomaly. *Journal of Financial and Quantitative Analysis,* 24(1), pp. 29-45.

Harris, M. & Raviv, A., 1993. Differences of Opinion Make a Horse Race. *Review of Financial Studies,* 6(3), pp. 473-506.

Hassanat, A. B., Abbadi, M. A. & Altarawneh, G. A., 2014. Solving the Problem of the K Parameter in the KNN Classifier Using an Ensemble Learning Approach. *International Journal of Computer Science and Information Security,* 12(8), pp. 33-39.

Hastie, T., Tibshirani, R. & Friedman, J., 2011. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction.* 2nd ed. New York: Springer.

Haugen, R. A. & Jorion, P., 1996. The January Effect: Still There after All These Years. *Financial Analysts Journal,* 52(1), pp. 27-31.

Hiemstra, C. & Jones, J. D., 1994. Testing for Linear and Nonlinear Granger Causality in the Stock Price-Volume Relation. *Journal of Finance,* 49(5), pp. 1639-1664.

Hirshleifer, D. & Shumway, T., 2003. Good Day Sunshine: Stock Returns and the Weather. *Journal of Finance,* 58(3), pp. 1009-1032.

Hirshleifer, D. & Teoh, S. H., 2003. Limited Attention, Information Disclosure, and Financial Reporting. *Journal of Accounting and Economics,* 36(1-3), pp. 337-386.

Hochberg, Y. & Tamhane, A. C., 1987. *Multiple Comparison Procedures.* 1st ed. New York: John Wiley & Sons.

Hoerl, A. E. & Kennard, R. W., 1970. Ridge Regression: Biased Estimation for Nonorthogonal Problems. *Technometrics,* 12(1), pp. 55-67.

Hong, H. & Stein, J. C., 2007. Disagreement and the Stock Market. *Journal of Economic Perspectives,* 21(2), pp. 109-128.

Hong, H. & Yu, J., 2009. Gone Fishin': Seasonality in Trading Activity and Asset Prices. *Journal of Financial Markets,* 12(4), pp. 672-702.

Huberman, G. & Regev, T., 2001. Contagious Speculation and a Cure for Cancer: A Nonevent That Made Stock Prices Soar. *Journal of Finance,* 56(1), pp. 387-396.

Jaffe, J. & Westerfield, R., 1985. The Week-End Effect in Common Stock Returns: The International Evidence. *Journal of Finance,* 40(2), pp. 433-454.

Jain, P. C. & Joh, G.-H., 1988. The Dependence between Hourly Prices and Trading Volume. *Journal of Financial and Quantitative Analysis,* 23(3), pp. 269-283.

John, K., Saadi, S. & Zhu, H., 2015. Liquidity Risk. In: H. K. Baker & G. Filbeck, eds. *Investment Risk Management (Financial Markets and Investments).* 1st ed. New York: Oxford University Press, pp. 141-154.

Johnson, B., 2010. *Algorithmic Trading & DMA: An Introduction to Direct Access Trading Strategies.* London: 4Myeloma Press.

Jones, C. P., Pearce, D. K. & Wilson, J. W., 1987. Can Tax-Loss Selling Explain the January Effect? A Note. *Journal of Finance,* 42(2), pp. 453-461.

Kahneman, D., 2012. *Thinking, Fast and Slow.* London: Penguin Group.

Kamstra, M. J., D. Levi, M. D. & Levi, M. D., 2003. Winter Blues: A SAD Stock Market Cycle. *American Economic Review,* 93(1), pp. 324-343.

Kamstra, M. J., Kramer, L. A. & Levi, M. D., 2000. Losing Sleep at the Market: The Daylight Saving Anomaly. *American Economic Review,* 90(4), pp. 1005-1011.

Kandel, E. & Pearson, N. D., 1995. Differential Interpretation of Public Signals and Trade in Speculative Markets. *Journal of Political Economy,* 103(4), pp. 831-872.

Karpoff , J. M., 1986. A Theory of Trading Volume. *Journal of Finance,* 41(5), pp. 1069-1087.

Karpoff, J. M., 1987. The Relation Between Price Changes and Trading Volume: A Survey. *Journal of Financial and Quantitative Analysis,* 22(1), pp. 109-126.

