Topics in Market Microstructure

A thesis submitted for the degree of Doctor of Philosophy in Economics

Department of Economics

University of Essex

Date of submission for examination (January 2019)

Topics in Market Microstructure

Abstract

The increase in trading volume raised concerns about the impact of algorithmic trading, which includes high frequency trading (HFT), on price discovery process and volatility. First chapter considers the regulatory debate concerning HFT which led the European Commission to suggest implementing a financial transaction tax (FTT) of 0.1% on all stock transactions. A simulation of pure electronic limit order book (E-LOB) finds support to implement FTT. Particularly, traders are found to trade more aggressively by increasing the volume traded to remain in a profitable position after tax. The market, nevertheless, ended up with higher trading volume, lower bid-ask spread and almost same price volatility. The second and third chapters rebuild the London Stock Exchange Electronic Order Book (SETS) in real-time for 5 different stocks for two consecutive months, July and August 2007 and utilise the Directional Changes (DC) methodology to track price trends. Notional Volume Weighted Average Price (NVWAP) concept is used to analyse the intraday dynamics of liquidity in the London Stock Exchange E-LOB; namely, the slopes of NVWAP curves and the volumes on both sides of the market (bid and ask) are studied. Second chapter observes that the shape of both sides of the order book changes during DC in predictable ways where changes in volumes and NVWAP curves' slopes revealed to be robust proxy to identify the prevailing market trends without prior knowledge of price. Third chapter assesses order book events' (OBE), i.e. submissions and cancellations, influence on the shape of the order book which generates price trends that may cause flash and mini-crashes. It revealed that OBE' effects are highly significant determinants of the change in cumulative return under normal price conditions while only buy side events are significant under extreme price conditions. These findings match the expected direction of change in cumulative return.

Table of Contents

1 Chapter One

Introduction

Trading in the stock markets has undergone a dramatic perfection over the 20th century onward due to the pressure of the rapid growth of financial markets (Foucault, 1999). The promotion of automatic and electronic trading in financial markets has replaced the old trading system which is based on floor trading and dealership markets. Electronic limit order books, indeed, were the successful reliable candidate to replace floor and dealership markets. What is even more, it becomes possible with the wide spread of the internet to access and benefit from different information anytime and anywhere, the thing that motivated a necessary reconsidering of market structures. Changing market structures would raise many questions that should be answered in order to guarantee a smooth trading. The way markets will change, who will benefit from these changes and what and how the relationship between sellers and buyers will be controlled are examples of questions raised under the name of transparency (Grover et al., 1999).

In the recent years, the wide evolution on electronic commerce in both Europe in the U.S. has led to introduce new trading platforms (e.g. Chi-X and EdgeX) in stock exchanges (Colliard and Foucault, 2012). These platforms allow traders to utilize technology to swiftly buy and sell their shares in the market. Briefly, in line with the notion of fast automated execution and high technology, it has become ordinary for market participants to fully electronically complete their transactions with identified counterparties who want to transact/trade in a matter of seconds. This ultimately sets the tone for emerging a new style of trading known as high frequency trading (HFT), where automated systems electronically buy and sell given that they will not hold a particular position for a long instance (a few seconds or less). High frequency traders (HFTs) are becoming key players in equity markets due to large trading volumes. However, the variation of trading volumes, high volatility of intraday prices (possible trigger of Flash Crashes) and the massive growth in limit order submissions and cancellations caused by HFTs have necessitated strict regulations to control the market (Cliff et al., 2011; Hagströmer and Nordén, 2012).

For many years, HFT firms have undertaken their transactions smoothly and gained billions without being investigated (in shadows). They were criticised and accused of damaging markets and hurting regular investors. Interestingly, a massive public attention is recently placed on HFT since the U.S. May 6, 2010 Flash Crash along with the large trading volumes increase of HFT strategies. In other words, the 'Flash Crash' of 2010 pushed HFT into the spotlight (Gomber et al., 2011; Guo, 2012; Easley et al., 2012a). Interestingly, as there is an impact of HFT on market quality, there has been a considerable effort on questioning market quality and HFT. Most of the research which has investigated HFT has concentrated on the market impact of HFT on market quality measures. Besides, another bulk of studies has inspected the suitability of the current market regulations in the light of HFT. Most of the literature argues in favour of HFT but still more to learn from it.

The present thesis consists of three chapters. The first chapter studies the impact of the ongoing debate on re-regulating the market structure and its effect on HFTs. In particular, the European Commission is considering applying a financial transaction tax (FTT) of 0.1% on all stock transactions. The European Commission believes that the financial sector should contribute more fairly to cover the crisis costs. This contribution comes in a form of cancelling the VAT exemption of most financial services and starting participating in these costs. Although HFT is believed to be good for market quality, HFTs would be negatively hit by such a tax as low-margin HFT strategies would remain unprofitable. More specifically, any regulatory policy would mainly affect market makers if it is directed at HFTs. The hypothesis to be tested is HFTs would be negatively hit by such a tax and therefore the volatility will increase. This study is an attempt to simulate a stock market with large number of interacting agents. This simulation seeks to investigate the aforementioned impact on the market in general and on HFTs in particular. This simulation is based on a pure order-driven market where asks and bids are directly matched in the limit order book.

The second and third chapters rebuild the London Stock Exchange Electronic Order Book (SETS) in real-time for 5 different stock holdings for two consecutive months, July and August 2007. The choice of these stocks considers including stocks from different sectors and different market capitalisations. The economic mechanism behind second and third studies is to continuously search for profitable algorithmic trading strategies and exploit trading opportunities. This, of course, has to be in line with the market regulations and fairness of trade. However, some trading strategies could be used to affect the price and set pressure on it to change in a certain direction to increase profitability of the trader (perpetrator). Predicting future price change to exploit profitable trading opportunities is considered to be lawful act. Yet, setting temporary pressure on the price (using order cancellations for example as a tool) to temporarily change in certain direction and then that pressure is deemed to be unlawful. Second and third studies will investigate these trading opportunities.

Therefore, the second chapter aims at modelling the shape of bid and ask side of the order book using the notional volume weighted average price (NVWAP) concept to analyse the intraday dynamics of liquidity in the London Stock Exchange (LSE) limit order book. Specifically, it checks whether the shape of bid and ask side of the order book could be used as a proxy to identify the prevailing market trend without prior knowledge of the price. Within an algorithmic trading system, as market state changes, these four statistics can be used to switch between different trading strategies.

By rebuilding the E-Order Book for London SETS for July and August in 2007, the third chapter reconstructs the demand and supply curves to analyse the market impact and the volatility with special focus on orders' cancellations. More specifically, this paper investigates how cancellations on both sides of the order book contribute to the market impact (price dynamics) explanation. It, also, aims at investigating the cancellations behaviour on both sides of the order book and their effect on market volatility (denoted by the bid-ask spread). To this end, different model specifications are defined and tested. All analysis is based on ultra-high-frequency data from the order book which covers the periods where the market was experiencing a downtrend price change (from peak to trough). Robust regressions[1], which control for efficiency under more realistic conditions, are run to check the validity of the regression results obtained.

The remaining of this thesis is organised as follows. Next section presents the first chapter, which studies "European Commission Transaction Tax and HFT". While the third section highlights the second chapter which discusses "Price Trends and Intraday Liquidity Dynamics in Limit Order Book", the fourth section emphasises the third chapter of this thesis which is "The Effect of Cancellations in Electronic Order Book". Section five concludes the thesis.

[1] See Andrews (1974).

2 Chapter Two

European Commission Transaction Tax and HFT

Abstract

The regulatory debate concerning high frequency trading has led the European Commission (EC) to suggest implementing a financial transaction tax (FTT) on all stock transactions. This proposed tax is set to be 0.1% on all stock transactions and will be implemented at the beginning of 2014 according to the EC. Building on the model created by Daniel and Cappellini (2006), we simulate a pure limit order market in order to find out how the proposed FTT would influence the market in general and high frequency traders (HFTs) in particular. We concentrate in this study on the main trading session during the trading day, namely the continuous double auction. Comparing the behaviour of the market before and after implementing this tax, I found support to the regulatory policy of the EC regarding introducing FTT. In particular, I found that traders tend to submit orders and start to trade more aggressively by increasing the volume traded in order to remain in a profitable position after considering the tax. The market, nevertheless, ended up with higher trading volume, lower bid-ask spread and almost same price volatility.

Keywords: Order Book Simulation; High Frequency Trading; Volatility; Financial Transaction Tax.

1. Introduction

For several decades, there has been an increased attention among economists with regard to financial economics in general and the market quality in terms of high frequency trading (HFT) in particular. This interest can be evidenced by the bulk of suggestions found in the literature that investigate and explain the role played by HFT in the financial markets, mainly in terms of the changes in market quality (Cvitanic and Kirilenko, 2010). Most academic studies on HFT pay extra attention to examining the effects on market quality. The majority of these papers have empirically investigated this impact and suggested that there is no evidence of negative effects of HFT on market quality. On the contrary, positive effects of HFT were noticed by a myriad of studies concerning different market parameters, liquidity and volatility (Gomber et al., 2011; Carroll, 2012).

High frequency traders (HFTs) are becoming key players in equity markets due to large trading volumes carried out by them. However, the variation of trading volumes, high volatility of intraday prices (possible trigger of Flash Crashes) and the massive growth in limit order submissions and cancellations caused by HFTs have necessitated strict regulations to control the market (Hagströmer and Nordén, 2012). For many years, HFT firms have undertaken their transactions rapidly and gained billions without being investigated. They were criticised and accused of damaging markets and hurting regular investors (Guo, 2012).

Nerudová and Dvořáková (2014) point out that as a result of the Financial Crisis, discussions about the possible taxation of the financial sector have started in the European Union following the United States in 2008. The EU Member States individually committed for a total of EUR 4.6 trillion[2] in 2009 to support the financial sector. Those public interventions

[2] This forms 39% of EU-27 GDP.

have significant budgetary consequences. That has imposed heavy burden on the present and future generations especially in Greece, Spain or Italy. Therefore there is a strong consensus in the European Union and internationally that financial sector should contribute more fairly to the public finance. In fact, recently there has been an ongoing debate on re-regulating the market structure. Moreover, it has been argued that taxes could be used as regulatory tools to mitigate the causes of the crisis, i.e. complex interaction of market failures, global monetary and financial imbalances and weak supervision.

Carroll (2012) argues that the problems aligned with HFT are due to the existing market structure not because of HFT itself, the thing that needs to be seriously considered by regulators. As regulators have begun to deal with the proposed problem, the European Commission is considering applying a financial transaction tax (FTT) of 0.1% on all stock transactions. The European Commission believes that the financial sector should contribute more fairly to cover the crisis costs. This contribution comes in a form of cancelling the VAT exemption of most financial services and starting participating in these costs. Although HFT is believed to be good for market quality, HFTs would be negatively hit by such a tax as low-margin HFT strategies would remain unprofitable. More specifically, any regulatory policy would mainly affect market makers if it is directed at HFTs.

The International Regulatory Strategic Group (IRSG) report, in August 2017, indicates that the tax does not differentiate between HFT and other forms of trading. Although Mortgage Backed Securities, Collateralised Debt Obligations, Credit Default Swaps and leverage of the banks were the assets which led to the crisis, none of them are covered under the FTT[3], which is instead aimed at HFT. Accordingly, it is very important for investors and interested

---

[33] See the Centre for Policy Studies website. Available online at: https://www.cps.org.uk/blog/q/date/2017/08/07/the-drawbacks-of-the-financial-transactions-tax/.

researchers to test the impact of this tax on stock markets and specifically on HFTs. This can be done using simulation platforms as there is no data yet. The hypothesis to be tested is HFTs would be negatively hit by such a tax as their trading costs will increase, and therefore the volatility will increase.

This study is an attempt to simulate a stock market with large number of interacting agents. This simulation seeks to investigate the aforementioned impact on the market in general and on HFTs in particular. In order to implement the empirical investigation, the present study implements a continuous double-auction mechanism, where agents can submit, asynchronously and at any time, limit or market orders to a single public book. Orders are sorted by price and then by time, as on the London Stock Exchange for instance. The simulation is conducted in the Matlab environment and builds on the model created by Daniel and Cappellini (2006) to implement the simulation. This simulation is based on a pure order-driven market where asks and bids are directly matched in the limit order book. Given the details of the market microstructure, the model used in this study is entirely defined by the characteristics of the agents responsible for generating the order flow, which makes it generic to be applied to any market.

The remainder of this paper proceeds as follows. Section 2 presents a brief survey of the literature that investigated HFT: its regulations and its impact on the market along with analysing the FTT. Section 3 provides a brief description of the main procedures of market design and data simulation that are used in this study. Results of data analysis and discussions obtained by checking the impact of FTT implementation on the stock market are given in section 4 while the last section, section 5, concludes and gives some suggestions for future research.

2. Litrauture Review

This section provides a brief survey of the literature that investigated HFT, its regulations and its impact on the market. It also presents the importance of the market regulations for participants and markets. This comes in line with emphasising the concept and the goal at which the authorities are aimed to be achieved. Two different case studies, however, are presented in which the two possible outcomes of implementing such a tax are represented, success and failure. These cases are (1) the Swedish case with a securities transaction tax (failure) and (2) the British case of applying a stamp duty (success). Finally, a comment on these experiences is provided.

2.1 Brief Survey

Interestingly, massive public attention is recently placed on HFT since the U.S. 2010 Flash Crash and the large trading volumes' increase associated with HFT strategies. In the academic literature and due to the data availability matters, little attention is paid to HFT (Brogaard, 2010; Gomber et al., 2011). Recently, there has been considerable effort on questioning market quality and HFT. Most of the research which has been carried out on HFT has highlighted the market impact of HFT on market quality measures. Besides, another bulk of studies has inspected the suitability of the current market regulations in light of HFT. However, there is still a gap between the results of academic literature and the beliefs predominant or sustained in media, public and even regulatory discussions on HFT impact on markets.

Additionally, Cvitanic and Kirilenko (2010) have investigated the 2010 Flash Crash. They include low frequency traders (LFT) (i.e. humans) in line with HFTs (i.e. machine to count for the speed of submitting or cancelling orders, modelled as an uninformed trader) to model an electronic market. They suggest that the HFTs may have an impact on the average transaction price and, hence, may change the whole transaction price distribution. Moreover, they glean some evidence on lower transaction price volatility hand in hand with improving the transaction price prediction. They, essentially, state that speed may explain the increase in market liquidity measures according to trading volume and inter-trade duration. These suggestions have been confirmed by many studies, such as Brogaard (2010), Hasbrouck and Saar (2010), Hendershott and Riordan (2011) and Groth (2011).

Carroll (2012), in turn, investigates the open issues on HFT in line with the available empirical literature, i.e. the widely debated issue of the benefits of HFT and its regulations. Briefly, he provides an evaluation of the regulations set to organize HFT after studying different market quality issues in the context of HFT and Flash Crash such as manipulation, liquidity, volatility and market efficiency. He concludes that the problems aligned with HFT are due to the existing market structure not because of HFT itself. Moreover, he states that HFT plays a positive role in the financial markets. He also suggests that regulators have begun to set policies that push towards clarifying and providing better understanding of HFT and transparency. He adds that the current, and proposed, regulations are a step in the right direction in that they are aimed towards mitigating risks related to HFT manipulation that comes from structural defects. Lastly, he suggests that regulators must not consider only those traders who are engaged in HFT but rather they must consider all traders who use automated strategies.

Guo (2012), besides, points out that HFTs are trying to give a good impression to regulators, public and other investors by stepping into the light. Full electronic trading in many stock markets has encouraged the researchers to pay more attention to automated trading strategies design and its impact on market quality. Considering the bad reputation of HFT in the media even with research evidences that suggest the good impact of HFT (see Carroll, 2012), authorities were urged to take an action regarding HFT and review the current regulations. These actions are supposed to ensure liquidity and prevent any manipulation. The European

Commission has introduced a FTT which in turn has raised many concerns about the impact of this tax on traders (HFTs in particular as they initially supposed to benefit the market) and financial markets. This tax has been announced by the European Commission (in 2012 to be applied by 2014) where it considers charging a FTT of 0.1% on all the exchanges of bonds or shares and 0.01% on derivatives contracts.

In the context of regulations, Chlistalla et al. (2011) make the point that there were a comprehensive number of legislations passed in both the US and Europe in the years preceding the crisis in an attempt to re-regulate the securities markets[4]. They acknowledge that Markets in Financial Instruments Directive (MiFID) is the representative of the regulation authority of securities markets in Europe where the Regulation National Market System (RegNMS) is its counterpart in the U.S. Both MiFID and RegNMS seek to strengthen the regulatory structure of the equity markets in Europe and the U.S. by pushing towards more competition between market participants. They are both supervising the markets and legislating updated regulations to achieve efficient, more integrated and liquid financial markets across all securities markets in Europe and the U.S. In this notion, it is worthy to cite Hanson (2011) who simulates a continuous double auction stock market and finds a justification of the regulators' concern regarding HFT and its impact on market quality. He finds that adding more participants who are HFTs to the market leads to higher volatility in the market especially in times of high volatility or unstable markets. In other words, HFTs would exaggerate the problem, but not cause it, the thing that needs to be seriously considered by regulators.

Hagströmer and Nordén (2012) point out that HFTs are becoming key players in equity markets due to large trading volumes. They go further by stating that the regulatory debate

---

[4] Kirelenko et al. (2011) who investigated the 2010 Flash Crash suggest that the problem is not in HFT but in the market structure.

concerning HFT should take into consideration the importance of distinguishing different HFT strategies in order to be able to investigate their influence on market quality. They make the point that HFTs is a heterogeneous group of traders where different strategies are being used. Specifically, they (Hagströmer and Nordén) use a unique data from NASDAQ OMX Stockholm to empirically provide that distinction for equity markets. They find that in the group of HFTs, the liquidity supply of market makers[5] is more often than opportunistic HFTs. Furthermore, market makers have higher order-to-trade ratios, lower latency, and lower inventory than opportunistic HFTs but they are both (in terms of strategies) mitigate intraday price volatility. However, the variation of trading volumes, high volatility of intraday prices (possible introduction to Flash Crashes) and the massive growth in limit order submissions and cancellations caused by HFTs have demanded strict regulations to control the market. On the other hand, they suggest that any introduced regulations, i.e. the FTT proposed by the European Commission, should take into consideration hitting market makers and therefore increasing market volatility.

2.2 The Financial Transaction Tax (FTT)

The debate of the securities transaction tax has been undergoing for a long time (Baltagi et al., 2006). In the context of the economic and financial crisis, the European Commission make the point that the financial sector should contribute more fairly to cover the crisis costs. This contribution comes in a form of cancelling the VAT exemption of most financial services and starting participating in the costs of the crisis (European Commission Website). As a consequence, a transaction tax on financial instruments has been proposed to be implemented by the beginning of 2014. The European Commission states that its proposal is mainly aimed

[5] Market making is one of the strategies used by HFTs (see Jones (2013) for more details).

at ensuring a fair contribution of financial institutions in the crisis costs as well as avoiding internal market fragmentation for financial transactions (equal for all participants). The European Commission documents that “[t]he rates of FTT to be applied by Member States may not be lower than (minimum rates): 0.1 % for all financial transactions other than those concerning derivatives agreements; 0.01 % for all financial transactions concerning derivatives agreements.”

Besides, Jones (2013) addresses that this proposal is aimed at raising revenue for the government but at the same time limiting HFT and other “excessive” trading. He also points out that this proposal is introduced by some policymakers who are still in doubt of the HFT value. He makes the point that transaction tax would increase the trading costs for investors and would also cause a fall in stock prices. This fall is due to the tax evaluation each time on the same share of stock when traded. Market participants would push stock prices to fall to preserve the same returns after-tax by the same amount they lost from their returns to the tax. He goes further and claims that transaction tax would negatively hit the stock market liquidity by increasing the bid-ask spread. This result is because market-makers would need to gain the tax by increase their bid-ask spreads. However, Schulmeister (2009) argues in favour of implementing such a financial tax. He suggests that implementing a small financial tax, which would be between 0.01% and 0.1%, would lead to mitigate price volatility over the short and the long run. This tax would also provide a considerable amount of revenues to governments.

Interestingly, Jones (2013) provides an example of a 1% transaction tax which has been introduced by Sweden in 1984 and then repealed after a short time. The effect of this introduction on the returns’ behaviour of Swedish equity during the period between 1980 and 1987 was investigated by Umlauf (1993). Consequently, after implementing the tax, the volatility was supposed to decline in response to the tax but not decreased. Stock price levels and turnover, however, dropped. Two years later, Sweden decided to increase the tax to 2%.

This rise results in a fall in the stock market by 5.3% at the time in which the tax was initially announced associated with about 60% movement of trading volume from Stockholm to London. Then, this transaction tax was removed in 1991 with a loss of the trading share equals to that which moved to London. Generally, regulators should be careful when implementing such a tax and try to learn from previous examples.

2.3 Case Studies on Financial Transaction Tax

In fact, any case study on financial transaction tax may be seen as misleading if, and only if, it is used to draw inferences and applications for a different environment. Studying international examples of tax implementations, however, would help us anticipate the possible changes in capital markets as a result of any tax activation. The Swedish experience with a securities transaction tax will be studied here, as a failure example, in line with the successful British application of a stamp duty.

2.3.1 The 1984 Swedish securities transaction tax

As legislation, Securities Transaction Tax (STT) in Sweden was effective from 1984 to 1991. The Swedish experience in the STT is widely treated in the financial literature as a failure (Schulmeister et al., 2008). In this section, we provide a brief picture for the Swedish STT experience and attempt to illustrate and identify what was behind its disappointing performance.

The Swedish financial sector has performed a dramatic growth in the early 80s the thing that accelerated the jealousy of the labour sector in the country. At the time, the labour sector was concerned about the "unjustifiable" earns of the young finance professionals compared to others where there should hypothetically be equal incomes in the society. This concern initially was based on the seemingly unproductive tasks, in terms of economic and social contributions, that these professionals performed according to the labour sector's point of view. Hence, they claimed a direct tax on domestic brokerage service providers without arguing in the inefficiency of the outcomes which result from trading in financial markets (Habermeier and Kirilenko, 2003 and Umlauf, 1993).

The story of this tax began in January 1984 when Swedish labour sector in 1983 claimed it. Sweden authorities decided after this proposal by the labours to introduce a 0.5% tax (50 basis points) on equities transactions on both purchases and sales which means 1% for the round trip. The Parliament approved this decision considering labour's pressure more than the objections and resistance of the Swedish Finance Ministry and business sectors. As a result, registered Swedish brokerage services were subject to this tax as they are holding the significant size of exchanges in the market. Moreover, when using Swedish brokerage, the tax is not only payable by domestic customers but also foreign customers had to pay. No tax was levied in case of transacting without dealer. In other words, small traders who trade infrequently or in small volumes are not subject to this tax (tax free) as well as gifts. Exchanges between brokers at the time of tax adoption were not included in the taxation category until 1987 as market makers were initially not considered as final consumers of domestic brokerage services. Rather, they were considered as intermediaries (Umlauf, 1993; Campbell and Froot, 1994; Schulmeister et al., 2008).

Additionally, Habermeier and Kirilenko (2003) and Schulmeister et al. (2008) acknowledge that the tax rate on stock options was 2% (200 basis points) for the "round-trip", i.e. 2% is a result of 1% taxed at the option premium in addition to another 1% for the exercise of the option. This additional 1% was justified in a way that exercising the stock option makes it a transaction in the underlying stock and therefore becomes subject to tax charge. The popular awareness about how much transactions are useful in different financial instruments has been reflected in tax coverage and rates, i.e. those transactions which involve equity options being the least useful. Generally, introducing or increasing any tax will lead assets to be devaluated by the investors accordingly in order to put the present value of the future tax payments into consideration. On the same day, an index fall of 2.2% was a result of the introduction of the tax.

The labour sector has placed more pressure, in early 1986, on the Parliament to drive it towards reconsidering the transaction taxes issue. Both the Finance Ministry and the financial sector held their positions in further opposing of any tax increase. Again, the labour sector achieved another victory over its opponents despite their opposition. This resulted in a tax increase of another 1% to be 2% on equity transactions on July 1, 1986 (Umlauf, 1993). This increase initiated a 0.8% index fall only compared to 2.2% at the time in which the tax was firstly introduced (Schulmeister et al., 2008).

On the other hand, tax revenues during the period 1984-1989 were not considered as bad in terms of growth but disappointing in levels though (see Table 1). Tax revenues were respectively SEK 820 million, SEK 1.17 billion, and SEK 2.63 billion, SEK 3.74 billion and SEK 4.01 billion for the years from 1984 to 1988. This corresponds to the following ratios of the total tax revenue in Sweden: 0.37, 0.45, 0.96, 1.17, and 1.21 percent respectively for the corresponding period (Bijlsma et al., 2012).

	1984	1985	1986	1987	1988
Equity tax revenues	0.82	1.17	2.63	3.74	4.01
Equity trading volume	71	83	142	125	115
Transaction tax rate	1%	1%	1%, 2%	2%	2%

Table 1: This table presents the Swedish transaction tax revenues and trading volume, 1984-1988 (SEK billions)

Source: Umlauf (1993), p. 4 (230 from the journal)

Introducing a new transaction tax increases each transaction’s cost (Campbell and Froot, 1994). This, therefore, may affect the investors’ behaviour in a way or another trying to avoid as much tax as possible. The possibility of not trading at all is always there though.

Schulmeister et al. (2008) make the point that the budgetary failure of the tax is mainly due to tax avoidance which also has effects on other taxes such as capital gain tax. Umlauf (1993) demonstrates that the huge drop in the capital gains tax revenues was a result of lower trading levels. He went further by indicating (p.229) that “[t]rading in Swedish government debt (which was also taxed) suffered so severely that taxes on bond trading were eventually removed. The interest rate options market evaporated with the imposition of taxes.” Thus tax avoidance negatively hit public revenues. Lower revenues were deepened after the migration of the most actively traded shares of large companies abroad under the pressure of higher costs.

On April 1, 1990, the turnover tax on fixed-income securities was abolished following its introduction on January 1, 1989. A percentage of 60% of the spot market trading volume (bonds and bills) has been recovered following the tax removal. Futures market recovered slowly for bonds but not for the bills. Tax rates were reviewed in Sweden on all remaining transaction taxes and cut by one half before their removal on December 1, 1991. The bad design was the major reason behind the failure in Swedish turnover tax which led to migration of trading volume. Moreover, tax avoidance were made relatively easy by making the tax liability subject to using Swedish brokers only whereas it is worldwide when trading in UK companies (Campbell and Froot, 1994; Schulmeister et al., 2008).

2.3.2 The British securities transaction tax (stamp duty)

Recalling the Swedish experience in STT which is widely considered in the literature as a failure, there still another STT case which is considered as a success. The British STT or stamp duty is fairly different from its Swedish counterpart. It is a worldwide tax in terms of coverage not only subject to the usage of domestic trading/brokers services. Besides, stamp duty is payable upon the ownership transfer of the financial asset. In other words, tax applies when registering the ownership of a financial instrument. Campbell and Froot (1994, p.280) argue that “an STT along British lines would be far more workable than a Swedish-style STT.” However, they also suggest that British-style STT is still likely to have critical behavioural affects additionally to the very optimistic revenue figure of $10 billion for a tax of 0.5%.

Stamp duty has been set to different rates few times over time. It was 2% before lowering it to 1% in August 1963. Then, in May 1974, it was set to 2% again the thing that did not last for so long. In April 1984, it has been set to 1% before setting it at a level of 0.5% in October 1986. Since 1986, when the “stamp duty reserve tax" (SDRT) was introduced as a replacement for the stamp duty, the tax rate has been set and kept at 0.5%. Both transfer documents and transacting agreements became subject to this tax under the rate of 0.5%, while it is 1.5% as an "exit charge". The introduction of SDRT restricted the tax avoidance and made the tax unavoidable even when trading overseas. Only shares which are transferred to clearance services to effectively avoid stamp duty and/or converted to financial products are subject to this "exit charge". Transactions in ordinary shares[6] are subject to the stamp duty as well as assets convertible to shares (i.e. unsecured loan stock). Futures and options transactions, nevertheless, are excluded. Exercising an option at the exercise price is taxable as it is considered as a purchase of ordinary shares. Few exemptions from stamp duty are made such as registered charities and market makers registered by the London Stock Exchange (LSE) when making a market. Member firms of the London International Futures and Options Exchange (LIFFE) were also exempted when hedging equity options positions or satisfying delivery obligations following the equity options exercise. This exemption is also available for intermediaries, who are defined as natural liquidity providers, trading on any UK recognized

[6] Any formal document produced is subject to the stamp duty tax.

investment exchange (Campbell and Froot, 1994; Saporta and Kan, 1997; Schulmeister et al., 2008).

In the era between 1997 and 2001, stamp duty has scored a much higher revenue growth than other taxes in the stock market before its fall in 2002 to 2004. Additionally, Schulmeister et al. (2008) addressed the extremely low tax collection costs for stamp duty where the average collecting cost is only 0.02 pence for each pound for SDRT compared to 1.11 pence per pound for all taxes collected by Inland Revenue. This is because of the electronic transactions system of the LSE which automatically applies tax on transactions. They also suggest few steps for the UK tax authorities to efficiently collect stamp duty when UK companies are listed on foreign stock markets. For example, charging interest on the stamp duty payable when returning the relevant legal documents to the UK on the share transactions that take place overseas.

2.3.3 Remarks on the tax introduction

Comparing the British with the Swedish experience clarifies the crucial role of the tax design. It is important not to face the large substitution effects which can be done by levying stamp duty paying no attention to the trade location or the investor. However, the British stamp duty is not perfect as there are potential adverse effects that need to be accounted for. Nevertheless, legislators, when setting any tax regime, should always assure that tax avoidance measures are kept at the minimum in order to pave the way for the tax to work well. This is because tax avoidance is a threat in terms of revenue bearing in mind that investors would adapt their behaviour accordingly. They should also keep in mind the academic evidence which indicates that (Jones, 2013) higher transaction taxes will lead to higher volatility, lower price efficiency, and worse liquidity. Put differently, STT is positively related with trading costs which may cause trading to move offshore as well as hitting stock prices.

3. Motivation

Nerudová and Dvořáková (2014) point out that as a result of the Financial Crisis, discussions about the possible taxation of the financial sector have started in the European Union following the United States in 2008. The EU Member States individually committed for a total of EUR 4.6 trillion[7] in 2009 to support the financial sector. That has imposed heavy burden on the present and future generations of some member states especially in Greece, Spain and Italy. Therefore there is a strong consensus in the European Union and internationally that financial sector should contribute more fairly to the public finance. In fact, recently there has been an ongoing debate on re-regulating the market structure. Moreover, it has been argued that taxes could be used as regulatory tools to mitigate the causes of the crisis, i.e. complex interaction of market failures, global monetary and financial imbalances and weak supervision.

"Naturally estimated revenues may vary considerably depending on the tax rate but also on the assumed effect of the tax on trading volumes. An official study by the European Commission suggests a flat 0.01% tax would raise between €16.4bn and €43.4bn per year, or 0.13% to 0.35% of GDP. If the tax rate is increased to 0.1%, total estimated revenues were between €73.3bn and €433.9bn, or 0.60% to 3.54% of GDP", European Commission (2011).

As regulators have begun to deal with the proposed problem, the European Commission is considering applying a financial transaction tax (FTT) of 0.1% on all stock transactions. The European Commission believes that the financial sector should contribute more fairly to cover the crisis costs. This contribution comes in a form of cancelling the VAT exemption of most financial services and starting participating in these costs. The International Regulatory Strategic Group (IRSG) report, in August 2017, indicates that the tax does not differentiate

[7] This forms 39% of EU-27 GDP.

between HFT and other forms of trading. Although Mortgage Backed Securities, Collateralised Debt Obligations, Credit Default Swaps and leverage of the banks were the assets which led to the crisis, none of them are covered under the FTT[8], which is instead aimed at HFT.

Carroll (2012) argues that the problems aligned with HFT are due to the existing market structure not because of HFT itself. Although HFT is believed to be good for market quality, HFTs would be negatively hit by such a tax as low-margin HFT strategies would remain unprofitable. More specifically, any regulatory policy would mainly affect market makers if it is directed at HFTs. Accordingly, it is very important for investors and interested researchers to test the impact of this tax on stock markets and specifically on HFTs. This can be done using simulation platforms as there is no data yet. The hypothesis to be tested is HFTs would be negatively hit by such a tax as their trading costs will increase, and therefore the volatility will increase.

Due to lack of real data on the effect of the proposed FTT, this study uses an asynchronous market simulator named NatLab, which stands for Natural Asynchronous-Time Event-Lead Agent-Based Platform. NatLab is a continuous asynchronous model in which thousands of individual traders interact through a central orders matching mechanism, just as it happens in real stock markets. Each trader has a unique decision function, which allows him/ her to trade at any time, to react to external news, to respond to price changes (or volume, volatility, etc.), and to consider the "fundamental price". This study is an attempt to simulate a stock market with large number of interacting agents. This simulation seeks to investigate the impact of FTT on the market in general and on HFTs in particular. In order to implement the empirical investigation, the present study simulates a limit order book of a market.

[8] See the Centre for Policy Studies website. Available online at: https://www.cps.org.uk/blog/q/date/2017/08/07/the-drawbacks-of-the-financial-transactions-tax/.

The simulation is based on a pure order-driven market where asks and bids are directly matched in the limit order book. Our market implements a continuous double-auction mechanism, where agents can submit, asynchronously and at any time, limit or market orders to a single public book. Orders are sorted by price and then by time, as on the London Stock Exchange for instance. Every agent acts as a simple trader. The model used in this study[9] is entirely defined by the characteristics of the agents responsible for generating the order flow, which make it suitable for different European markets.

4. Data Sources, Definitions and Market Design

This section introduces the initial steps and the way followed to generate and prepare the data for analysis purposes. It also outlines the definitions upon which I rely in order to design and build the market. Briefly, it states how the artificial market is built.

4.1 Data Source

In order to implement an empirical investigation, the present study has relied upon simulated data which is represented by detailed market data. In fact, this study simulates a limit order book of a market in order to extract useful information from it. The simulation is conducted in the Matlab environment and builds on the platform (Figure 1) created by Daniel and Cappellini (2006) to implement the simulation. This simulation is based on a pure order-driven market where asks and bids are directly matched in the limit order book. It aims to provide a better comprehension to traders' behaviour, intraday market events and price changes at the level of limit order book. Also, it aims to identify and forecast the impact of implementing a financial transaction tax on the market as well as on the traders, HFTs in particular.

[9] See Daniel and Cappellini (2006) for more information on the agents' order placement and cancellation strategies, as well as their patterns of activation.

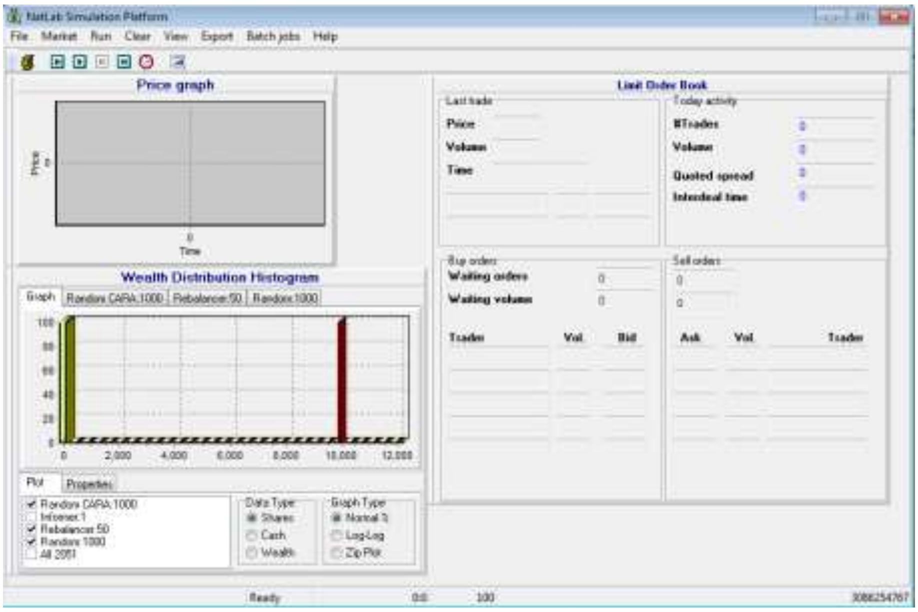


Figure 1: Daniel and Cappellini (2006) simulation platform, "Natlab" where the five best bids and asks, last trade, total market activity and Bid-Ask spread can be observed.

4.2 Definitions and Market Design

For simplicity, I assume that the built market trades only one stock. For one stock in a specific time t, the limit order book can be described as follows:

$$\beta_n \leq \cdots \beta_3 \leq \beta_2 \leq \beta_1 < \alpha_1 \leq \alpha_2 \leq \alpha_3 \leq \cdots \alpha_m \qquad (1)$$

where β_i and α_j represent buy limit orders (bids) and sell limit orders (asks) respectively (Hasbrouck, 1993). Each limit order from the above has a limit price, a size and a time stamp for arrival time in the book. All limit orders, additionally, are queued and are waiting in the book to their turn to be executed where the book depth and the tick size are assumed and set to 1024 orders and 0.01 Pound respectively. These queued orders are usually arranged or sorted according to their price, arrival time and then their size. This process would vary from market to market, but this study applies the LSE's principle which is the price, arriving time and then order size respectively (or first come first served basis). The difference between the highest bid β_1 and the lowest offer α_1 is called the spread.

$$Spread = \alpha_i - \beta_i \qquad (2)$$

Accordingly, β_1 will be executed if it arrives market or limit sell order to the book which offers a lower price than β_1. If this is the case, the transaction will be completed and β_1 becomes the new market price.

This study's dataset, therefore, covers all the transactions, bids, offers, volume and all market details which are incorporated with this stock. The feature of this artificial stock market is that agents or traders are engaged in continuous trading session through a centralised limit order book. The number of agents, besides, is constant and will be set to N =2050 traders who are all risk-averse agents (this will be described in the next section). The main difference between market simulated data and the real market data is the identification of agents. In other words, any agent who completes a transaction can be identified in the simulated market. In real market data, however, identity is not available subject to regulations and anonymity reasons.

Due to the interaction between agents' demand and supply in the market, the price of the stock will fluctuate. This fluctuation is purely due to the non-linear interactions between traders' demand and supply and is highlighted in the limit order book. For the sake of simplicity again, I assume that the market is a closed market, i.e., there is no money or information inflow or outflow along with constant number of traders. With this in mind, this assumption will let the price tend to change randomly around the steady state equilibrium price P*. Initially, this price is set by the total available amount of cash and shares and is defined as:

$$P^* = \frac{N*C_0}{N*S_0} = \frac{C_0}{S_0} \qquad (3)$$

where for each agent/trader of the N traders, C_0 and S_0, respectively, denote the initial endowments of cash and shares.

Alternatively, the trading day can be calibrated to the way real markets are organised, for example, the LSE trading day (this day excludes the opening and closing auctions) which lasts for 7.5 hours from 9:00 am to 4:30 pm. Thus, the duration of trading day, in seconds, can be directly estimated to T = 27000 seconds, which corresponds to 7.5 hours of continuous trading × 60 minutes × 60 seconds. The total trading days simulated forms two weeks of official trading plus one day. Put differently, we run the simulation for 300,000 seconds (Figure 2) assuming that each trading day equals to T = 27000 seconds as mentioned above to end up with roughly 11 days of continuous trading, i.e., approximately two weeks of real trading.

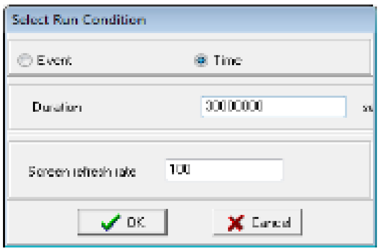


Figure 2: Setting the simulation run time

To control for external events or incoming information in general, I run each simulation twice with and without tax assuming: (a) a fully closed market with no news effect or no external events are allowed once, and (b) allowing for incoming information to the market the other time. The probability of receiving good news is assumed to be Pr = 0.5; whereas Q = 1 – Pr represents bad news. Looking back at the 2010 Flash Crash, the idea of controlling for the news (see Figure 3 below) becomes justifiable in that on the day of the Flash Crash, the U.S. stock markets opened down and trended down most of the day. Leis (2012) acknowledges that on that specific day, financial markets have experienced a pressure at market open caused by the upsetting political and economic news concerning the European debt crisis in general and doubts over the Greek economic (debt) problem in particular. Furthermore, Zhang (2012) after analysing a high frequency data from NASDAQ suggests that HFT profits are affected by news shocks.

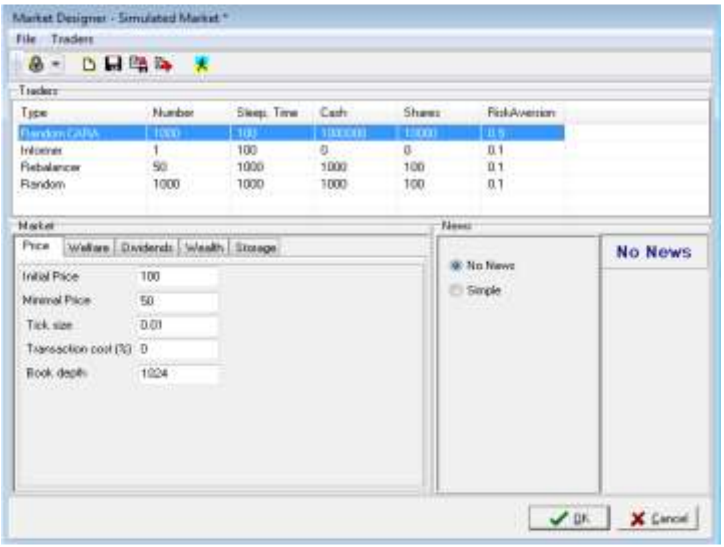


Figure 3: This figure shows the traders population and how to control for news flow, tick size, book depth and transaction tax.