Keerthi, S. S. & Lin, C.-J., 2003. Asymptotic Behaviors of Support Vector Machines with Gaussian Kernel. *Neural Computation,* 15(7), pp. 1667-1689.

Keim, D. B., 1983. Size-Related Anomalies and Stock Return Seasonality: Further Empirical Evidence. *Journal of Financial Economics,* 12(1), pp. 13-32.

Keim, D. B. & Stambaugh, R. F., 1984. A Further Investigation of the Weekend Effect in Stock Returns. *Journal of Finance,* 39(3), pp. 819-835.

Kemal, S. & Starks, L. T., 1998. The Stock Price–Volume Relationship in Emerging Stock Markets: The Case of Latin America. *International Journal of Forecasting,* 14(2), pp. 215-225.

Kim, C.-W. & Park, J., 1994. Holiday Effects and Stock Returns: Further Evidence. *Journal of Financial and Quantitative Analysis,* 29(1), pp. 145-157.

King, M. A. & Wadhwani, S., 1990. Transmission of Volatility between Stock Markets. *Review of Financial Studies,* 3(1), pp. 5-33.

Klibanoff, P., Lamont, O. & Wizman, T. A., 1998. Investor Reaction to Salient News in Closed-End Country Funds. *Journal of Finance,* 53(2), pp. 673-699.

Krämer, W. & Ralf, R., 1997. Stocks and the Weather: An Exercise in Data Mining or yet Another Capital Market Anomaly?. *Empirical Economics,* 22(4), pp. 637-641.

Krämer, W. & Runde, R., 1996. Stochastic Properties of German Stock Returns. *Empirical Economics,* 12(2), pp. 281-306.

Kunkel, R. A., Compton, W. S. & Beyer, S., 2003. The Turn-of-the-Month Effect Still Lives: The International Evidence. *International Review of Financial Analysis,* 12(2), pp. 207-221.

Kyle, A. S., 1985. Continuous Auctions and Insider Trading. *Econometrica,* 53(6), pp. 1315-1335.

Lakonishok, J. & Levi, M., 1982. Weekend Effects on Stock Returns: A Note. *Journal of Finance,* 37(3), pp. 883-889.

Lakonishok, J. & Maberly, E., 1990. The Weekend Effect: Trading Patterns of Individual and Institutional Investors. *Journal of Finance,* 45(1), pp. 231-243.

Lakonishok, J. & Smidt, S., 1984. Volume and Turn-of-the-Year Behavior. *Journal of Financial Economics,* 13(3), pp. 435-455.

Lakonishok, J. & Smidt, S., 1988. Are Seasonal Anomalies Real? A Ninety-Year Perspective. *Review of Financial Studies,* 1(4), pp. 403-425.

Leys, C. et al., 2013. Detecting Outliers: Do Not Use Standard Deviation around the Mean, Use Absolute Deviation around the Median. *Journal of Experimental Social Psychology,* 49(4), pp. 764-766.

Liano, K. & White, L. R., 1994. Business Cycles and the Pre-Holiday Effect in Stock Returns. *Applied Financial Economics,* 4(3), pp. 171-174.

Loughran, T. & Schultz, P., 2004. Weather, Stock Returns, and the Impact of Localized Trading Behavior. *Journal of Financial and Quantitative Analysis,* 39(2), pp. 343-364.

Müller, L., Schiereck, D., Simpson, M. W. & Voigt, C., 2009. Daylight Saving Effect. *Journal of Multinational Financial Management,* 19(2), pp. 127-138.

Maberly, E. D. & Pierce, R. M., 2003. The Halloween Effect and Japanese Equity Prices: Myth or Exploitable Anomaly. *Asia-Pacific Financial Markets,* 10(4), pp. 319-334.

Marquardt, D. W. & Snee, R. D., 1975. Ridge Regression in Practice. *American Statistician,* 29(1), pp. 3-20.

Martikainen, T., Perttunen, J. & Puttonen, V., 1995. Finnish Turn-of-the-Month Effects: Return, Volume, and Implied Volatility. *Journal of Futures Markets,* 15(6), pp. 605-615.

MathWorks, 2016. *Understanding Support Vector Machine Regression.* [Online]
Available at: http://uk.mathworks.com/help/stats/understanding-support-vector-machine-regression.html?refresh=true
[Accessed 10 03 2016].

McInish, T. H. & Wood, R. A., 1992. An Analysis of Intraday Patterns in Bid/Ask Spreads for NYSE Stocks. *Journal of Finance,* 47(2), pp. 753-764.

Mehdian, S. & Perry, M. J., 2002. Anomalies in US Equity Markets: A Re-Examination of the January Effect. *Applied Financial Economics,* 12(2), pp. 141-145.

Meneu, V. & Pardo, A., 2004. Pre-Holiday Effect, Large Trades and Small Investor Behaviour. *Journal of Empirical Finance,* 11(2), pp. 231-246.

Menzly, L. & Ozbas, O., 2010. Market Segmentation and Cross-Predictability of Returns. *Journal of Finance,* 65(4), pp. 1555-1580.

Mills, T. C. & Coutts, A. J., 1995. Calendar Effects in the London Stock Exchange FT–SE Indices. *European Journal of Finance,* 1(1), pp. 79-93.

Moosa, I. A., Silvapulle, P. & Silvapulle, M., 2003. Testing for Temporal Asymmetry in the Price-Volume Relationship. *Bulletin of Economic Research,* 55(4), pp. 373-389.

Morgan, I. G., 1976. Stock Prices and Heteroscedasticity. *Journal of Business,* 49(4), pp. 496-508.

Morgenstern, O. & Granger, C. W. J., 1963. Spectral Analysis of New York Stock Market Prices. *Kyklos,* Volume 16, pp. 1-27.

Moro, E. et al., 2009. Market Impact and Trading Profile of Large Trading Orders in Stock Markets. *Physical Review E,* 80(6), pp. 1-8.

MSCI, 2014. *MSCI Completes February 2014 ASR Agreement.* [Online]
Available at: http://ir.msci.com/releasedetail.cfm?releaseid=847313
[Accessed 27 10 2015].

MSCI, 2015. *MSCI Equity Indexes August 2015 Index Review.* [Online]
Available at: https://www.msci.com/eqb/pressreleases/archive/MSCI_Aug15_QIRPR.pdf
[Accessed 27 10 2015].

Narang, R. K., 2009. *Inside the Black Box: The Simple Truth about Quantitative Trading.* 1st ed. London: Wiley Finance.

NeverLossTrading, 2014. *Algorithmic Trading.* [Online]
Available at: http://neverlosstrading.com/Algorithmic%20Trading.html
[Accessed 09 06 2014].

Ng, L. & Wang, Q., 2004. Institutional Trading and the Turn-of-the-Year Effect. *Journal of Financial Economics,* 74(2), pp. 343-366.

Nikkinen, J., Sahlström, P., Takko, K. & Äijö, J., 2009. Turn-of-the-Month and Intramonth Anomalies and U.S. Macroeconomic News Announcements on the Thinly Traded Finnish Stock Market. *International Journal of Economics and Finance,* 1(2), pp. 3-11.

Nippani, S. & Medlin, B. W., 2002. The 2000 Presidential Election and the Stock Market. *Journal of Economics and Finance,* 26(2), pp. 162-169.

Nutti, G., Mirghaemi, M., Treleaven, P. & Yingsaeree, C., 2011. Algorithmic Trading. *Computer,* 44(11), pp. 61-69.

NYSE Euronext, 2016. *NYSE Group Turnover.* [Online]
Available at:
http://www.nyxdata.com/nysedata/asp/factbook/viewer_edition.asp?mode=table&key=3307&category=3
[Accessed 23 03 2016].

Ogden, J. P., 1990. Turn-of-Month Evaluations of Liquid Profits and Stock Returns: A Common Explanation for the Monthly and January Effects. *Journal of Finance,* 45(4), pp. 1259-1272.

Pardo, R., 2008. *The Evaluation and Optimization of Trading Strategies.* 2nd ed. Hoboken, New Jersey, USA: John Wiley & Sons, Inc..

Pearce, D. K., 1996. The Robustness of Calendar Anomalies in Daily Stock Returns. *Journal of Economics and Finance,* 20(3), pp. 69-80.

Pettengill, G. N., 2003. A Survey of the Monday Effect Literature. *Quarterly Journal of Business and Economics,* 42(3), pp. 3-27.