4.3 Market Participants

The number of agents previously mentioned, i.e., N =2050 agents, has been considered in the market analysis so as to take into account the diversity of the traders which is likely to be determined by the number of contracts transacted by each trader. Therefore, these types of traders will be included in the market which this study simulates. This artificial market has been totally set with agents who are mainly liquidity providers. Random traders are risk-averse agents and trade with the utility function of:

$$U_i(\pi_i) = -e^{-\rho_i \pi_i} \qquad (4)$$

where ρ_i is the CARA coefficient (see Vives (2008) for detailed CARA-Gaussian model). The return π_i for each trader i is computed according to the CARA-Gaussian model as:

$$\pi_i = (\theta - p)x_i \qquad (5)$$

from buying x_i units at price of p and share random fundamental value of θ. I classify market participants into three categories according to the manner by which each group trades. The market, hence, mainly includes: HFTs group, intermediaries and opportunistic group and small traders group. The identification of HFTs is of importance. HFTs identification, in the literature, has been done following either exchanges id or data-driven definitions classifications. Thus, the current study will follow Kirilenko et al. (2011) who use a data-driven classification to identify HFTs as the top 7% of intermediaries with highest trading volume. In general, HFTs are the most skilled group (see Vacha et al., 2012) where I assume that this group represents traders who traded the top 10% to 20% of the highest volume (auctions) in the market. This assumption combines between Kirilenko et al. (2011) and the law of vital few for Pareto.

Market type / stats	No tax, no news	With tax, no news	No tax, with news	With tax and news
Volume mean	41.19	40.19	22.25	22.41
Volume stdev	76.12	84.60	62.89	55.41
Volume mode	5	5	5	5
Volume 80th percentile	97	93	9	12
Volume 90th percentile	127	125	91	94

TABLE 2: This table shows the frequency statistics for the mean, standard deviation, mode and percentiles used to determine the truncation value

Technically, dataset will be truncated according to transacted volume at level of the value of the 80th percentile (TABLE 2) which corresponds to each market type in order to allow for HFTs group's investigations. The reason behind why we care for the top 20% traders is Pareto Efficiency Law or Law of the Vital Few (80-20 rule) in that studying 20% of the participants could make inferences about 80% of market changes. This is because the top 20% are the richest (see United Nations Development Program, 1992). On the other hand, small traders are those who traded 5 shares or less each time, whereas intermediaries and opportunistic traders are those who fall in neither of the abovementioned groups. Agents can only view or access the current state of the market, namely current market price, best bid and best offer available. Agents, accordingly, can place market or limit orders randomly around viewed prices. Furthermore, based on the fact that stock markets' activity is affected by external events or news (see Leis, 2012; Kirilenko et al., 2011), I introduce in the simulations, additionally to random agents, an informer trader to allow for informed trading. Random agents are market liquidity providers while informed traders are subject to place orders aggressively to the book.

5. Simulation Results

The major aim of this section is to introduce, highlight and analyse the results conducted from the simulated data. These findings will be analysed vis-à-vis the target objective of this study, namely investigating the impact of the FTT proposed by the European Commission. This can be implemented by analysing the simulated data as a whole to address the impact of this proposed tax on the market trading activity, or specifically, by analysing the transacted volume during the trading period.

Beyond the prevision of the dataset acquired, data truncation will be employed in order to address the tax impact upon large traders. For so doing, the following procedures have been followed. I have run the simulation four times. The simulation has been run for the first time before implementing the transaction tax along without allowing for external events flow (closed market assumed). The second time, on the other hand, I have run the simulation after implementing the transaction tax but without allowing for the flow of external events, too. As a result, the market remains closed. Consequently, it is possible to investigate the impact of the transaction tax on the market in general and on the traders in particular. It is not always the case, however, that the market is isolated from the external environment. This ultimately leads to re-run the simulation two more times. Since there is an impact for the external events or incoming news on the market, I have run the simulation for the third time allowing this time for external events to flow but not allowing for the tax to be implemented. The simulation has been run for the last time, not surprisingly, after allowing for the transaction tax and flow of external events together.

5.1 Closed Market (no news flow)

In general, the cases in which the simulated market have been run without allowing for news flow are studied. More specifically, these cases include two datasets for simulated market: one is before implementing the transaction tax, whereas the other is after implementing the transaction tax.

4.1.1 before transaction tax implementation

Bid			Ask		
Rank	Volume	Price	Price	Volume	Rank
1	5	80.86	80.87	5	1
2	146	80.66	80.90	5	2
3	1	80.55	80.92	63	3
4	5	80.54	81.00	152	4
5	5	80.50	81.04	5	5
Total Number of Trades		117 384	Total Number of Volume Traded		4 834 749
Number of Trades carried out by top 20% (including HFTs Group)		25 526	Number of Volume Transacted by top 20% (including HFTs Group)		3 882 322
Number of Unexecuted Buy Orders which is transferred to the Next Day		312	Number of Unexecuted Sell Orders which is transferred to the Next Day		285
Volume of Unexecuted Buy Orders		45 647	Volume of Unexecuted Sell Orders		40 561

Table 3: This table depicts the limit order book snapshot when market is closed and tax is not imposed. It shows the five best bids and asks on both sides of the market at the end of the last trading day simulated with further information from the book during the two weeks of trading. The volumes carried out from each side of market are reported at the end of the last trading day simulated.

As Table 3 above best illustrates, it can be noticed that the difference between the highest bid β1 and the lowest offer is α1 is: α1−β1 = 80.87 – 80.86 = 0.01. Since this number, which is Bid-Ask Spread, is relatively small, it means that the stock is highly liquid. Furthermore, according to Mills and Markellos (2008), the volatility of the return of the stock could be conducted by applying the following formula given that p_t is the price at time t:

$$\widehat{\sigma}^2 = \frac{1}{n-1}\left(\sum_{t=1}^{n} r_t^2 - \frac{\log(p_n/p_0)^2}{n}\right) \qquad (6)$$

where σ^2 is the true return variance and p_0 and p_n are respectively the first and last prices observed while r_t is the return which is calculated as:

$$r_t = p_t - p_{t-1} \qquad (7)$$

This reveals that the mean-adjusted estimator corresponds to $\hat{\sigma}^2 = 0.0255$. That is to say the price volatility of this stock is equal to 15.98%. Figure 4 below indicates the stock price movement during the simulated period. It shows a slight decrease in the stock price in the first week, while it mirrors nearly a linear movement with a low volatility.

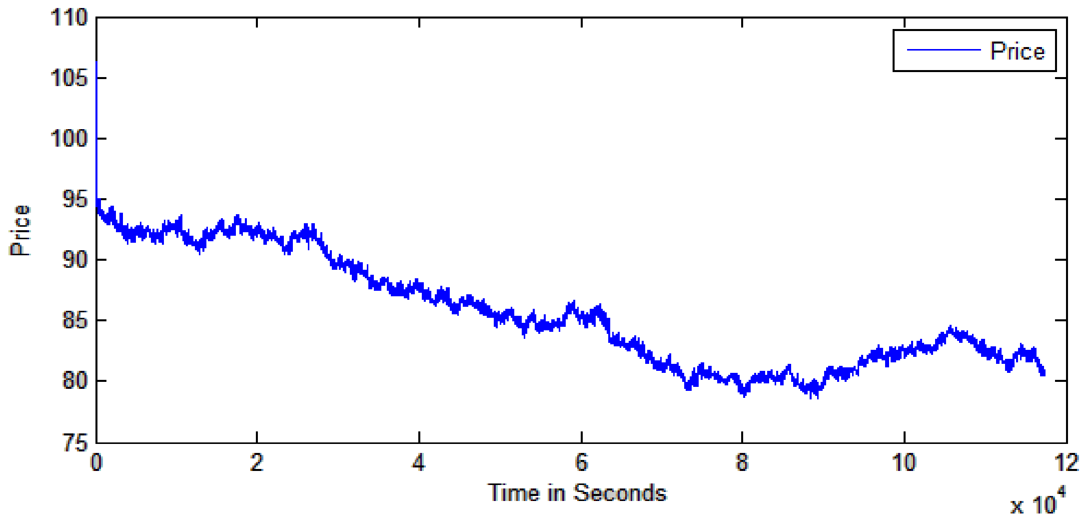


Figure 4: The stock price movement during the simulated perioud before allowing for tax

4.1.2 after transaction tax implementation

Bid			Ask		
Rank	Volume	Price	Price	Volume	Rank
1	4	91.90	91.93	9	1
2	5	91.88	92.48	4	2
3	5	91.83	92.49	5	3
4	5	91.83	92.50	12	4
5	114	91.83	92.66	120	5
Total Number of Trades		132 739	Total Number of Volume Traded		5 334 513
Number of Trades carried out by top 20% (including HFTs Group)		27 365	Number of Volume Transacted by top 20% (including HFTs Group)		4 304 650
Number of Unexecuted Buy Orders which is transferred to the Next Day		291	Number of Unexecuted Sell Orders which is transferred to the Next Day		273
Volume of Unexecuted Buy Orders		42 867	Volume of Unexecuted Sell Orders		39 206

Table 4: This table shows the limit order book snapshot when market is closed and tax is imposed. It presents the five best bids and asks on both sides of the market at the end of the last trading day simulated with further information from the book during the two weeks of trading. The volumes carried out from each side of market are reported at the end of the last trading day simulated.

A deep examination of Table 4 discloses that the difference between the highest bid (β1) and the lowest offer (α1) in this case is: α1−β1 = 91.93 – 91.90 = 0.03 compared to 0.01 before implementing the transaction tax. Bid-Ask Spread, interestingly, is relatively small which means that the stock is highly liquid. The volatility of the stock return, besides, (i.e. the mean-adjusted estimator) is equal to $\hat{\sigma}^2 = 0.0342$. This ultimately means that the price volatility of this stock is equal to 18.46% compared to 15.98% before implementing the tax (a small difference though). Additionally, comparing the total number of volume traded in the market during the period of study before and after implementing the transaction tax suggests that this number has increased by nearly ≈ 11% (from 3,882,322 to 4,304,650) after implementing that tax. One possible explanation for this increase is that HFTs and part of the intermediaries start to trade aggressively in order to remain in a profitable situation. Furthermore, the result goes in line with Colliard and Foucault (2012) who studied a limit order market in which traders have to pay a trading fee to the market owner before they start trading. It also matches the findings of Malinova and Park (2011) who based their analysis on trading fee changes on the Toronto Stock Exchange (TSX) on two distinct dates in 2005. The stock price movement during the simulated period is displayed in the Figure 5 below. The price reflects higher volatility compared to the previous market which was running before implementing the tax. However, this volatility is still small with a price downtrend during the simulated period.

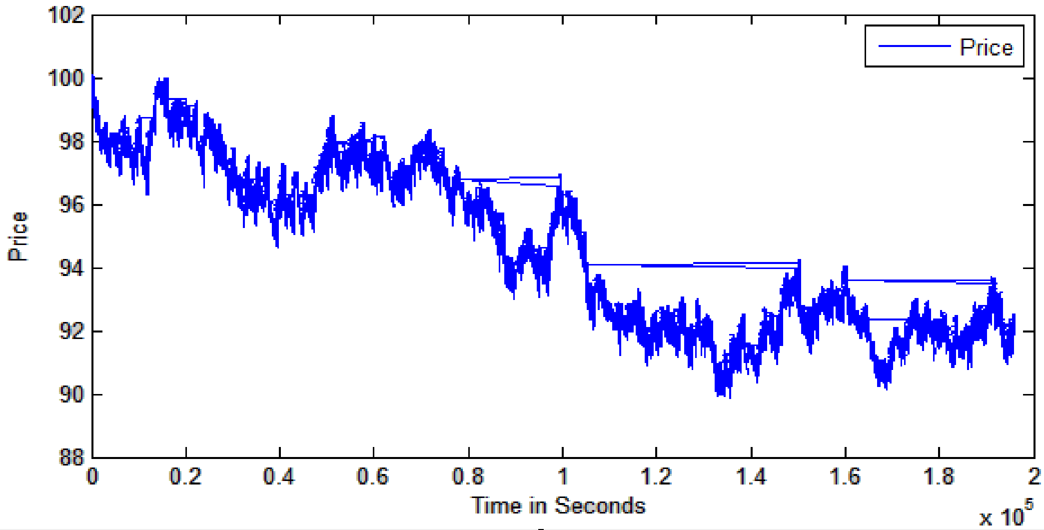


Figure 5: The stock price movement during the simulated perioud after allowing for tax

5.2 Open Market (with news flow)

In a related manner, I consider the case in which the simulated market has been run allowing, this time, for news flow. On a deepened level, this case, also, includes two datasets for simulated market: one is before implementing the transaction tax while the other is after implementing the transaction tax.

4.2.1 before transaction tax implementation

Bid			Ask		
Rank	Volume	Price	Price	Volume	Rank
1	5	85.22	85.76	110	1
2	57	85.21	85.80	159	2
3	5	85.15	85.88	163	3
4	5	85.15	85.95	143	4
5	5	85.14	85.98	134	5
Total Number of Trades		192 271	Total Number of Volume Traded		4 278 590
Number of Trades carried out by top 20% (including HFTs Group)		28 883	Number of Volume Transacted by top 20% (including HFTs Group)		3 520 603
Number of Unexecuted Buy Orders which is transferred to the Next Day		415	Number of Unexecuted Sell Orders which is transferred to the Next Day		329
Volume of Unexecuted Buy Orders		58 641	Volume of Unexecuted Sell Orders		49 509

Table 5: This table depicts the limit order book snapshot when market is open and tax is not imposed. It shows the five best bids and asks on both sides of the market at the end of the last trading day simulated with further information from the book during the two weeks of trading. The volumes carried out from each side of market are reported at the end of the last trading day simulated.

Looking at Table 5, it reveals that the difference between the highest bid (β1) and the lowest offer (α1). This difference is calculated to be equal to: $\alpha 1 - \beta 1 = 85.76 - 85.22 = 0.54$. It is of interest to state that the Bid-Ask Spread has increased after allowing for the external events compared to the market where news flow is not allowed. However, the spread is still relatively small, the thing that means that the stock is still considered as highly liquid. The volatility measure of the stock return, besides, (i.e. the mean-adjusted estimator) has been calculated to be equal to $\widehat{\sigma}^2 = 0.0045$. This eventually means that 6.73% represents the price volatility of this stock. Compared to the measurement value of volatility in the closed market which equals to 15.98%, the stock price volatility becomes clearly lower after allowing for informed trading or external events flow. Thus, we can argue in favour of the good impact for HFTs on the market. This goes in line with many researches who have gleaned some evidence upon the good impact of HFT on the market quality (e.g. Chaboud et al., 2009; Brogaard, 2010; Hendershott et al., 2011; Hendershott and Riordan, 2011). Figure 6 below shows the stock price changes during the simulated period where it demonstrates a downtrend price movement for the whole simulated period.

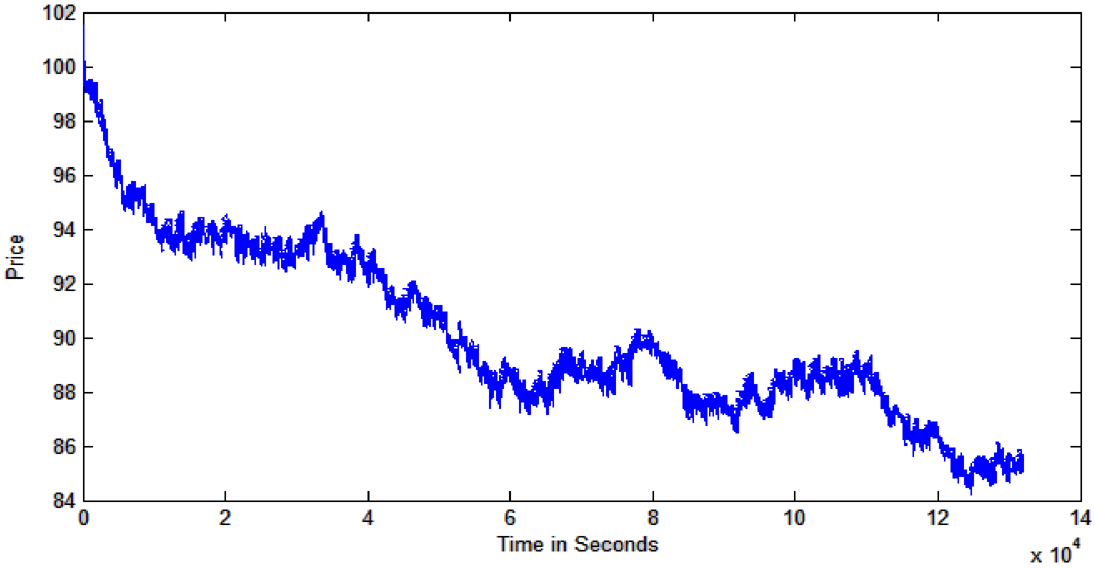


Figure 6: The stock price movement during the simulated period after allowing for news and before allowing for tax

4.2.2 after transaction tax implementation

Bid			Ask		
Rank	Volume	Price	Price	Volume	Rank
1	134	90.07	90.28	22	1
2	170	90.00	90.31	246	2
3	46	89.77	90.34	135	3
4	123	98.72	90.38	101	4
5	155	98.72	90.38	104	5
Total Number of Trades		196 589	Total Number of Volume Traded		4 406 250
Number of Trades carried out by top 20% (including HFTs Group)		29 585	Number of Volume Transacted by top 20% (including HFTs Group)		3 635 196
Number of Unexecuted Buy Orders which is transferred to the Next Day		329	Number of Unexecuted Sell Orders which is transferred to the Next Day		311
Volume of Unexecuted Buy Orders		49 755	Volume of Unexecuted Sell Orders		43 448

Table 6: This table depicts the limit order book snapshot when market is open and tax is imposed. It shows the five best bids and asks on both sides of the market at the end of the last trading day simulated with further information from the book during the two weeks of trading. The volumes carried out from each side of market are reported at the end of the last trading day simulated.

Table 6 above displays that the difference between the highest bid (β1) and the lowest offer (α1) is α1−β1 = 90.28 – 90.07 = 0.21. The Bid-Ask Spread in this case, interestingly, is relatively small which, also, means that the stock is highly liquid. After allowing tax implementation, the spread becomes smaller, lowering from 0.54 to 0.21 compared to 0.03 for the closed market where news flow is not allowed. The mean-adjusted estimator for the volatility of the stock return equals to $\widehat{\sigma}^2 = 0.0052$ which means that the price volatility of the stock is equal to 7.23% compared to 6.73% before applying the transaction tax where they seem to be indifferent. The volatility numbers suggest that volatility measures have not been strongly affected by the tax implementation. Moreover, the total number of volume traded in the market during the period of study before and after implementing the transaction tax suggests that this number has increased by nearly 3% (from 4,278,590 to 4,406,196 ≈ 3%) after implementing that tax. Observing lower spread, roughly same volatility and higher trading volume, I would suggest pushing forward into transaction tax implementation. This suggestion stands partially against the claim made by Hagströmer and Nordén (2012). They argue that the FTT proposed by the European Commission would increase market volatility by hitting market makers in a way that makes most HFT strategies remain unprofitable. Figure 7 below demonstrates the stock price movement during the simulated period. The stock price had a smooth downtrend in the first week of trading, the thing that started to change, also smoothly, into an uptrend afterwards with a low volatility.

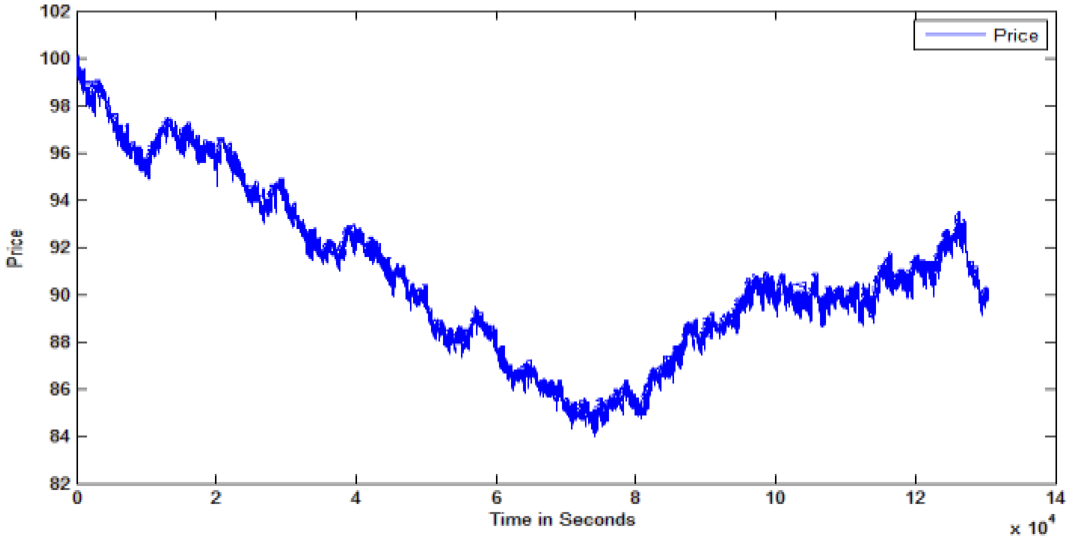


Figure 7: The stock price movement during the simulated period after allowing for news and tax

To sum up, for the purpose of analysis, and since this study has exclusively focused on the impact of the FTT proposed by the European Commission on the market and its participants, the market has been limited to 2050[10] traders who traded twice: once without paying any tax, and the other with tax paying. Accordingly, market was assumed to be closed market once, where an external event or information is not allowed to flow, whereas news flow is allowed another time. This specific choice of traders could be attributed to assure the variety of traders

[10] This number is 2052 when accounting for the two informers added to the market to allow open market assumption test.

in the market. Consequently, the analysis of the different market conditions reveals that the proposed FTT may not negatively hit the market. Rather, it increased traded volume without negatively affecting the bid-ask spread or volatility in the market.