Pinegar, M. J., 2002. Losing Sleep at the Market: Comment. *American Economic Review,* 92(4), pp. 1251-1256.

Pitman, E. J. G., 1937. Significance Tests Which May Be Applied to Samples from Any Populations. *Supplement to the Journal of the Royal Statistical Society,* 4(1), pp. 119-130.

Platt, J. C., 1998. *Sequential Minimal Optimization: A Fast Algorithm for Training Support Vector Machines,* Redmond: Microsoft Research.

Pope, P. F. & Yadav, P. K., 1992. The Impact of Option Expiration on Underlying Stocks: The UK Evidence. *Journal of Business Finance & Accounting,* 19(3), pp. 329-344.

Poterba, J. M. & Weisbenner, S. J., 2001. Capital Gains Tax Rules, Tax-Loss Trading, and Turn-of-the-Year Returns. *Journal of Finance,* 56(1), pp. 353-368.

Reinganum, M. R., 1983. The Anomalous Stock Market Behavior of Small Firms in January: Empirical Tests for Tax-Loss Selling Effects. *Journal of Financial Economics,* 12(1), pp. 89-104.

Ritter, J. R., 1987. The Buying and Selling Behavior of Individual Investors at the Turn of the Year. *Journal of Finance,* 43(3), pp. 701-717.

Ritter, J. R. & Chopra, N., 1989. Portfolio Rebalancing and the Turn-of-the-Year Effect. *Journal of Finance,* 44(1), pp. 149-166.

Rogalski, R. J., 1978. The Dependence of Prices and Volume. *Review of Economics and Statistics,* 60(2), pp. 268-274.

Rogalski, R. J., 1984. New Findings Regarding Day-of-the-Week Returns over Trading and Non-Trading Periods: A Note. *Journal of Finance,* 39(5), pp. 1603-1614.

Rosenberg, M., 2004. The Monthly Effect in Stock Returns and Conditional Heteroscedasticity. *American Economist,* 48(2), pp. 67-73.

Rosset, S., Perlich, C. & Zadrozny, B., 2007. Ranking-Based Evaluation of Regression Models. *Knowledge and Information Systems,* 12(3), pp. 331-353.

Roweis, S. & Ghahramani, Z., 1999. A Unifying Review of Linear Gaussian Models. *Neural Computation,* 11(2), pp. 305-345.

Rozeff, M. S. & Kinney, W. R. J., 1976. Capital Market Seasonality: The Case of Stock Returns. *Journal of Financial Economics,* 3(4), pp. 379-402.

Rutledge, D. J. S., 1979. Trading Volume and Price Variability: New Evidence on the Price Effects of Speculation. In: B. A. Goss, ed. *Futures Markets: Their Establishment and Performance.* New York: New York University Press, pp. 137-156.

Sadath, A. & Kamaiah, B., 2011. Expiration Effects of Stock Futures on the Price and Volume of Underlying Stocks: Evidence from India. *IUP Journal of Applied Economics,* 10(3), pp. 25-38.

Saunders, E. M. J., 1993. Stock Prices and Wall Street Weather. *American Economic Review,* 83(5), pp. 1337-1345.

Schäfer, D. & Strauss, D., 2014. Forex Trading Speeds up Moves to Computer Platform. *Financial Times*, 03 03.

Schwert, W. G., 2003. Chapter 15: Anomalies and Market Efficiency. In: G. Constantinides, M. Harris & R. M. Stulz, eds. *Handbook of the Economics of Finance.* Amsterdam: Elsevier B.V., pp. 937-972.

Sharpe, W. F., 1966. Mutual Fund Performance. *Journal of Business,* 39(1), pp. 119-138.

Sias, R. W. & Starks, L. T., 1995. The Day-of-the-Week Anomaly: The Role of Institutional Investors. *Financial Analysts Journal,* 51(3), pp. 58-67.

Siegel, J. J., 2008. Chapter 18: Calendar Anomalies. In: J. J. Siegel, ed. *Stocks for the Long Run.* 3rd ed. New York: McGraw-Hill, pp. 299-315.

Smirlock, M. & Starks, L., 1985. A Further Examination of Stock Price Changes and Transactions Volume. *Journal of Financial Research,* 8(3), pp. 217-226.