6. Conclusions

This study is an attempt to empirically examine the impact of transaction tax implementation on the stock market by modelling a pure limit order market with a heterogeneous set of players. This set of players includes traders who are liquidity takers and providers where they buy and sell a single asset in the market. Inspired by previous studies which simulate the order book and those investigating real markets, an order-driven market is simulated using a double auction in order to let supply and demand match. The artificial market n this study is populated exclusively with zero-intelligence agents who randomly place orders. Yet, the specifications of the limit order book in this artificial market has led to generate a non-random price. The novelty of the current research lies in the use of simulated order flow data to assess the impact of the proposed FTT on the market quality as well as on the market participants. This study finds support to the regulatory policy of the European Commission regarding introducing a FTT of 0.1% on all stock transactions. In particular, traders are found to tend to submit orders and start to trade more aggressively by increasing the volume traded in order to remain in a profitable position after considering the tax. The market, nevertheless, ended up with higher trading volume, lower bid-ask spread and almost same price volatility. This may suggest that the proposed tax does not affect HFTs in a negative way. On the contrary, it induced them to perform more trading volume. This may mitigate the EU concerns regarding tax implementation which would reduce the profit margin of HFT strategies and, therefore, they remain unprofitable. However, more research on this topic needs to be undertaken, preferably using real data, in order to clearly understand the association between tax and HFT.

3 Chapter Three

Price Trends and Intraday Liquidity Dynamics in Limit Order Book

Further evidence from the London Stock Exchange Electronic Order Book (SETS)

Abstract

In this study, we rebuild the London Stock Exchange Electronic Order Book (SETS) in real-time for 5 different stocks. This paper's aim is to model the shape of bid and ask sides of the order book using Notional Volume Weighted Average Price (NVWAP) concept. Using data for five different stocks from London SETS over a period of two consecutive months, July and August 2007, this paper analyses the intraday dynamics of liquidity in the London Stock Exchange (LSE) limit order book and constructs the demand and supply curves to predict future price trend. The results suggest that the shape of bid and ask side of the order book is revealed to be robust proxy to identify the prevailing market trend without prior knowledge of the price.

Keywords: Order Book Rebuild; High Frequency Trading; Volatility; Liquidity Dynamics; Price Trend; Predict Future Price Trend

1 Introduction

In recent years, the wide evolution in electronic trading has led to introduce new trading platforms in stock exchanges (Colliard and Foucault, 2012). These platforms allow traders to utilize technology to swiftly buy and sell their shares in the market. Thus, utilizing the fast automated execution and high technology, market participants complete their transactions fully electronically with identified counterparties who want to transact/trade in a matter of seconds. This ultimately sets the tone for emerging a new style of trading known as high frequency trading (HFT), where automated systems electronically buy and sell. HFT strategies update their orders very quickly (a few seconds or less) and have no over-night positions (Gomber et al., 2011).

Market microstructure is closely related to the investments field, which studies the process under which investors' latent demand and supply are ultimately translated into prices and volumes. Noticeably, the rapid technological, structural and regulatory changes in the securities industry world-wide have placed more interest in market microstructure. These structural shifts are caused by complex processes such as new technological innovations, the substantial increase in trading volume, changes in the regulatory environment and the creation of new financial instruments.

An important function of financial markets is to produce information. One important issue in recent market microstructure research has been whether knowledge of the structure of the electronic limit order book is informative regarding future price movements. Informational research in microstructure contains investigating the price formation and price discovery. Essentially, this topic is concerned with looking inside the "black box" where latent demands are translated into realized prices and volumes. There has been an extensive empirical research on the informational content of the electronic limit order book (E-LOB). This research is relevant to predicting price movements in the stock markets (see Madhavan, 2000; Kane, 2011; and Kozhan and Salmon, 2012).

As a trading mechanism of financial markets, electronic limit order markets are gaining more importance. All limit orders are collected in the order book at any given point of time. Orders come into the book throughout the day at the time they are submitted to the market. These orders are removed from the book as they are executed, cancelled, or expired. To match supply and demand on a daily basis, stock exchanges all over the world use trading rules and mechanisms. Following progress in technology, those rules and mechanisms not only differ between exchanges but can also change in time. Generally, traders use computer based algorithms to buy (sell) a position while endeavouring to stick to a client's selected benchmark. Trading algorithms are continuously looking for profitable opportunities to trade. Therefore, analysing each market event and its impact on liquidity at a micro-level is of great interest to market participants (see Degryse et. al., 2005; Gilles, 2006; and Frei and Westray, 2015).

One of the oldest and most popular benchmarks used in computer based algorithms is Volume Weighted Average Price (VWAP). For several reasons, the VWAP benchmark is very popular amongst both brokers and clients. It is very simple to calculate, so it facilitates the post-trade reporting. It helps breaking large orders into smaller ones which reduces market impact/volatility as it mitigates the demand for large liquidity. Besides, it is considered as a "fair benchmark price" at any given time interval. Malik and Markose (2012) have introduced the concept of Notional Volume Weighted Average Price (NVWAP). It refers to the average expected transaction price for a given order of volumes priced at different prices. They argue that the information conveyed by the NVWAP curves and more specifically the price trend indicator (the statistics about the state of the limit order book) defined in their work is quite useful as an indicator of prevailing market trends. They emphasise that the change in the slope of the NVWAP curves as well as the change in the total volume on each side of the limit order book can be used in an algorithmic trading strategy to exploit different trading opportunities.

The objective of this study is to analyse the intraday dynamics of liquidity in the London Stock Exchange (LSE) limit order book for five different stocks over a period of two consecutive months, July and August 2007. These two months fairly represent the period before the Financial Crisis in 2007 whereby unlike August (when the financial crisis triggered), the month of July is a fairly stable month. This study will model the shape of bid and ask sides of the order book using NVWAP concept using different stock holdings from the London SETS. To this end, I replicate Malik and Markose (2012) methodologies to construct the demand and supply curves for different stocks from LSE to predict future price trend. After that, I perform out-of-sample tests using different time intervals to see how accurately this indicator would predict future price trend. Then, I compare the obtained results against their findings.

This study proceeds as follows. The following section presents a brief survey of the literature that investigated electronic order books and computational market microstructure. Section three describes London SETS Data. Data and methodology are discussed in the fourth section. Section 5 focuses on the analysis and results. Lastly, section 6 concludes.

2 Literature Review

This section surveys the technological and technical improvements that proceed alongside with financial markets. It, also, covers the HFT and its impact on financial markets. Then, it briefly discusses the E- LOB and its informational content. Finally, it describes VWAP as a trading benchmark.

2.1 Technology, Algorithmic Trading and High Frequency Trading

Briefly, over the years, automated order execution systems have continued to be developed allowing instant information processing and setting the human intervention or action taking to the minimum. These systems, as a result, become known in financial industry as “algo” or algorithms (Brownlees et al., 2011). With respect to Algorithmic Trading, it is noteworthy to highlight that Chabound et al. (2009) defined it as trading platforms which are directly run on computers at high frequency basis. Hendershott and Riordan (2009) consider the process of making trading decisions, submitting orders and adjusting submitted orders using computer algorithms as algorithmic trading. Put differently, the term algorithmic trading was generally defined by Domowitz and Yegerman (2005) as equity orders executed automatically by computers using direct market-access channels with the goal of achieving a specific standard. They, also, imply that programme trading and rules-based trading could be used as alternative terms to demonstrate algorithmic trading. Moreover, Chordia et al. (2008) make the point that over the recent years, algorithmic trading was behind the increase in trading volume as well as the drop in the average trade size. The increase in trading volume may explain the concerns about the impact of algorithmic trading on price discovery process and volatility.

Cliff et al. (2011) place particular attention on market impact when executing large orders. They suggest that reducing this market impact is one of the most important motivations which inspire developing automated execution systems. Arguably, execution of very large orders by a trader affects the market price as it encloses a sudden sharp increase in supply (demand) which may cause a negative (positive) impact on price and an interaction from the counterparties. In other words, they define the case when the price moves against the trader before his deal is done as market impact. This negative price movement is due to the size of his order along with quick execution. Briefly, it has become ordinary for market participants to fully electronically complete their transactions in a matter of seconds. This ultimately sets the tone for emerging a new style of trading known as high frequency trading (HFT), where automated systems electronically buy and sell given that they will not hold a particular position for a long instance (a few seconds or less).

In sophisticated technology driven markets, HFT is an important aspect when it comes to price formation process. In the recent years, a significant amount of scientific and technical research has placed more attention on optimizing more reliable systems to deal with HFT based on high frequency data (HFD) (Easley et al., 2012). Briefly, Leinweber (2009) points out that always there are some traders or investors who are much faster than others due to technological breakthroughs used by these traders. Within a similar vein, Easley et al. (2012) and Fabozzi et al. (2011) agree with Leinweber (2009) that speed is a very important factor to the success of HFT.

2.2 High Frequency Trading and Market Impact

Interestingly, public attention is recently immensely placed on HFT since the U.S. May 6th, 2010 Flash Crash along with the large trading volumes increase of HFT strategies (Gomber et al., 2011). Easley et al. (2012) state that the ‘Flash Crash’ of 2010 pushed HFT into the spotlight. Accordingly, in sophisticated technology driven markets, HFT is an important aspect when it comes to price formation process. A significant amount of scientific and technical research has placed more attention on optimizing more reliable systems to deal with HFT based on high frequency data (HFD). Briefly, Leinweber (2009) points out that there are always some traders or investors who are much faster than others due to technological breakthroughs used by these traders. Within a similar vein, Easley et al. (2012) and Fabozzi et al. (2011) agree with Leinweber (2009) that speed is a very important factor to the success of HFT.

Cvitanic and Kirilenko (2010) have investigated the impact of HFT on market quality. They include low frequency traders (i.e. humans) in line with high frequency trader (i.e. machine to count for the speed of submitting or cancelling orders, modelled as an uninformed trader[11]) to model an electronic market. They suggest that the high frequency trader may have an impact on the average transaction price and, hence, may change the whole transaction prices distribution. Moreover, they glean some evidence[12] that lower transaction prices volatility is linked to improving the transaction prices prediction. They, essentially, state that speed may explain the increase in market liquidity measures according to trading volume and intertrade duration. Gomber et al. (2011) point out that most academic studies on HFT are concentrated on examining the effects on market quality. They add that the majority of these papers have empirically investigated this impact and suggested that there is no evidence for negative effects of HFT on market quality. In contrast, positive effects of HFT were noticed by most studies concerning the most important market parameters, liquidity and volatility.

Easley et al. (2012) go further and make the point that speed was not the only reason that stands behind the wide attention paid to HFT. For so doing, they argue that "what lies at the centre of HFT is a change in paradigm" starting from the point that reality is more complex in that "today's high frequency markets are not the old low frequency markets on steroids" (p.1). They suggest that because the authorities have legislated a new set of rules allowing for highly

[11] This follows the classical notion of symmetric information.
[12] These suggestions have been confirmed by many studies, such as Jarnecic and Snape (2010), Brogaard (2010), Hasbrouck and Saar (2010) and Groth (2011).

technological competitors to enter the market, these changes have paved the way for an "arms race" between competitors. This race is about developing technological and quantitative strategies (HFT strategies) that allow them to meet the demands of market participants and, at the same time, could excavate for the last penny of profitability from trading. Precisely, HFTs during their process of fulfilling the demands of market participants, they, most importantly, care about the volume (market share). However, measuring or accounting for time is something intuitive. In terms of HFT, the trader is utilizing information revealed and playing a game of making the best possible move (before competitors), not to move as fast as possible (computers will do that).

Nevertheless, Fabozzi et al. (2011) maintain that the speed of HFT has come out with new risks associated with it. These risks have been extensively discussed by Sornette and Von der Becke (2011). They emphasise that according to Tabb Group[13], HFT, the ultra-high-speed version of algorithmic trading, is estimated to account for over 77% of transactions in the UK market, but lower estimates of about 25% are there for futures in 2010. They find that following the May 6th Flash Crash in 2010, HFTs were subsequently mostly cleared from having caused the crash. This conclusion is confirmed by Foucault (2016). With the notion of The Flash Crash of May 6th, 2010, The Treasury Flash Crash of October 15th, 2014 and The Exchange Traded Funds (ETFs) Flash Crash of August 24th, 2015, Foucault (2016) attributes the occurrence of these events to the automation of trading and structural changes in market organisation but not to HFT per se. He also suggests that regulation of HFT should be aimed at specific trading strategies rather than fast trading in general.

[13] See http://www.bloomberg.com/news/2011-01-24/high-frequency-trading-is-77-of-u-k-market-tabb-group-says.html.

2.3 The information content of the electronic limit order book (LOB)

One important issue in recent market microstructure research has been whether knowledge of the structure of the limit order book is informative regarding future price movements. The topic of informational research in microstructure contains a very wide range of topics such as price formation and price discovery, including both static issues (e.g. the determinants of trading costs) and dynamic issues (where prices come to impound information over time). Essentially, this topic is concerned with looking inside the "black box" where latent demands are translated into realized prices and volumes. There has been an extensive empirical research on the informational content of the limit order book. This research is relevant to predicting price movements in the stock markets.

An important function of financial markets is to produce information. In his own words, Foucault (2016) explained how information is produced by stating "(t)hat is, asset prices aggregate informed investors' signals and thereby convey information for real decisions, e.g. investment (see Bond, Edmans, and Goldstein, 2012)." He suggests that making prices more informative is a potential benefit of informed trading which makes real decisions more efficient. He also acknowledges that the progress in information and trading technologies have contributed to the development of HFTs; those traders whose trading strategies rely on extremely fast reaction to market events.

There is a growing body of literature on market microstructure with emphasis on the information content of the electronic limit order book (LOB). LOB is a record of all unexecuted orders to buy (or sell) a given quantity of a stock at or below (above) a specified price (see Madhavan, 2000; Kane, 2011; and Kozhan and Salmon, 2012). Moreover, Cao et al (2009) analysed data from the Australian Stock Exchange using error correction model. They suggest that order book is reasonably informative. They found that the contribution to price discovery from the best bid and ask prices and the last transaction price is estimated at only 22%.

Moreover, they illustrated the relevance between predicting future short-term returns movements and order imbalances between demand and supply along the book. These results are confirmed by Harris and Panchapagesan (2005) who analysed data from the New York Stock Exchange and investigated whether the limit order book is informative about future price dynamics.

2.4 VWAP as a trading benchmark and the proposed improvement

Seeking to emphasise financial market success factors, Malik and Markose (2012) suggest that[14] the success depends on the market ability to determine the accurate price for traded assets. They claim that matching the demand and supply efficiently and effectively can determine a proper trading price. In order to optimally construct trading strategies, they acknowledge that a comprehensive knowledge of market microstructure is required so that quantitative modelling of price and liquidity dynamics can be done after accounting to ultimate transaction costs. Berkowitz et al (1988) have proposed the Volume Weighted Average Price (VWAP) with the aim of providing adequate methodologies to analyse and extract limit order book information content and adequately use this information for optimal trading strategies design. The VWAP is a measure of market impact and is calculated over each trading day. The VWAP is used as a benchmark to assess the trade-off between cost/benefit from market impact and the desire for immediacy of execution.

The concept of VWAP as a trading benchmark is well understood by investors and extensively used in the literature. Generally, VWAP[15] is a measure of the average price at

[14] According to: Berkowitz et al. (1988); Almgren and Chriss (2000); Kissell and Glantz (2003); Madhavan (2002a); Hasbrouck (2004) which are cited by Malik and Markose (2012).

[15] VWAP's definition given a series of prices and volumes is $VWAP = \frac{\sum_j P_j \times Q_j}{\sum_j Q_j}$ where VWAP is the volume weighted average price, P_j is the price of trade j, Q_j is the quantity of trade j and j is each individual trade that takes place over the defined period of time.

which a stock is traded over a trading horizon. It is particularly useful when to break a bulk order into several smaller orders according to the historical data and is computed to be optimum. It uses historic transaction prices and volumes (see Berkowitz et al, 1988; Bansal et al, 2010; and Kato, 2015).

Malik and Markose (2012) adjust the VWAP logic and apply their approach to the extant limit orders using data from the London Stock Exchange (LSE). Building on the VWAP and the development of quantitative framework, they proposed an enhanced measure, namely, the Notional Volume Weighted Price (NVWAP). NVWAP is a statistic that applies to the information to the electronic order books in terms of the limit orders posted on the bid and ask sides. It measures the change in the shape of empirical liquidity from both supply and demand curves and refers to the average expected transaction price. The difference between the best price and the NVWAP is the premium a trader has to pay for executing an order with a volume larger than the total available at best price. In other words, it is the cost of immediacy.

3 London Stock Exchange (LSE) Data

As a matter of fact, LSE is considered as one of the largest market exchanges in the world (Zovko, 2008). Many British and international stocks are traded in this market. Accordingly, studying the LSE is quite beneficial as it entails the common characteristics of large equity markets. LSE operates different types of systems that deal with high and low liquid securities. Two trading systems for high liquid securities are being used in the LSE, namely electronic limit order book (E-LOB) and retail service provider (RSP). Put differently, E-LOB is known as on-book market whereas RSP or quotation market is known as o-book market.

At the LSE, Stock Exchange Electronic Trading System (SETS) session is a trading service offered by the LSE for the on-book trading. Besides, SETS is based upon the electronic open limit order book and considered from the most liquid electronic order books in Europe. Stock

Exchange Automated Quotation System (SEAQ) session, on the other hand, is a trading service offered by the LSE for the o-book trading. Nevertheless, LSE offers a service for less liquid securities compared to those traded on SETS. This service is Stock Exchange Electronic Trading Service – quotes and crosses (SETSqx) (Kim, 2008; Zovko, 2008; London Stock Exchange (LSE) website). The following screenshot best illustrates how the trading platform screen looks like.

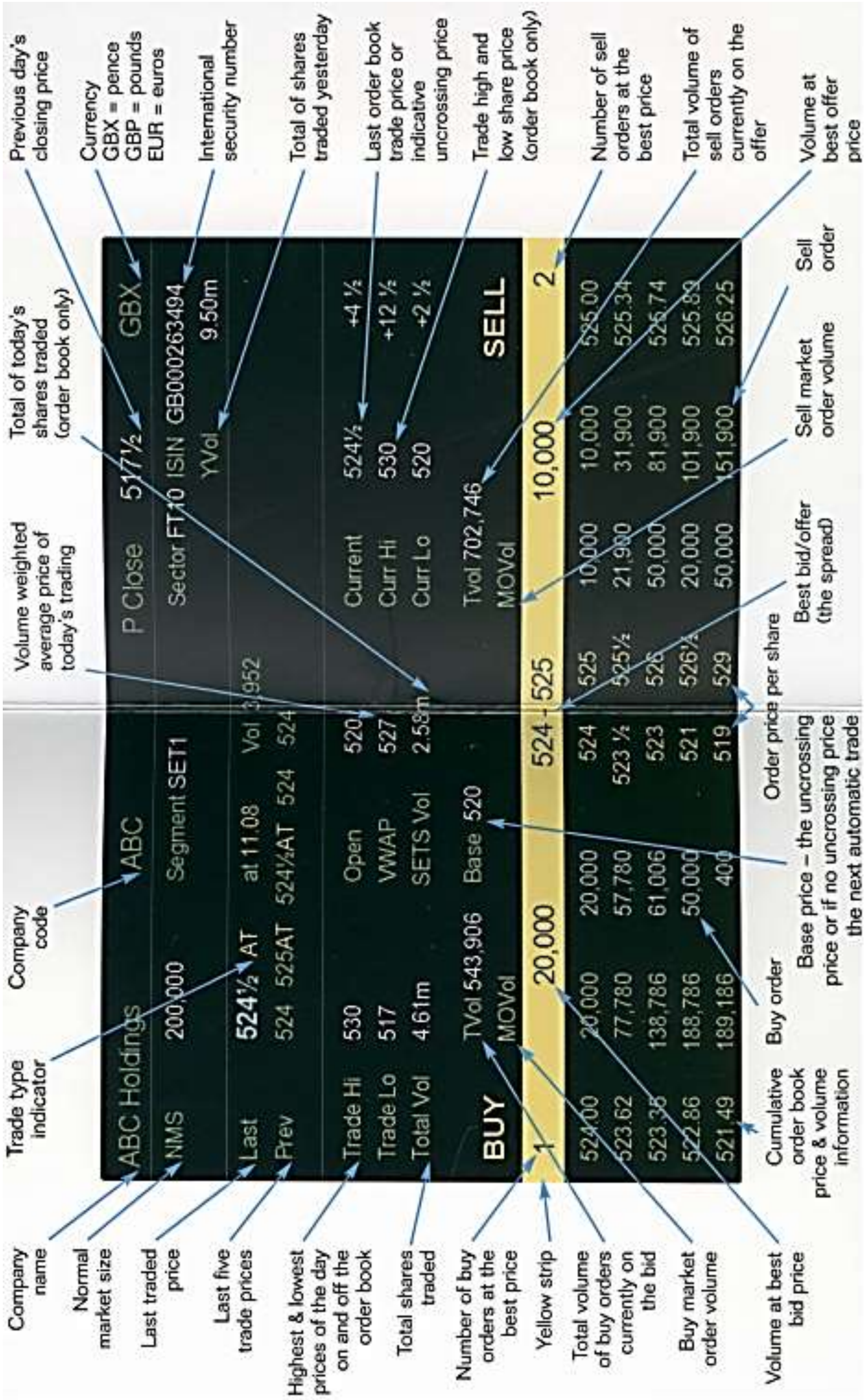


Figure 8: Trading screen example for the SETS (the electronic open limit order book of the London Stock Exchange) with some displayed traders and some general stock information. (Source: Zovko, 2008)

3.1 SETS Developments

Several order types are being traded and supported in one of the most liquid electronic order books in Europe, the London SETS. The Millennium Exchange trading system of SETS has significantly improved in terms of dealing with latency and different types of orders. These order types include:

Passive Only Orders ensure that if any part of the order can be executed, the order would be cancelled immediately when entered. This type of orders can also be used by participants to limit the accepted incoming orders to those priced within an indicated number of visible price points of the BBO. The order, besides, will also be cancelled if it becomes active due to a price change. This order type addresses participant's demand and offers additional alternatives to traders and investors.

Stop Orders, also called a stop-loss order, allow participants to set a specific price at which they will buy or sell a stock when this price is met; this specific price is usually referred to as stop price. At the point that stop price been hit, stop order will become a **Market Order** and execute at best. Market orders exceeding the volume at the best price are allowed to '*walk up the book'*. These orders are matched by standing limit orders beyond the best bid and ask prices. Buy/sell stop orders are generally used by investors to protect a profit or limit a loss.

Stop Limit Orders combine the features of both stop orders and those of limit orders. As soon as a stop price is reached, a stop-limit order will become a limit order and then executed (buy or sell) at no worse than this limit price. This order type provides investors with ultimate control over picking the best time to *fill or kill* the order. As with all limit orders, however, trade execution is not guaranteed as long as the stock/commodity does not meet the stop price.

Hidden Limit Orders [also called “iceberg” or “undisclosed” order (Moinas, 2006)] are the type that provides traders with possibility of submitting anonymous limit orders on to the order book. Traders are allowed to submit *iceberg* orders when in need to buy and sell large amounts of securities. Traders can divide their large orders into smaller chunks so that other traders only see a small portion of the full order at a time. The *iceberg* order reduces liquidity which is the impact on price movements resulted from significant changes in a stock’s supply and demand. Liquidity suppliers, accordingly, will have the ability to choose whether to display the volume, price of the order or both of them to other participants. Hidden orders are able to interact with other orders regardless whether they are displayed or hidden on the order book. Limit Orders can only be hidden when entered once they meet the Large Order Threshold (LOT) stated in Millennium Exchange Business Parameters of LSE. LOT, hence, means that orders to be hidden should be equal to or above the LOT at the point of entry. In the notion of Iceberg Order, large order is divided into smaller parts whether by traders’ choice or automatically (automated program). This action aims to reduce any unwanted price movement when traders want to buy or sell large amounts of securities. Hidden limit orders will be completely hidden tough.

Mid Price Pegged Orders allow participants to anonymously submit and price their non-displayed orders into the book at a limit price which is the true mid-price or the average of the BBO (best bid and offer). At the time of opening auction and during intra-day auctions, Pegged orders entered will automatically be injected in the next period of continuous trading. Only Mid Price Pegged Orders which meet the LOT of the LSE could be entered. Using Mid Price Pegged Orders would enable participants to be more flexible and have higher ability to match submitted orders with incoming ones (higher execution) along with limiting information leak and increasing the probability of reducing the total trading cost. If the BBO, for instance, is £12.01 to $12.02, the mid-point peg order will be priced at true mid-point at £12.015. At the end of the trading day, all pegged orders will be deleted including those which entered during the closing auction.

Executable Quotes are orders where market makers have to register in individual SETS securities in order for them to be able to enter this type of orders into order book securities. Executable Quotes entered by market makers, also, have to be named, fully visible, specifically priced and sized as required, electronically executable and dual sided quotes on entry.

In response to customers demand, LSE on its market has introduced a new functionality which is Closing Price Crossing session in April 2012 and set SETS to support this session. This functionality will provide participants with a five-minute opportunity to execute or cancel stock order at the closing price determined during the exchanges' day after the closing auction conclusion. This five-minute opportunity will take place directly after the exchange's closing auction between 16.35 and 16.40. Matched orders are executed at the closing price whereas after the closing auction, the remaining orders on either book will remain to the new closing price crossing auction to be available for execution. Orders, however, which are entered with a time-in-force value or good for auction tag will not be available for execution in the next auction. During the session, newly entered orders would be priced at the closing price or market order.

3.2 Trading Day and Order Matching Mechanism in the SETS

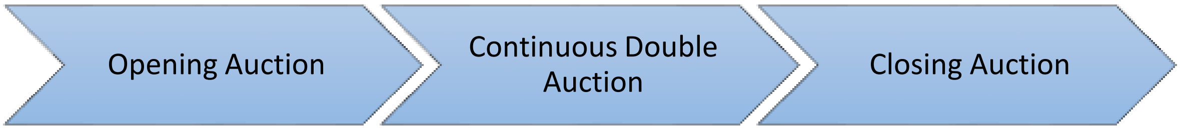


Figure 9: Sections of the Trading Day in SETS

In the LSE, the trading day of the SETS (on-book) trading session starts at 8:50 with an opening auction which lasts for 10 minutes. During the opening auction, market participants place buy and sell orders while still no execution occurs. Orders submitted by market participants (traders) are mainly differentiated by their execution priority in the SETS into two types, namely limit and market orders. An order is a limit order if the price and size of the order is predetermined. Besides, limit orders are executed if there is a match between the order price and the market price whereas non matched orders will remain in the limit order book waiting for the incoming orders arrival. Limit orders are executed at a specified price or better. On the other hand, an order is a market order if the order size is predetermined but rather it would be executed at the best currently available market price rather than a predetermined one. The priority in execution in the first place is to the market orders as they arrive to the central limit order book while the execution priority for the limit orders depends upon how far they are from the clearing price and then on the arrival order. However, not surprisingly, if the supplied volume was insufficient for the clearance of all the market orders, priority will be based upon the order submission time. The clearing price (or the market opening price) is the price calculated at the end of the auction that maximizes the trade volume. Each order will have a new timestamp each time it has a modification in its price or (and) size before execution and therefore will be assigned a new priority. Yet, the timestamp and execution priority of the order will remain the same if it is partially executed (Zovko, 2008; Kim, 2008).

Based on their trading preferences, market participants will choose between placing limit or market orders. Market order, interestingly, is more preferable to the market participants if they seek for immediate execution. This choice would be risky in terms of orders executed at undesirable price(s) though. However, if they have concerns about execution price, choosing a limit order will be better. Choosing limit order also entails a risk of long time waiting till execution, or even no execution at all.

The market enters the main trading phase which is “the continuous double auction” period after the end of the opening auction. All orders submitted during the auction and not cleared will be shifted to the order book. Zovko (2008), also, make the point that market participants can continuously submit buy or sell orders and matched transactions will be cleared directly. At the same time, cancelation of any untransacted limit orders is guaranteed for traders at any time. In the case of large market order and insufficient volume at the best price, it can be broken into multiple limit orders with multiple prices in order to be transacted.

On deepened level, SETS is being set to suspend the trading system in case of an unusual large price move (“more than 10%-20% difference from the last transaction price”). When this trading suspension occurs, SETS will enter an auction period identical to the opening auction in order to provide traders with time to process potentially new information. At the end of the trading day, trading session will start to end at 16:30 with another auction period described at the end of previous sub section.

4 Methodology and Data

In this study, I reconstruct the historic limit order books of London Stock Exchange’s (SETS) platform over a period of two consecutive months July and August 2007. I base the analysis on these two months because the month of July was a quiet month where the market was not volatile. The market in the month of August was volatile due to the Financial Crisis which started August 2007-9. In their words, Eigner and Umlauft (2015) suggest that “The

Great Financial Crisis of 2007-2009 was the most severe economic crisis since the Second World War. Only the Great Depression was similar in severity and length."

The information conveyed by the NVWAP curves is greatly useful as an indicator of prevailing market trends. This information suggests that the market activity on each side of the limit order book directly affects the shape of liquidity supply and demand curves. These curves, in turn, depict the expectations of market participants regarding current and future price movements. Hence, the cumulative impact of market events is expected to have observable changes in these curves. In this study, I will investigate the state of the limit order book using data from 5 different stocks. More specifically, I will check whether the change in the slope of the NVWAP curves, and the change in the total volume on each side of the limit order book can be used within an algorithmic trading strategy to exploit trading opportunities. Then I'll compare the results against Malik and Markose (2012) results as well as out-of-sample tests.

To this end, this study presents results for five stocks from the London SETS. The stocks are chosen to represent different market capitalisation levels and market sectors in the sample studied. The following table (Table 7) provides a basic description of these stocks. All data is based on 5 minutes intervals where $\Delta t = 5\ min$. So, each interval is of the shape (t, t + Δt) following the literature as in Weber and Rosenow (2005), Hendershott et al (2011) and Malik and Markose (2012). With this intervals approach, the impact of daily opening and closing auctions is excluded as I only evaluate price return between 8:01 AM and 4:29 PM. This approach filters out both auctions and overnight trading effect on one hand, and removes the impact of corporate actions (e.g. dividends, stock splits, etc) on the other hand.

Stock	Ticker	Industry Sector
HSBC Holdings	HSBA	Banks
British Petroleum Plc	BP	Petroleum
Tesco Plc	TSCO	Retail
Astra Zeneca Plc	AZN	Pharmaceutical Products
Vodafone Group Plc	VOD	Telecommunications

Table 7: List of stocks used in this study from different market sectors

To construct the liquidity demand and supply curves, I follow Malik and Markose (2012) methodologies where I find the turning/reversal points (identified as peaks and troughs) in the cumulative price return within each trading month of the two selected months. Peaks and troughs in cumulative price are located by setting a minimum price return change threshold of 25bp to identify local peaks and troughs. Peaks and troughs utilise the Directional Changes (DC) approach and are shown in Figure 10 below. The stock price return, r, is calculated following the formula:

$$r = \log(P_{m,t+1}) - \log(P_{m,t}) \qquad (1)$$

where P_m is the mid price, and given as $P_m = \frac{P_{bid}+P_{ask}}{2}$.

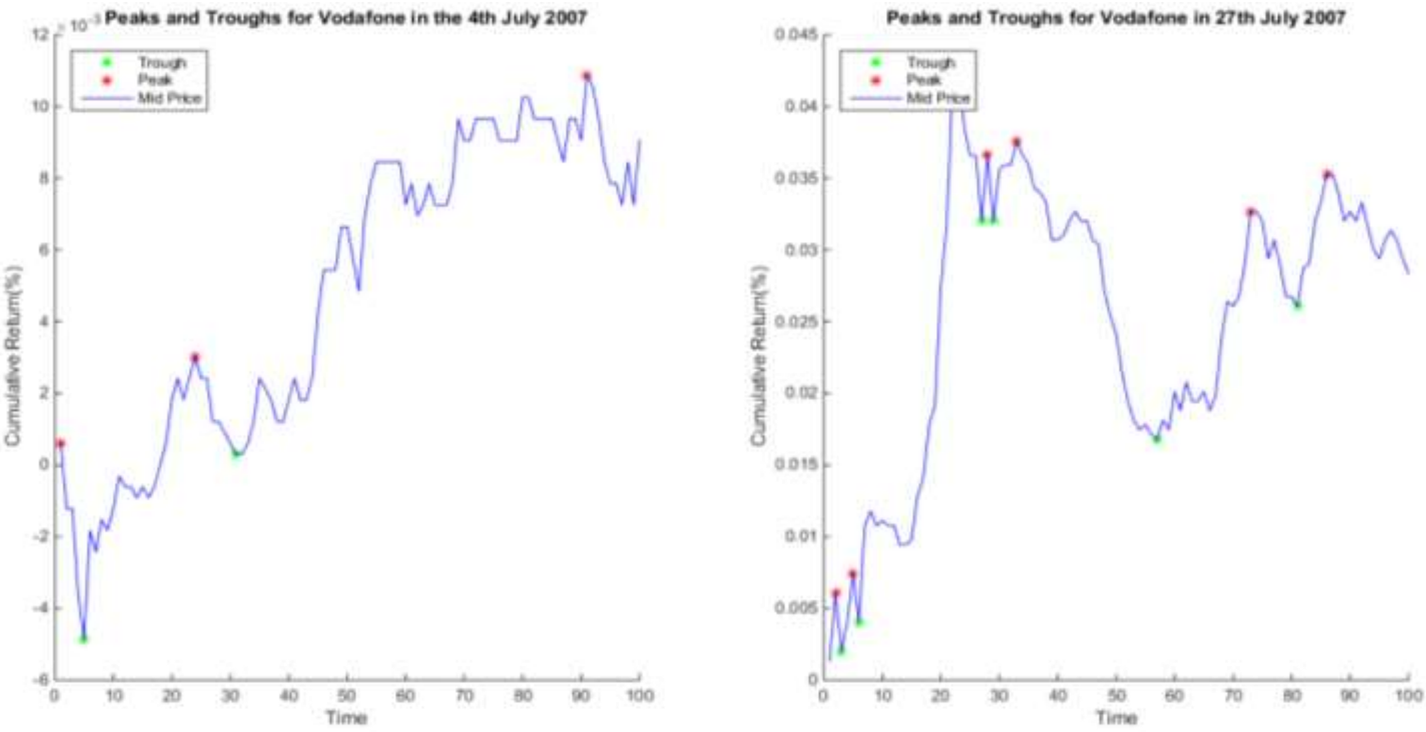


Figure 10: Peaks and Troughs calculated with a 25 basis points threshold in July 2007 for Vodafone stock in two different days. The 4th (right panel) and the 27th (left panel), where market volatility varies from quite (left) to volatile (right) market

At the selected cumulative price return peaks or troughs, I take a snapshot of the order book at time t and then calculate the NVWAP which is the average expected transaction price for a trade size equal to the cumulative volume at the i-th level as follows:

$$NVWAP_i = \frac{\sum_{j=1}^{i} P_j \times Vol_j}{\sum_{j=1}^{i} Vol_j} \qquad (2)$$

where P_j is the price and Vol_j is the volume at order j in the book.

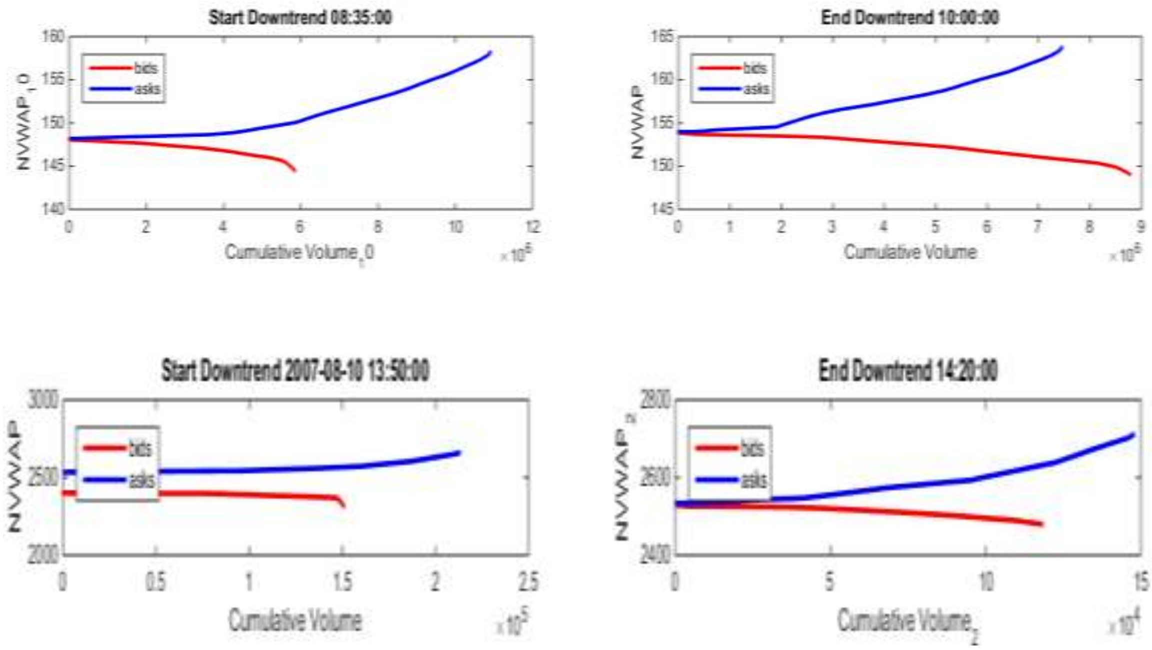


Figure 11: The NVWAP curves in the top panel show the start (8:35 AM) and end (10:00 AM) of downtrend on 27th of July 2007 for Vodafone stock. The ask side NVWAP curve gradually expands and becomes flatter; the bid side NVWAP curve contracts and steepens. Besides, the NVWAP curves in the bottom panel show the start (13:50 PM) and end (14:20 PM) of downtrend on the 10th of August for the AZN stock. NVWAP curves follow the same patterns as mentioned above.

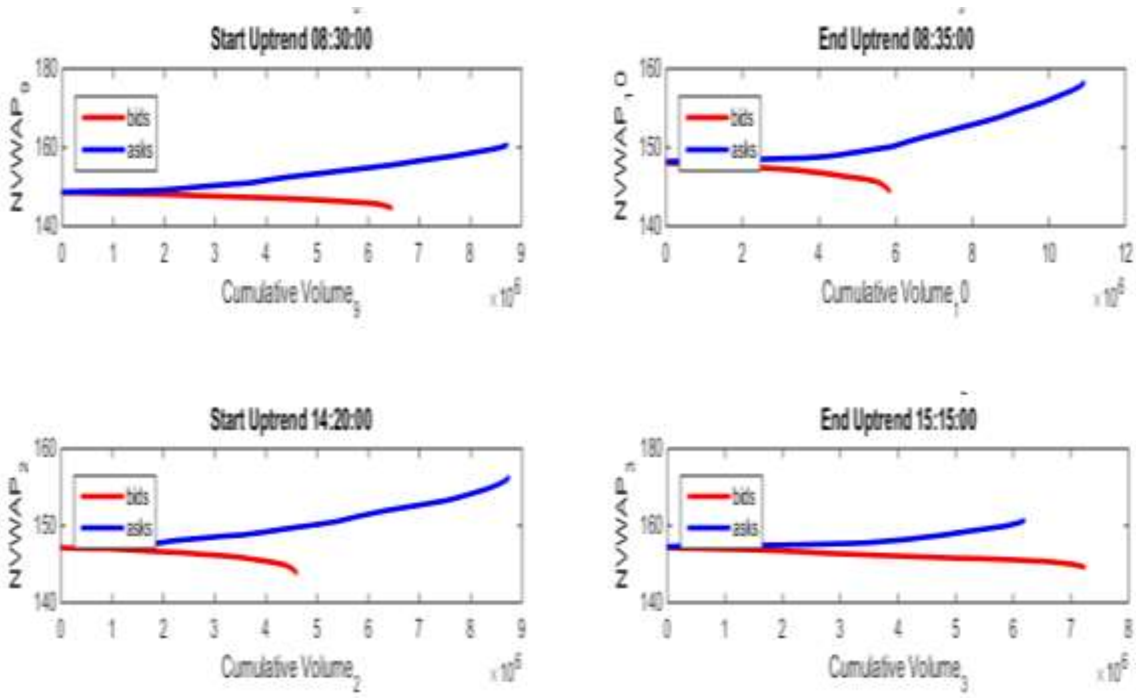


Figure 12: The NVWAP curves in the top panel show the start (8:30 AM) and end (08:35 AM) of uptrend on 27th of July 2007 for Vodafone stock. The bid side NVWAP curve gradually expands and becomes flatter. However, the ask side NVWAP curve contracts and steepens. Further, the NVWAP curves in the bottom panel show the start (14:20 PM) and end (15:15 PM) of uptrend on the 10th of August for the AZN stock. The bid side NVWAP curve and the ask side NVWAP follow the same pattern as expected.

It is expected that when the market price falls (downtrend), the bid side gradually decreases. The bid side NVWAP curve slowly contracts and becomes steeper. The NVWAP curve on ask (liquidity supply) side curve then slowly expands and becomes flatter. This effect can be seen in Figure 11 in the shape of the NVWAP curves on the bid side where the curves slowly contract and become steeper. The snapshots in Figure 12 show the opposite case, when the market price increases.