Smirlock, M. & Starks, L., 1986. Day-of-the-Week and Intraday Effects in Stock Returns. *Journal of Financial Economics,* 17(1), pp. 197-210.

Steeley, J. M., 2001. A Note on Information Seasonality and the Disappearance of the Weekend Effect in the UK Stock Market. *Journal of Banking & Finance,* 25(10), pp. 1941-1956.

Stoll, H. R. & Whaley, R. E., 1997. Expiration-Day Effects of the All Ordinaries Share Price Index Futures: Empirical Evidence and Alternative Settlement Procedures. *Australian Journal of Management,* 22(2), pp. 139-174.

Sukumar, N. & Cimino, A., 2012. *European Trading Volumes Rise before Last Options Expiry.* [Online] Available at: http://www.bloomberg.com/news/2012-12-18/european-trading-volumes-rise-before-last-options-expiry.html
[Accessed 19 05 2014].

Sullivan, R., Timmermann, A. & White, H., 2001. Dangers of Data Mining: The Case of Calendar Effects in Stock Returns. *Journal of Econometrics,* 105(1), pp. 249-286.

Tauchen, G. E. & Pitts, M., 1983. The Price Variability-Volume Relationship on Speculative Markets. *Econometrica,* 51(2), pp. 485-505.

Thaler, R. H., 1987. Amomalies: The January Effect. *Journal of Economic Perspectives,* 1(1), pp. 197-201.

The Trade, 2007. Understanding Index Front Running. *The Trade*, 12.

Tibshirani, R., 1996. Regression Shrinkage and Selection via the Lasso. *Journal of the Royal Statistical Society. Series B (Methodological),* 58(1), pp. 267-288.

Torre, N. G. & Ferrari, M. J., 1998. The Market Impact Model. *Horizons, The Barra Newsletter,* Volume 167.

Treleaven, P., Galas, M. & Lalchand, V., 2013. Algorithmic Trading Review. *Communications of the ACM,* 56(11), pp. 76-85.

U.S. Securities and Exchange Commission, 2013. *Equity Market Structure Literature Review Part I: Market Fragmentation,* Washington: U.S. Securities and Exchange Commission.

Vapnik, V., 1999. *The Nature of Statistical Learning Theory.* 2nd ed. New York: Springer-Verlag New York, Inc..

Vapnik, V. N., 1999. An Overview of Statistical Learning Theory. *IEEE Transactions on Neural Networks,* 10(5), pp. 988-999.

Vergin, R. C. & McGinnis, J., 1999. Revisiting the Holiday Effect: Is It on Holiday?. *Applied Financial Economics,* 9(5), pp. 477-482.

Vipul, 2005. Futures and Options Expiration-Day Effects: The Indian Evidence. *Journal of Futures Markets,* 25(11), pp. 1045-1065.

Wang, K., Li, Y. & Erickson, J., 1997. A New Look at the Monday Effect. *Journal of Finance,* 52(5), pp. 2171-2186.

Westerfield, R., 1977. The Distribution of Common Stock Price Changes: An Application of Transactions Time and Subordinated Stochastic Models. *Journal of Financial and Quantitative Analysis,* 12(5), pp. 743-765.

Woord, R. A., McInish, T. H. & Ord, K. J., 1985. An Investigation of Transactions Data for NYSE Stocks. *Journal of Finance,* 40(3), pp. 723-739.

World Federation of Exchanges, 2012. *2011 WFE Market Highlights.* [Online] Available at: http://www.world-exchanges.org/files/file/stats%20and%20charts/2011%20WFE%20Market%20Highlights.pdf [Accessed 20 05 2014].

World Federation of Exchanges, 2013. *2012 WFE Market Highlights.* [Online] Available at: http://www.world-exchanges.org/files/statistics/2012%20WFE%20Market%20Highlights.pdf [Accessed 20 05 2014].

Ying, C. C., 1966. Stock Market Prices and Volumes of Sales. *Econometrica,* 34(3), pp. 676-685.

Zou, H. & Hastie, T., 2005. Regularization and Variable Selection via the Elastic Net. *Journal of the Royal Statistical Society. Series B (Statistical Methodology),* 67(2), pp. 301-320.