The market impact or the premium a trader pays for executing a volume larger than the amount available at best price is defined as:

$$Y_i = \left(\frac{NVWAP_i - P_m}{P_b}\right) \times 10000 \qquad (3)$$

where P_m and P_b are mid- and best-price respectively.

The steepening/flattening effect is captured by estimating the slope of the NVWAP curves. The slope of the NVWAP curve is estimated by the regression:

$$log(Y_i) = log\beta_0 + \beta_1 X_i \qquad (4)$$

where Y_i is the market impact and $X_i = \frac{\sum_{j=1}^{i} Vol_j}{ADV} \times 100$ which is the cumulative volume normalized to the i-th order at which the market side that a trader wants to trade. The exponential transformation $Y = \beta_0 e^{\beta} 1^{X}$ is performed to ensure that the supply curve is strictly increasing for $\beta_0, \beta_1 > 0$, as assumed by definition.

The hypothesis this study tests is that in an uptrend, change in slope of the ask curve is greater than the change in the slope of the bid curve $\Delta log(slope^{ask}) > \Delta log(slope^{bid})$. While in a downtrend, change in slope of the bid curve is greater than the change in the slope of the ask curve, i.e. $\Delta log(slope^{bid}) > \Delta log(slope^{ask})$. The change in the slope of the curve captures both the volume available and any additional cost (benefit) results from buying (selling) in the rising (falling) market. These changes in the slopes of NVWAP curves cause contraction/expansion and steepening/flattening effects. Capturing these effects allows to test the previous hypothesis.

Over a predefined interval, the contraction and expansion of the curves is measured by calculating the change in the total volume available on both sides of the limit order book. In an uptrend, the change in the total volume on the bid side is greater than the change in the total volume on the ask side, i.e. $\Delta log(Q^{bid}) > \Delta log(Q^{ask})$. However, in a downtrend, the change in the total volume on the ask side of the order book is greater than the change in the total volume on the bid side, i.e. $\Delta log(Q^{ask}) > \Delta log(Q^{bid})$. As the price starts rising, investors will compete to consume liquidity quickly to take advantage of the potential higher returns. This competition will attract higher volume to the demand side. At the same time, volume on the supply side will decrease as less traders become willing to sell.

To capture these effects, e.g. the steepening and flattening as well as the contraction and expansion of the curves, I will build on the DC methodology and use different snapshots for two stocks over two months, July and August in 2007. I estimate the slope coefficient (β_1) at each peak and trough covering both uptrends (from trough to peak) and downtrends (from peak to trough) intervals. Each of these uptrend and downtrend intervals has a price return of at least ±25bp. Then, I calculate the following four statistics $\Delta log(slope^{bid})$, $\Delta log(slope^{ask})$, $\Delta log(Q^{bid})$and $\Delta log(Q^{ask})$ at the start and the end of each uptrend and downtrend interval. Lastly, I will compare the different slopes and quantities at each side of the order book against each other to check whether this study's results support Malik and Markose's (2012) results.

5 Analysis and Results

This section proceeds in two divisions. Firstly, a test is implemented to the price trend indicator. Then, in the other division, an out-of-sample analysis is performed to test for the efficiency and robustness for this indicator.

5.1 preliminary analysis of the market trend indicator

As previously mentioned, this analysis is based on the order book data for five different stocks which are traded at the LSE in 2007. Over a period of two months, i.e. July and August, I identify two days at each month. These days/dates are selected depending on how volatile each day was. The aim is to test the hypothesis at different market settings. As discussed above, the month of July was fairly a stable month compared to the month August which was a volatile month due to the emersion of the Financial Crisis at the time. Following the peaks and troughs approach, I will take different snapshots as discussed in Section 4 above. I estimate the slope coefficient (β_1) at each peak or trough for the two different market states, the uptrend (from trough to peak) and downtrend (from peak to trough) intervals. Each of these intervals, uptrend and downtrend, has a price return of at least ±25 basis points to be identified. I, then, calculate the $\Delta log\left(slope^{bid}\right)$, $\Delta log(slope^{ask})$, $\Delta log\left(Q^{bid}\right)$ and $\Delta log(Q^{ask})$ statistics at the start and the end of each uptrend and downtrend intervals. Table 9, Table 16 and Table 17 show how these calculations are done.

Figure 13, Figure 14, Figure 15 and Figure 16 below illustrate the variation in NVWAP curves at start and end of each uptrend and downtrend intervals as discussed above. Two snapshots from each month are presented. These snapshots are selected at the points of highest change in returns for both uptrends and downtrends. From Figure 13 and Figure 14 below, we note that strongly similar patterns of steepening and contraction in NVWAP-Ask accompanied by flattening and expansion in NVWAP-Bid are observed at all eight downtrend intervals. From Figure 13, The $slope^{ask}$ values at start and end of the third downtrend interval 08:05 to 08:20 are 0.74 and 1.00 respectively. The Q^{ask} at these two points are 6,170,708 and 5,688,603 respectively. The $\Delta log(slope^{ask})$ and $\Delta log(Q^{ask})$ are 0.20 and -0.13. The positive $\Delta log(slope^{ask})$ reflects the steepening of the NVWAP-Ask curve and negative $\Delta log(Q^{ask})$ suggests that the curve is contracting. Also, the $slope^{bid}$ values at start and end of the third downtrend interval are 1.37 and 1.39 respectively. Q^{bid} at these two points are, respectively, 3,254,123 and 4,065,445. The $\Delta log\left(slope^{bid}\right)$ and $\Delta log\left(Q^{bid}\right)$ are 0.05 and 0.15. The positive $\Delta log\left(slope^{bid}\right)$ reflects the flattening of the NVWAP-Bid curve while the positive $\Delta log\left(Q^{bid}\right)$ value suggests that the curve is expanding.

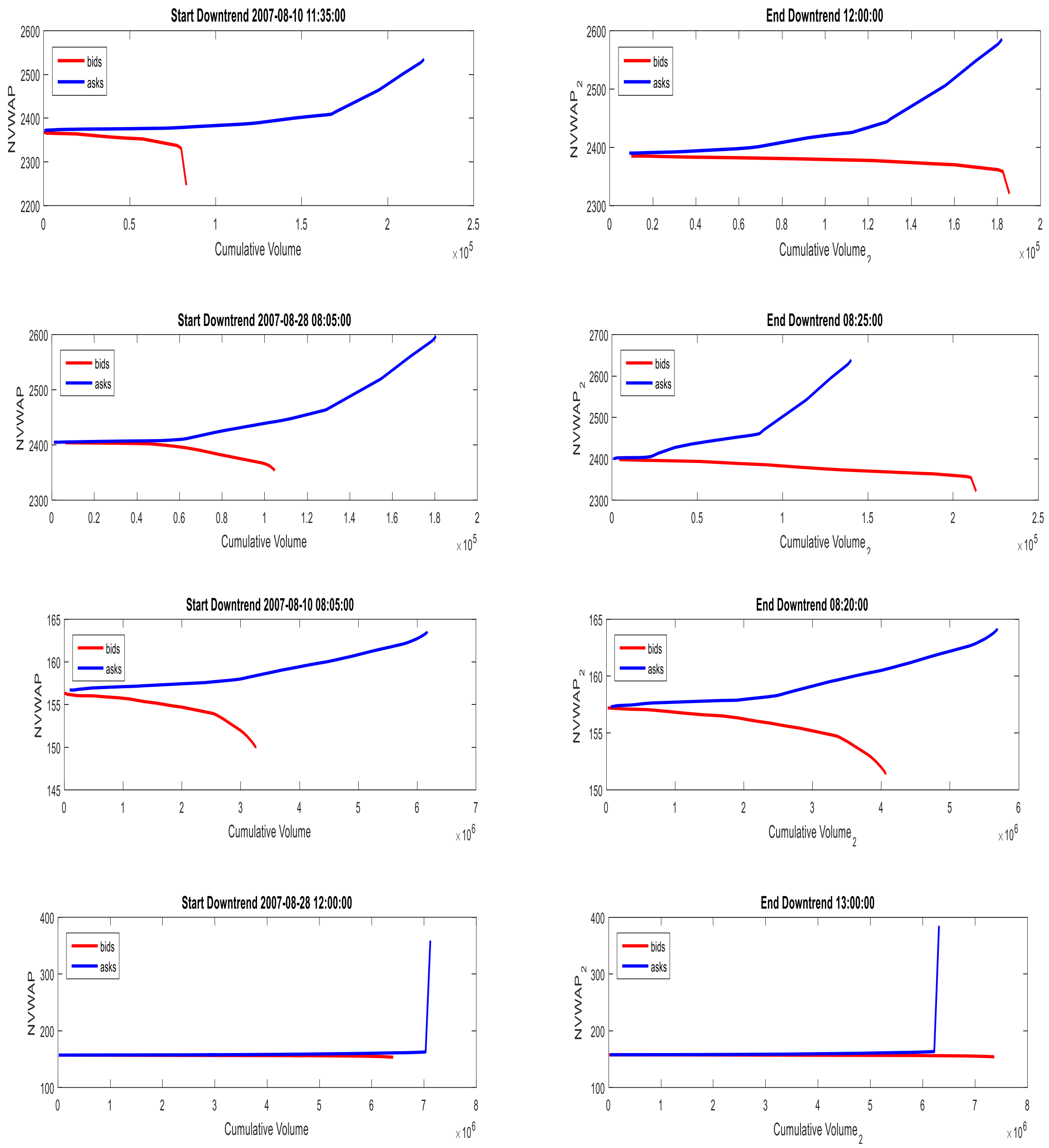


Figure 13: NVWAP curves for the stocks of AZN and Vodafone respectively at the start and end of the downtrend in the month of August 2007. The start and end timestamp for the each snapshot are noted above each curve. The top two panels are for AZN stock, while the bottom two panels are for Vodafone stock.

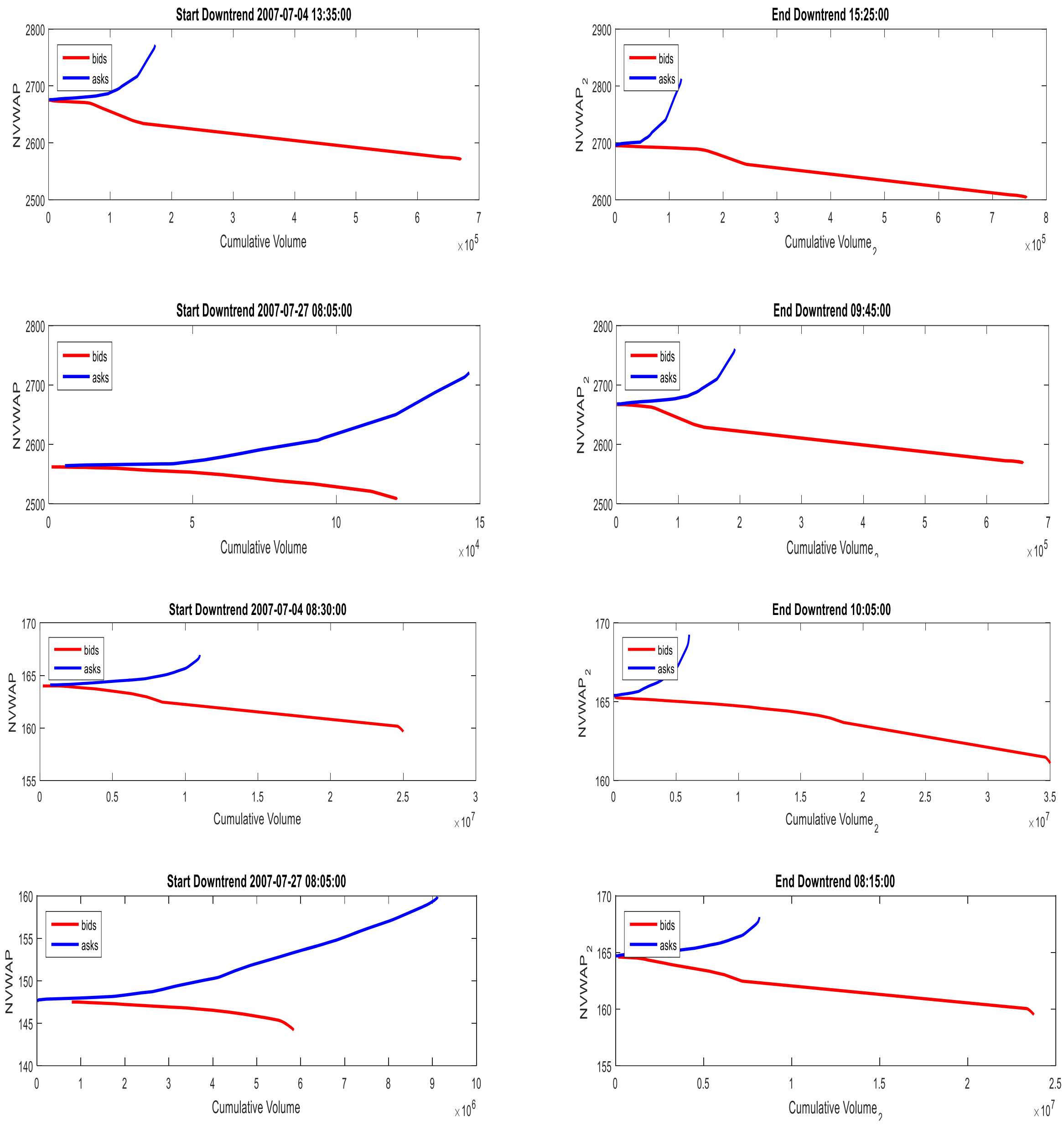


Figure 14: NVWAP curves for the stocks of AZN and Vodafone respectively at the start and end of the downtrend in the month of July 2007. The start and end timestamp for the each snapshot are noted above each curve. The top two panels are for AZN stock, while the bottom two panels are for Vodafone stock.

Similarly, from the first downtrend interval 13:35 to 15:25 in Figure 14, the $slope^{ask}$ values at start and end of the interval are 1.42 and 1.95 respectively. Q^{ask} at these points are 173,286 and 122,631. $\Delta log(slope^{ask})$ and $\Delta log(Q^{ask})$ are calculated to be 0.07 and -0.05 respectively. The NVWAP-Ask curve steepens due to the positive value of $\Delta log(slope^{ask})$. Besides, the negative value of $\Delta log(Q^{ask})$ suggests that the curve is contracting. Further, the $slope^{bid}$

values at start and end of the third downtrend interval are 0.31 and 0.28 respectively. Q^{bid} at these two points are 670,549 and 763,407. $\Delta log(slope^{bid})$ and $\Delta log(Q^{bid})$ are 0.02 and 0.01. The positive $\Delta log(slope^{bid})$ suggests that the NVWAP-Bid curve is flattening, while the positive $\Delta log(Q^{bid})$ value reflects the expansion of the curve.

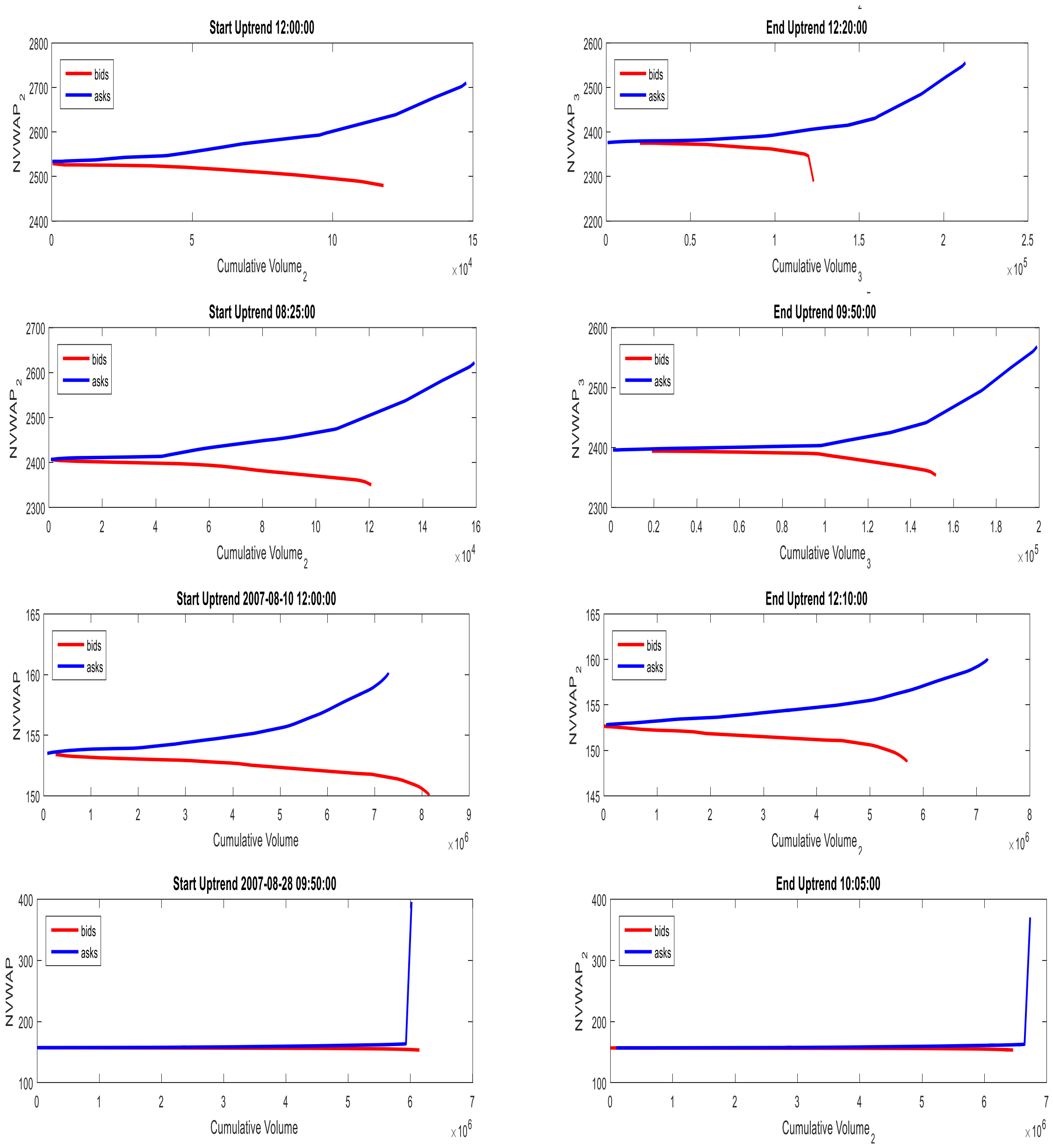


Figure 15: NVWAP curves for the stocks of AZN and Vodafone respectively at the start and end of the uptrend in the month of August 2007. The start and end timestamp for the each snapshot are noted above each curve. The top two panels are for AZN stock on the 10th and the 28th respectively, while the bottom two panels are for Vodafone stock.

On the other hand, Figure 15 and Figure 16 depict a strongly similar patterns of steepening and contraction in NVWAP-Bid accompanied by flattening and expansion in NVWAP-Ask following an uptrend. These effects are observed at all eight downtrend intervals. The value of the $slope^{bid}$ at start and end of the last uptrend interval 09:50 to 10:05 in figure 6 are 0.81 and 0.76 respectively, while the Q^{bid} at these points are 6,144,675 and 6,459,550. The $\Delta log\left(slope^{bid}\right)$ and $\Delta log\left(Q^{bid}\right)$ are reported at -0.09 and 0.01. The negativity in $\Delta log\left(slope^{bid}\right)$ suggests the flattening of the NVWAP-Bid curve and the positive $\Delta log\left(Q^{bid}\right)$ reflects the expansion of the curve. Likewise, the $slope^{ask}$ value at start and end of the last uptrend interval in Figure 15 are 1.10 and 0.96 respectively. As for Q^{ask} values at these points, they are 6,018,278 and 6,733,815. $\Delta log(slope^{ask})$ and $\Delta log(Q^{ask})$ are 0.04 and -0.007 respectively. The previous numbers suggest that the positive $\Delta log(slope^{ask})$ reflects the NVWAP-Ask curve's steepening. Moreover, the negative $\Delta log(Q^{ask})$ suggests that the curve is contracting.

This pattern is also clear in the statistics for the last interval in Figure 16. $Slope^{bid}$ at start and end of the last uptrend interval 08:15 to 08:20 are 0.79 and 0.66 respectively. Q^{bid} at these two points are 6,976,201 and 6,343,944. $\Delta log\left(slope^{bid}\right)$ and $\Delta log\left(Q^{bid}\right)$ are at -0.18 and -0.09. The negativity in $\Delta log\left(slope^{bid}\right)$ suggests the flattening of the NVWAP-Bid curve. However, the negative $\Delta log\left(Q^{bid}\right)$ doesn't really reflect the expansion of the curve as expected. Likewise, the $slope^{ask}$ value at start and end of the last uptrend interval in Figure 16 are 0.75 and 0.64. The Q^{ask} values at these two points are 8,808,778 and 10,288,709. Values of $\Delta log(slope^{ask})$ and $\Delta log(Q^{ask})$ respectively are -0.16 and 0.15. The negative $\Delta log(slope^{ask})$ doesn't reflect the NVWAP-Ask curve's steepening. Also, the positive $\Delta log(Q^{ask})$ suggests that the curve is not contracting.

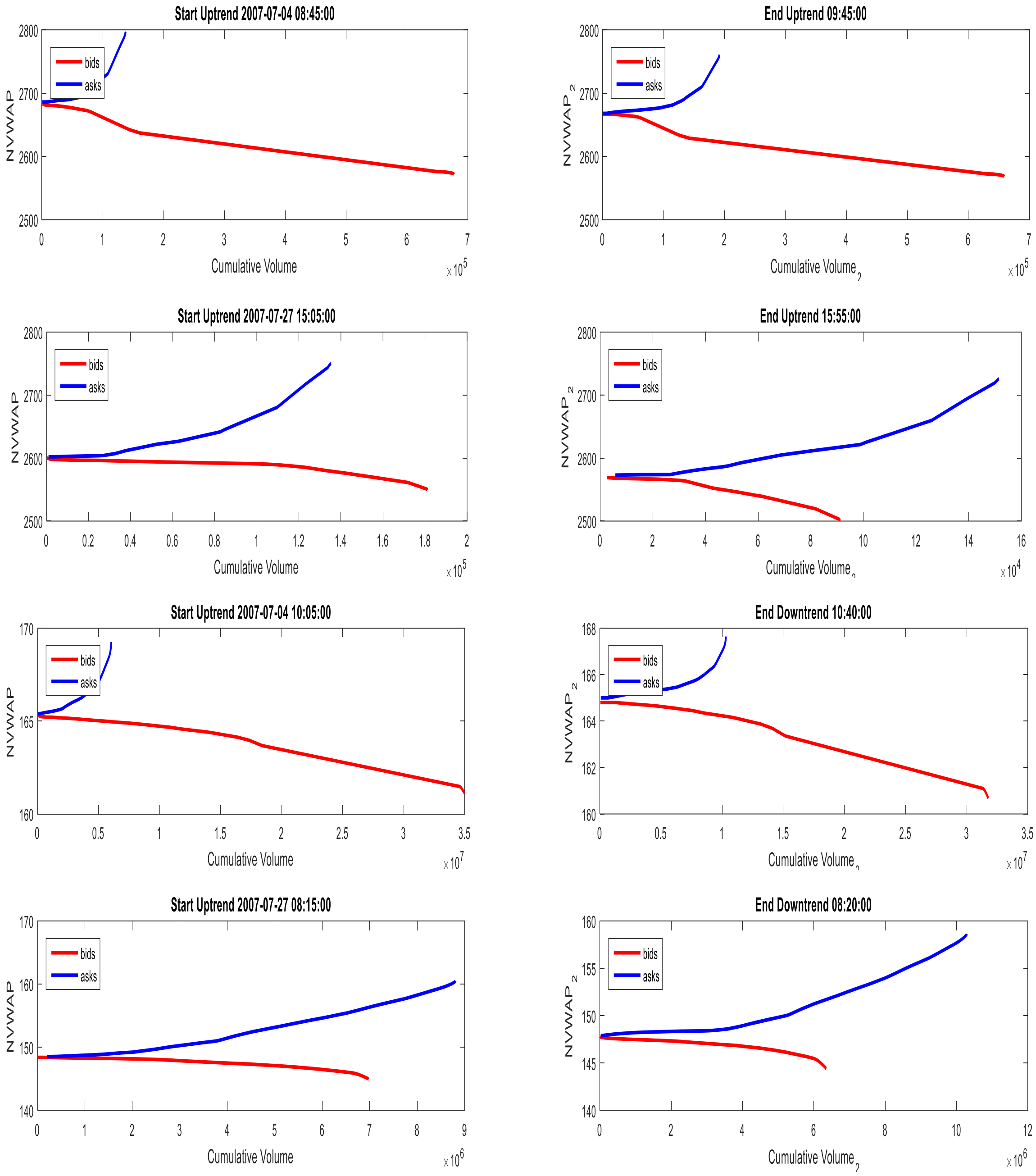


Figure 16: NVWAP curves for the stocks of AZN and Vodafone respectively at the start and end of the uptrend in the month of July 2007. The start and end timestamp for the each snapshot are noted above each curve. The top two panels are for AZN stock, while the bottom two panels are for Vodafone stock.

A sample of uptrend and downtrend intervals for the Vodafone stock in July 2007 is shown in Table 9 below. All uptrend and downtrend intervals for the Vodafone stock in August 2007 and for the AZN stock in July 2007 are shown in Table 16 and Table 17 in the appendix. For uptrend intervals, most $\Delta log(slope^{bid})$ values are negative and most $\Delta log(Q^{bid})$ are positive. Moreover, $\Delta log(slope^{ask})$ values are positive and $\Delta log(Q^{ask})$ are negative. On the other hand, for downtrend intervals, most $\Delta log(slope^{bid})$ values are positive while most

$\Delta log(Q^{bid})$ are negative. Further, $\Delta log(slope^{ask})$ values are negative, however, $\Delta log(Q^{ask})$ are positive.

The analysis cover all uptrend and downtrend intervals which are identified by reversal points at a minimum change in return of ±25bp for all five stocks over a 2 month period, July and August 2007. Overall, the detailed results are consistent with the explanation of the behaviour of the four statistics for three out four uptrend and downtrend intervals analysed above. To sum up, Table 8 summarises the results (comparisons) from the identified intervals of the overall (results) is presented in below. The number of steepening/flattening (S/F) behaviour of the NVWAP curves is counted as the total number of observations where $\Delta log(slope^{bid})$<$\Delta log(slope^{ask})$ in an uptrend, and $\Delta log(slope^{ask})$<$\Delta log(slope^{bid})$ in a downtrend. On the other hand, the contraction and expansion (C/E) behaviour of NVWAP curves is defined as the total number of observations where $\Delta log(Q^{bid})$>$\Delta log(Q^{ask})$ in an uptrend, and $\Delta log(Q^{ask})$>$\Delta log(Q^{bid})$ in a downtrend.

As illustrated in Table 8 below, steepening/flattening (S/F) and contraction/expansion (C/E) of NVWAP curves' behaviour has identified the price trend as expected in 86.61% of the total observations for all five stocks.

Stock	S/F	C/E	Total	S/F (%)	C/E (%)
HSBA	269	274	282	95.39	97.16
BP	361	366	385	93.77	95.06
TSCO	258	263	290	88.97	90.69
AZN	335	338	388	86.34	87.11
VOD	255	264	277	92.06	95.31
Total	**1478**	**1505**	**1722**	**% 85.83**	**% 87.40**

Table 8: Summary of the results of steepening/flattening (S/F) and contraction/expansion (C/E) statistics used to measure the change in the shape of NVWAP curves for all five stocks over a period of 2 consecutive months, July and August 2007. Column S/F is the number of observations where steepening/flattening of NVWAP curves correctly identified the price trend. Column C/E is the number of observations where contraction/expansion of the NVWAP curve correctly identified the price trend. Column Total is the total number of uptrend and downtrend intervals with at least ±25 basis points price return change. Column S/F (%) is the S/F as percentage of total uptrends and downtrends. Column C/E (%) is the C/E as percentage of total observations.

Table 8 above shows that steepening/flattening behaviour has identified the price trend as expected and ranged from 86.34% for AZN to 95.39% for HSBA out of the total observations. In the same way, the contraction/expansion effect has captured the trend and ranged from 87.11% for AZN to 97.16% for HSBA out of the total observations. These results are consistent with the findings of Malik and Markose (2012).

Uptrend – NVWAP – Bid										
Date_time(t)	**Date_time(t+1)**	**Mid_Price$_t$**	**Mid_Price$_{t+1}$**	**Dlog(Mid_Price)**	**Slope_Bid(t)**	**Slope_Bid(t+1)**	**Dlog(Slope_Bid)**	**Q_Bid(t)**	**Q_Bid(t+1)**	**Dlog(Q_bid)**
04/07/2007 10:05	04/07/2007 10:40	165.35	164.9	-0.0027	0.1360	0.15808	0.1500	34996400	31757533	-0.0971
27/07/2007 08:15	27/07/2007 08:20	148.45	147.85	-0.0040	0.7932	0.6593	-0.1848	6976201	6343944	-0.0950
27/07/2007 08:30	27/07/2007 08:35	148.65	148.15	-0.0033	0.7061	0.7720	0.0892	6447547	5835569	-0.0997
Uptrend – NVWAP – Ask										
Date_time(t)	**Date_time(t+1)**	**Mid_Price$_t$**	**Mid_Price$_{t+1}$**	**Dlog(Mid_Price)**	**Slope_Ask(t)**	**Slope_Ask(t+1)**	**Dlog(Slope_Ask)**	**Q_Ask(t)**	**Q_Ask(t+1)**	**Dlog(Q_Ask)**
04/07/2007 10:05	04/07/2007 10:40	165.35	164.9	-0.0027	0.9725	0.4388	-0.7957	6045396	10309776	0.5337
27/07/2007 08:15	27/07/2007 08:20	148.45	147.85	-0.0040	0.7470	0.6375	-0.1585	8808778	10288709	0.1552
27/07/2007 08:30	27/07/2007 08:35	148.65	148.15	-0.0033	0.8192	0.6442	-0.2403	8711915	10907526	0.2247
Downtrend – NVWAP – Bid										
Date_time(t)	**Date_time(t+1)**	**Mid_Price$_t$**	**Mid_Price$_{t+1}$**	**Dlog(Mid_Price)**	**Slope_Bid(t)**	**Slope_Bid(t+1)**	**Dlog(Slope_Bid)**	**Q_Bid(t)**	**Q_Bid(t+1)**	**Dlog(Q_bid)**
04/07/2007 08:30	04/07/2007 10:05	164.05	165.35	0.00789318	0.2110	0.1360	-0.4389	25004913	34996400	0.3361
27/07/2007 08:05	27/07/2007 08:15	147.55	148.45	0.0060811	1.0629	0.7932	-0.2927	5840890	6976201	0.1776
27/07/2007 08:35	27/07/2007 10:00	148.15	153.95	0.0384026	0.7720	0.5389	-0.3594	5835569	8772116	0.4076
Downtrend – NVWAP - Ask										
Date_time(t)	**Date_time(t+1)**	**Mid_Price$_t$**	**Mid_Price$_{t+1}$**	**Dlog(Mid_Price)**	**Slope_Ask(t)**	**Slope_Ask(t+1)**	**Dlog(Slope_Ask)**	**Q_Ask(t)**	**Q_Ask(t+1)**	**Dlog(Q_Ask)**
04/07/2007 08:30	04/07/2007 10:05	164.05	165.35	0.00789318	0.5219	0.9725	0.6222	11000721	6045396	-0.5986
27/07/2007 08:05	27/07/2007 08:15	147.55	148.45	0.0060811	0.6198	0.7470	0.1867	9115481	8808778	-0.0342
27/07/2007 08:35	27/07/2007 10:00	148.15	153.95	0.0384026	0.6442	0.9296	0.3667	10907526	7449644	-0.3812

Table 9: Steepening/ Flattening and Contraction/Expansion statistics for the VOD stock data July 2007.

5.2 The out-of-sample analysis (robustness check)

This analysis covers any arbitrary time interval with a fixed time duration with prediction taking place every 5, 10 and 15 minutes of the first time stamp. Results are reported for all 5 stocks over a 2 month period, July and August 2007. The methodology requires at least two consecutive timestamps for the order book to predict the price trend at the next future timestamp. Consecutive timestamps used are taken following fixed time durations of 5, 10 and 15 minutes time difference.

For an interval of 5 minutes difference between the consecutive timestamps, Table 10 below presents the total number of uptrend and downtrend price movement predicted by the price trend indicator in the month of August 2007. Table 11 depicts the results for the month of July 2007 for the same interval (5 minutes). The NVWAP curves' behaviour has identified the price trend as expected in 61.42% of the total observations for all five stock holdings in the month of August 2007 compared to 69.61% of the total observations in the month of July 2007. The price trend indicator has interestingly identified the price trend as expected, according to Table 10, in only 25.99% for AZN, which is a less liquid stock. Additionally, it ranged between 65.31% for TSCO to 75.27% for VOD out of the total observations. In the same way, Table 11 demonstrates that the price trend was correctly captured in 65.13% for TSCO compared to 73.29% for VOD out of the total observations.

Interestingly, using different time intervals, Table 10, Table 11, Table 12, Table 13, Table 14 and Table 15 suggest that the predictive power of the change in the shape of NVWAP curves for predicting the price trend decreases as markets are more volatile. Furthermore, as the month of August is considered as a volatile month, the predictive power of the change in the shape of NVWAP curves for the least liquid stock (AZN) was a bit low at nearly 26% only compared to nearly 71% in the month of July (before the start of the Financial Crisis). The predictive power of the change in the shape of NVWAP curves for all other stocks was fairly high in the two month. These results suggest the robustness and efficacy of the price trend indicator calculated using the NVWAP methodology.

	Total observations	Actual matches (correct predictions) for 5 min interval		Total actual/correct matches	Percentage of the actual matches from the total observations
Stock		Uptrend	Downtrend		%
HSBA	2220	818	778	1596	71.89
BP	2220	772	752	1524	68.65
TSCO	2220	715	735	1450	65.31
AZN	2220	293	462	755	25.99
VOD	2220	845	826	1671	75.27
Total	**11100**	**2940**	**3878**	**6818**	**61.42**

Table 10: Summary of the results of predictive power of the change in the shape of NVWAP curves for predicting the price trend for 5 minutes interval for all five stocks over the month of August 2007. Column Actual matches is the number of observations where steepening/flattening and contraction/expansion of NVWAP curves correctly identified the price trend for an interval of 5 minutes. This column reports uptrend and downtrend price movements. Moreover, column Total actual matches is the sum of all correctly identified uptrend and downtrend price movements. Lastly, column Percentage of the actual matches represents the percentage of observations where steepening/flattening and contraction/expansion of NVWAP curves correctly identified the price trend.

	Total observations	Actual matches (correct predictions) for 5 min interval		Total actual/correct matches	Percentage of the actual matches from the total observations
Stock		Uptrend	Downtrend		%
HSBA	2220	762	776	1538	69.28
BP	2220	764	777	1541	69.41
TSCO	2220	705	741	1446	65.13
AZN	2220	773	802	1575	70.94
VOD	2220	805	819	1627	73.29
Total	**11100**	**3809**	**3918**	**7727**	**69.61**

Table 11: Summary of the results of predictive power of the change in the shape of NVWAP curves for predicting the price trend for 5 minutes interval for all five stocks over the month of July 2007. Column Actual matches is the number of observations where steepening/flattening and contraction/expansion of NVWAP curves correctly identified the price trend for an interval of 5 minutes. This column reports uptrend and downtrend price movements. Moreover, column Total actual matches is the sum of all correctly identified uptrend and downtrend price movements. Lastly, column Percentage of the actual matches represents the percentage of observations where steepening/flattening and contraction/expansion of NVWAP curves correctly identified the price trend.

Stock	Total observations	Actual matches (correct predictions) for 10 min interval		Total actual/correct matches	Percentage of the actual matches from the total observations
		Uptrend	Downtrend		%
HSBA	1110	410	385	795	71.16
BP	1110	376	397	773	69.64
TSCO	1110	350	366	716	64.50
AZN	1110	143	154	297	26.77
VOD	1110	407	418	825	74.32
Total	**5550**	**1686**	**1720**	**3406**	**61.28**

Table 12: Summary of the results of predictive power of the change in the shape of NVWAP curves for predicting the price trend for 10 minutes interval for all five stocks over the month of August 2007. Column Actual matches is the number of observations where steepening/flattening and contraction/expansion of NVWAP curves correctly identified the price trend for an interval of 10 minutes. This column reports uptrend and downtrend price movements. Moreover, column Total actual matches is the sum of all correctly identified uptrend and downtrend price movements. Lastly, column Percentage of the actual matches represents the percentage of observations where steepening/flattening and contraction/expansion of NVWAP curves correctly identified the price trend.

Stock	Total observations	Actual matches (correct predictions) for 10 min interval		Total actual/correct matches	Percentage of the actual matches from the total observations
		Uptrend	Downtrend		%
HSBA	1110	415	372	787	70.90
BP	1110	389	385	774	69.73
TSCO	1110	355	372	727	65.49
AZN	1110	382	409	791	71.26
VOD	1110	398	415	813	73.24
Total	**5550**	**1939**	**1953**	**3892**	**70.12**

Table 13: Summary of the results of predictive power of the change in the shape of NVWAP curves for predicting the price trend for 10 minutes interval for all five stocks over the month of July 2007. Column Actual matches is the number of observations where steepening/flattening and contraction/expansion of NVWAP curves correctly identified the price trend for an interval of 10 minutes. This column reports uptrend and downtrend price movements. Moreover, column Total actual matches is the sum of all correctly identified uptrend and downtrend price movements. Lastly, column Percentage of the actual matches represents the percentage of observations where steepening/flattening and contraction/expansion of NVWAP curves correctly identified the price trend.

Stock	Total observations	Actual matches (correct predictions) for 15 min interval		Total actual/correct matches	Percentage of the actual matches from the total observations
		Uptrend	Downtrend		%
HSBA	740	256	278	534	72.16
BP	740	269	247	516	69.73
TSCO	740	236	244	480	64.86
AZN	740	96	108	204	27.57
VOD	740	294	265	559	75.54
Total	**3700**	**1151**	**1142**	**2293**	**61.97**

Table 14: Summary of the results of predictive power of the change in the shape of NVWAP curves for predicting the price trend for 15 minutes interval for all five stocks over the month of August 2007. Column Actual matches is the number of observations where steepening/flattening and contraction/expansion of NVWAP curves correctly identified the price trend for an interval of 15 minutes. This column reports uptrend and downtrend price movements. Moreover, column Total actual matches is the sum of all correctly identified uptrend and downtrend price movements. Lastly, column Percentage of the actual matches represents the percentage of observations where steepening/flattening and contraction/expansion of NVWAP curves correctly identified the price trend.

Stock	Total observations	Actual matches (correct predictions) for 5 min interval		Total actual/correct matches	Percentage of the actual matches from the total observations
		Uptrend	Downtrend		%
HSBA	740	261	260	521	70.40
BP	740	254	268	522	70.54
TSCO	740	237	261	498	67.30
AZN	740	247	277	524	70.81
VOD	740	266	279	545	73.65
Total	**3700**	**1265**	**1345**	**2610**	**70.54**

Table 15: Summary of the results of predictive power of the change in the shape of NVWAP curves for predicting the price trend for 15 minutes interval for all five stocks over the month of July 2007. Column Actual matches is the number of observations where steepening/flattening and contraction/expansion of NVWAP curves correctly identified the price trend for an interval of 15 minutes. This column reports uptrend and downtrend price movements. Moreover, column Total actual matches is the sum of all correctly identified uptrend and downtrend price movements. Lastly, column Percentage of the actual matches represents the percentage of observations where steepening/flattening and contraction/expansion of NVWAP curves correctly identified the price trend.

Generally, the detailed results are consistent with the explanation of the behaviour of the four statistics which form the price trend indicator. On average, 65.82% of the price changes for all stocks in the 2 months period were correctly predicted by the proposed indicator. This percentage is fairly high, which suggests the accuracy and the precision of this indicator as a tool to predict price movements in the stock market. This tool can be effectively utilised to form profitable trading strategies.

6 Conclusion

This study aims at analysing the intraday dynamics of liquidity in London SETS electronic limit order book for five different stock holdings over a period of two months, July and August 2007, due to data limitation. This study applies the Malik and Markose (2012) definitions of NVWAP to model the shape of bid and ask side of the order book to predict the market trend under different time intervals. The prediction methodology consists of four statistics which are $\Delta log(slope^{ask})$, $\Delta log\left(slope^{bid}\right)$, $\Delta log\left(Q^{bid}\right)$ and $\Delta log(Q^{ask})$.

The empirical analysis confirms the that information from the NVWAP along with the resulted liquidity supply and demand curves in the limit order book reveal consistent observable market behaviour during both uptrend and downtrend intervals. In fact, examining the change in the shape of NVWAP curves for five different stocks over a period of two consecutive months, July and August 2007, suggests that the aforementioned four statistics have correctly identified prevailing market trend in 86.61% of the total observation on average. Thus, these four statistics indicator is revealed to be robust measure to identify the prevailing market trend without prior knowledge of the price. These results go in line with Malik and Markose (2012) findings. Within an algorithmic trading system, as market state changes, these four statistics can be used to switch between different trading strategies.

This study can be extended by investigating the change in order submission sides along with cancellations in both sides in the limit order book. Understanding the order cancellation on each side of the order book will help to reveal more information about price changes in the market.

7 Appendix

Date_time	Slope_Bid(t)	Slope_Bid(t+1)	Dlog(Slope_Bid)	Q_Bid(t)	Q_Bid(t+1)	Dlog(Q_Bid)	Slope_Ask(t)	Slope_Ask(t+1)	Dlog(Slope_Ask)	Q_Ask(t)	Q_Ask(t+1)	Dlog(Q_Ask)
10/08/2007 08:05	1.37	1.44	0.05	3254123	3770995	0.15	0.74	0.91	0.20	6170708	5439794	-0.13
10/08/2007 08:20	1.39	1.41	0.02	4065445	3792570	-0.07	1.00	1.01	0.01	5688603	6114593	0.07
10/08/2007 08:45	2.51	1.92	-0.27	2947751	3567969	0.19	0.63	0.71	0.11	8385237	7566272	-0.10
10/08/2007 09:20	1.21	1.15	-0.05	4893506	4592816	-0.06	1.05	1.57	0.41	5326098	4724770	-0.12
10/08/2007 09:45	1.74	1.40	-0.21	3661629	4228186	0.14	0.79	0.82	0.04	6677490	6160040	-0.08
10/08/2007 10:10	0.87	0.82	-0.05	5178074	5266898	0.02	1.06	0.95	-0.12	5947307	5914503	-0.01
10/08/2007 11:35	1.79	1.37	-0.27	2859632	3877915	0.30	0.53	0.75	0.34	7680536	6989393	-0.09
10/08/2007 11:45	1.14	1.63	0.35	4489468	3316229	-0.30	0.83	0.62	-0.29	7033693	9157207	0.26
10/08/2007 11:50	1.63	0.77	-0.75	3316229	5812013	0.56	0.62	0.62	-0.01	9157207	8149647	-0.12
10/08/2007 12:00	0.60	0.74	0.20	8153153	6693109	-0.20	0.73	0.66	-0.11	7289708	8010201	0.09
10/08/2007 12:10	0.78	0.65	-0.17	5694197	6143574	0.08	0.73	0.64	-0.12	7205918	7409633	0.03
10/08/2007 12:35	0.60	0.67	0.11	7015777	6297316	-0.11	0.82	0.62	-0.29	7182225	7622034	0.06
10/08/2007 13:00	0.81	0.82	0.01	6091717	5835121	-0.04	0.72	0.64	-0.12	7439944	7825786	0.05
10/08/2007 14:00	0.59	0.62	0.05	7111153	6135572	-0.15	1.01	0.63	-0.47	6827081	7873537	0.14
10/08/2007 14:05	0.62	0.64	0.03	6135572	6360365	0.04	0.63	0.78	0.21	7873537	7220321	-0.09
10/08/2007 14:25	0.55	0.71	0.25	7509365	6274604	-0.18	1.30	0.82	-0.46	4914590	6790133	0.32
10/08/2007 14:40	0.69	0.69	0.01	6014669	6220684	0.03	0.94	0.92	-0.03	6850647	5698444	-0.18
10/08/2007 15:05	0.66	0.63	-0.04	7470163	6250194	-0.18	1.17	0.68	-0.55	4940238	7854195	0.46
10/08/2007 15:10	0.63	0.58	-0.09	6250194	7225677	0.15	0.68	0.94	0.33	7854195	6180254	-0.24
10/08/2007 15:15	0.58	0.66	0.14	7225677	6015799	-0.18	0.94	0.82	-0.14	6180254	7052619	0.13
10/08/2007 15:25	0.59	0.64	0.08	5924151	6808219	0.14	0.60	0.77	0.25	7964908	6502650	-0.20
10/08/2007 15:30	0.64	0.60	-0.08	6808219	6018456	-0.12	0.77	0.60	-0.24	6502650	7623769	0.16
10/08/2007 15:55	0.79	0.64	-0.22	6045821	6264747	0.04	0.76	0.66	-0.13	7876556	8459956	0.07
10/08/2007 16:15	0.79	0.65	-0.19	6675774	5700891	-0.16	0.92	0.77	-0.17	7722146	6541244	-0.17
10/08/2007 16:25	0.78	1.03	0.28	6605043	3785076	-0.56	0.76	0.89	0.16	7364437	5661905	-0.26

Steepening/ Flattening and Contraction/Expansion statistics for the Vodafone stock data August 2007 (Cont.)												
Date_time	Slope_Bid(t)	Slope_Bid(t+1)	Dlog(Slope_Bid)	Q_Bid(t)	Q_Bid(t+1)	Dlog(Q_Bid)	Slope_Ask(t)	Slope_Ask(t+1)	Dlog(Slope_Ask)	Q_Ask(t)	Q_Ask(t+1)	Dlog(Q_Ask)
28/08/2007 08:25	1.04	1.05	0.01	5381253	5422252	0.01	0.87	0.99	0.13	6465820	6346551	-0.02
28/08/2007 09:50	0.82	0.74	-0.09	6144675	6212672	0.01	1.10	1.15	0.04	6018278	5974961	-0.01
28/08/2007 10:05	0.76	0.75	-0.01	6459550	6865693	0.06	0.96	1.03	0.07	6733815	6208278	-0.08
28/08/2007 11:15	0.70	0.71	0.01	6974814	6440052	-0.08	1.16	1.07	-0.08	5961906	6226548	0.04
28/08/2007 12:00	0.86	0.68	-0.23	6411618	6192453	-0.03	0.85	0.95	0.12	7121114	7058960	-0.01
28/08/2007 13:00	0.73	0.79	0.08	7364317	7080544	-0.04	0.98	0.96	-0.02	6308823	6544674	0.04

Table 16: Steepening/ Flattening and Contraction/Expansion statistics for the Vodafone stock data August 2007.

Date_time	Slope_Bid(t)	Slope_Bid(t+1)	Dlog(Slope_Bid)	Q_Bid(t)	Q_Bid(t+1)	Dlog(Q_Bid)	Slope_Ask(t)	Slope_Ask(t+1)	Dlog(Slope_Ask)	Q_Ask(t)	Q_Ask(t+1)	Dlog(Q_Ask)
04/07/2007 08:45	0.26	0.29	0.10	677295	674167	-0.005	1.56	1.68	0.07	137745	143007	0.04
04/07/2007 09:45	0.34	0.33	-0.03	658829	662526	0.01	1.24	1.36	0.10	191912	187641	-0.02
04/07/2007 12:25	0.28	0.29	0.05	690575	687812	0.00	1.93	1.88	-0.03	138321	158272	0.13
04/07/2007 13:35	0.31	0.31	0.02	670549	680380	0.01	1.42	1.53	0.07	173286	164789	-0.05
04/07/2007 15:25	0.28	0.29	0.02	763407	758132	-0.01	1.95	2.12	0.08	122631	124879	0.02
27/07/2007 08:05	1.68	1.31	-0.25	121048	128240	0.06	1.78	1.28	-0.33	146225	145142	-0.01
27/07/2007 09:45	1.51	1.65	0.09	130288	120540	-0.08	2.39	2.21	-0.08	103416	122356	0.17
27/07/2007 10:20	1.95	1.73	-0.12	99621	116319	0.15	1.48	2.00	0.30	142291	143460	0.01
27/07/2007 10:25	1.73	2.09	0.19	116319	108343	-0.07	2.00	1.48	-0.30	143460	181082	0.23
27/07/2007 11:20	3.06	1.96	-0.44	82663	102601	0.22	1.76	1.81	0.03	165404	140905	-0.16
27/07/2007 11:50	1.56	1.77	0.12	113410	107362	-0.05	1.87	1.00	-0.63	129737	225106	0.55
27/07/2007 12:35	2.84	2.32	-0.20	84421	89650	0.06	1.41	1.69	0.18	192301	163164	-0.16
27/07/2007 14:20	1.84	1.85	0.01	131980	124222	-0.06	1.96	1.46	-0.30	147147	162250	0.10
27/07/2007 14:40	1.46	0.80	-0.61	136847	186906	0.31	1.12	1.65	0.39	182374	168048	-0.08
27/07/2007 15:05	0.82	1.79	0.78	181153	122065	-0.39	1.93	1.59	-0.19	135252	168555	0.22
27/07/2007 15:55	2.10	2.07	-0.01	92969	91051	-0.02	1.52	1.42	-0.07	151457	151394	-0.0004
27/07/2007 16:10	1.31	1.26	-0.03	127544	123763	-0.03	1.55	1.48	-0.05	148259	157227	0.06

Table 17: Steepening/ Flattening and Contraction/Expansion statistics for the AZN stock data July 2007.

4 Chapter Four

The Effect of Cancellations in The Electronic Order Book

Study on London Stock Exchange Electronic Order Book (SETS)

Abstract

The electronic order book (E-LOB) has no designated market maker for the liquidity provision. Traders, therefore, in E-LOB markets are known to provide fleeting liquidity with sudden withdrawals of positions. Order submissions and order cancellations can be driven by price discovery strategies as well as those that can influence the shape of the order book to bring about price trends that increase the profitability of traders' positions. Some of the latter can tip over to price manipulation, and these are known to include strategies like spoofing and layering, which have been alleged to cause flash and mini-crashes. This study tries to uncover some of the effects of the observable multitude of order cancellations using a sample of stocks from London Stock Exchange Electronic Trading System (SETS). Thus, it inspects whether cancellations contribute to the size of price drops to the extent that this phenomenon may lead to market volatility. In light of the Directional Change (DC) methodology, the sampling of the order book is done on the basis of a Peak-Trough methodology to avoid any potential loss of information that results from aggregating data containing all order events occurring within an interval. Further, a new measure of bid-ask spread is proposed based on the slopes of the Notional Volume Weighted Average Price (NVWAP) curves. All analysis is based on ultra-high-frequency data from the London SETS order book which investigates the periods where the market was experiencing a downtrend price move (peak to trough). The price impact of type of event, namely new buy, new sell, cancellation of buy or cancellation of sell, may not show any statistical significance outside of the Peak-Trough and Trough-Peak analysis. This study suggests that buy order submissions decrease and their cancellations increase during a downtrend price move. Their effects are highly significant determinants of the change in cumulative return when price is decreasing under both extreme and normal price conditions. Sell new orders are significant determinant of the change in cumulative return though. On the other hand, sell order submissions' increase and their cancellations' decrease are significant determinants of the change in cumulative return under normal price conditions. However, the effect of the changes on both sides of the market has no significant impact on the proposed bid-ask spread measure under extreme and normal price conditions. These findings go in line with the expected sign or direction of change in cumulative return. Lastly, cancellations, generally, increase on both sides of the market as the trading day approaches its end regardless of the volatility.

Keywords: Order Book Rebuild; High Frequency Trading; Volatility; Liquidity Dynamics; Price Trend; Order Book Cancellations; Bid-Ask Spread

1 Introduction

The recent Flash Crash and the subsequent investigations have focussed on the relation between order flow and price changes (Eisler et al., 2012). Orders submitted to the market are characterised as one of the following: firstly, fully or partially matched resulting in a trade, secondly, remain in the order book till they are matched (fully or partially) or expired or, thirdly, cancelled (deleted). Kirilenko et al. (2011) highlight that cancelation of orders was a large contributor to the Flash Crash. The fleeting nature of orders has raised many concerns, and therefore led to more research on the order submission process. The US Department of Justice report[16] (2015, p. 22), in the case "United States of America v. Navinder Singh Sarao", alleged that "Sarao's use of the dynamic layering technique was particularly intense in the hours leading up to the Flash Crash… Sarao began this cycle by placing … five sell orders nearly simultaneously… The orders were replaced or modified more than 19,000 times before Sarao cancelled them, without having executed any of them." This indicates that order submissions and cancellations can be used to manipulate the market.

Traditionally, Baruch and Glosten (2013) point out that limit order book cancellations have been treated as part of the order book imbalance measures. Moreover, the ability to revise a decision is a fundamental mechanism in the modern price discovery process. However, the fact that (see Table 22) almost 90% of all limit orders are cancelled before execution, most of them within a couple of seconds, sheds the light on the role which cancellations play in modern financial markets (see Hasbrouck and Saar, 2009; Huang and Hautsch, 2011; Scholtus and van Dijk, 2015). Therefore, regulators such as SEBI (Securities and Exchange Board of India[17])

---

[16] Available online at: https://www.justice.gov/sites/default/files/opa/press-releases/attachments/2015/04/21/sarao_criminal_complaint.pdf.

[17] That is in a "Discussion paper on 'Strengthening of the Regulatory framework for Algorithmic Trading & Co-location"; available online on: https://www.sebi.gov.in/sebi_data/attachdocs/1470393485587.pdf.

has proposed the imposition of minimum resting time of 500 milliseconds (ms) for orders, whereby an order will not be allowed to be amended or cancelled.

Interestingly, Farmer et al. (2004) and Finer et al. (2016) have argued that order cancellation has become as important as submission itself, mainly with regard to trading strategies, liquidity shortfalls and their effects on large price movements. Morck et al. (2000) make the point that trades affect the price in the direction of their transactions, whereby price fluctuates with excess demand and excess supply. Additionally, Paddrik et al. (2012) have made the point that according to data provided by the Commodity and Futures Trade Commission, high frequency traders (HFTs) place 60% of their orders 1 tick away or less from the last trade price. The analysis of this chapter shows that an average of 77% of the new submissions where placed 1 tick away or less from the last trade price in July 2007 for instance.

Treleaven et al. (2013) point out that, in the finance industry, the focus is on improving the profitability of trading algorithms as well as improving their ability to identify patterns in the market data besides, minimising transaction costs and limiting the amount and type of risk. Within an algorithmic trading system, as market state changes, the change in the shape of bid and ask sides of the order book can be used to switch between different trading strategies. Understanding the order cancellation on each side of the order book will help to reveal more information about price changes in the market. Moreover, at the time of the May 6th, 2010 Flash Crash, Paddrik et al. (2012) describe that algorithmic trading in the U.S. was thought to be accountable for more than 70% of trading volume.

Studying high frequency trading (HFT) is, therefore, crucial to comprehend modern market behaviour and order cancellations. The definition of HFT is not really coherent in the literature. It ranges from a very general, covering almost 95% of trading activities, to extremely specific definition, including only a small fraction of market actors. HFTs' strategies are characterised by holding usually extremely short term positions and making money by a high volume of many tiny returns rather than on the basis of high returns with single trades. To achieve these small profits, they need to rapidly submit orders to the order book and then cancel (delete) them (Aldridge, 2013; Baruch and Glosten, 2013; Gomber et al., 2011; Lhabitant and Gregoriou, 2015; Security and Exchange Commission, 2010).

Montgomery (2016) argue that HFTs act to move quotes by submitting and quickly cancelling limit orders to rapidly profit from trading the induced transitory mispricing. Interestingly, at the best price, a cancellation of the last limit order widens the spread by either increasing the best ask or decreasing the best bid. He highlights that there are different strategies/techniques[18] to manipulate the market price. One strategy is called Fictitious Orders, which is entering orders without an intention to execute them. Another strategy is spoofing which is very similar to the former strategy while the third is called Layering, which is the most complex one. Nevertheless, not all orders' submissions and cancellations are aimed at manipulating the market. Aquilina et al (2017) make the point that order flow is informative in itself. In light of cancellations, they emphasise that traders may submit additional buy and sell orders to the order book while trying to discover the efficient price in a gradual learning process.

Furthermore, Lhabitant and Gregorius (2015) underline that to understand the reasons behind the high order cancellation rate, one has to realise the behaviour intentions of market participants and the circumstances behind them. This behaviour is noted as trading strategy, however, not everything is covered by specific trading strategies. Generally speaking, there are two driving forces which drive cancellations. These factors differ between making money and preventing losses. Therefore, all reasons for order cancellation are classified into active execution for offence reasons (making money) or active execution for defence reasons

[18] These strategies are described in the SEC indictment for market manipulation by Sarao in the run inspired by the 2010 Flash Crash. See https://www.sec.gov/news/studies/2010/marketevents-report.pdf

(preventing losses). For a limit order book market, Obizhaeva and Wang (2005) have pointed out that when trading times are chosen optimally, the dynamics of the supply/demand play a key role in determining the optimal execution strategy.

The main innovation to the literature is to use the methodology from the previous chapter where it was found that that order book had characteristic changes in its shape depending on periods of price rises and price falls. Periods when cumulative returns exceed a certain threshold (in basis points) in the up- or downward directions, respectively, have been called directional changes (DC), i.e. trough to peak and peak to trough (see Tsang, 2017 for more about DC). Peak-Trough methodology has been introduced and applied to avoid any potential loss of information results when aggregating data containing all order events occurring within an interval. Further, to study the role of order submissions and cancellations in the order book under peak to trough and trough to peak, I use the tool developed in the previous chapter to indicate the prevailing market trends. This tool uses the information conveyed by the Notional Volume Weighted Average Price (NVWAP) curves. NVWAP curves provide a method to model the shape of the entire order book on the basis of the average price at cumulative volumes at given time stamps. These curves, in turn, depict the expectations of market participants regarding current and future price movements.

Put differently, the information conveyed by the NVWAP order book curves (in the bid and ask sides) suggests that the market activity on each side of the limit order book directly affects the shape of liquidity supply and demand curves. Consequently, the change in the slope of the NVWAP curve ($slope^{bid}$ and $slope^{ask}$) captures both the volume available and any additional cost (benefit) results from buying (selling) in the rising (falling) market. These changes in slopes of NVWAP curves cause steepening/flattening effects. On the other hand, the contraction/expansion effects of the NVWAP curves are measured by calculating the changes in the total volume available on both sides of the limit order book (Q^{bid} and Q^{ask}).

In a downtrend, change in slope of the bid curve is greater than the change in the slope of the ask curve, i.e. $\Delta log\left(slope^{bid}\right) > \Delta log(slope^{ask})$. The slope coefficient (β_1) is estimated at each peak and trough accounting for both uptrends (from trough to peak) and downtrends (from peak to trough) intervals. Each of these uptrend and downtrend intervals has a return of at least ±25bp. Then, I calculate the following four statistics $\Delta log\left(slope^{bid}\right)$, $\Delta log(slope^{ask})$, $\Delta log\left(Q^{bid}\right)$and $\Delta log(Q^{ask})$ at the start and the end of each uptrend and downtrend interval. Hence, the change in the slope of the NVWAP curves, and the change in the total volume on each side of the limit order book can be used within an algorithmic trading strategy to exploit trading opportunities. It is expected that when the market price falls (downtrend), the bid side gradually decreases. The bid side NVWAP curve slowly contracts and becomes steeper. The NVWAP curve on ask (liquidity supply) side curve then slowly expands and becomes flatter.

The present study investigates the cancellations that occurred on both sides of the limit order book, buy and sell, using a sample of stocks from London Stock Exchange examining the aforementioned change (NVWAP slopes). The categories of order events investigated, under conditions of price falls, include new sells, new buys, cancellation buys and cancellation sells. So, it is hypothesised that for large price falls, new buys decrease while new sells increase; also, cancellations of buy increase but cancellations of sells decrease for large price falls. These categories of order submissions and cancellations are studied in the context of the peak and trough analysis applied in the previous chapter. These lead to precise hypotheses where the differential behaviours of cancellations are investigated when price is falling (downtrend). To establish the stylised facts, the peak to trough data is differentiated into: (1) small change when the change in cumulative return is more than 25 basis points but less than 50, (2) medium change in cumulative return when the change is more than 50 basis points but less than 100 and (3) large change in cumulative return when it changes by more than 100 basis points.

I also introduce a new measure of volatility based on volume weighted concept of bid-ask spread[19]. This is estimated as the change in the logarithm of slope of the NVWAP curves (bid and ask) from peak to trough. Furthermore, this study extends Mayston et al. (2008) work. One extension is by dividing the trading day into four periods instead of a time dummy which accounts only for the end of the trading day. The new time dummy[20] will account for the time of the day in which the order has been submitted to the order book or cancelled to identify whether cancellations trend varies with trading hours. Another extension is redefining the resiliency variable to be the duration involved in cumulative return from peak to trough. Flash Crashes, for instance, involve large price drops in very short time periods, i.e. duration. This variable (duration) is used to normalise/adjust the volumes of cancellations and submissions on both sides of the market from peak to trough.

Thus, this study follows the literature by investigating the variables which contribute to the explanation of the market impact and volatility. These variables are, namely, volume, change in return, change in price and time/duration. By rebuilding the E-Order Book for 5 stocks from London SETS in July and August 2007, this study reconstructs the demand and supply curves to analyse the price impact and fluctuations with special focus on cancellations. More specifically, this paper's aim is to investigate how cancellations on both sides of the order book contribute to the market impact measured as the duration adjusted cumulative return of peak-trough directional changes. It, also, aims at investigating the cancellations behaviour on both

---

[19] The bid-spread at the best is very competitive at 0.5 basis points on average.

[20] This variable clarifies the times where more orders have been submitted to the order book by traders from outside the UK, namely from the US or any other country. As these traders join the trading session, they may affect the number of submissions of new orders as well as orders cancellations.

sides of the order book and their effect on market volatility based on the changes of the NVWAP slopes between bid and ask sides.

To recap, this study investigates the effect of cancellations on the market to the extent that this phenomenon may lead to market crashes. It is expected that during a downtrend price move, cancellations on the buy side increase while cancellations on the sell side decrease. The cancellations behaviour affect both the size of the price drop as well as volatility by broadening them which might cause a flash crash to the market. Particularly, this study inspects whether:

1. Cancellations contribute to the size of price drops (have large market impact).
2. Cancellations increase the market volatility.

To this end, different model specifications are defined and tested comparing scenarios where variables are/aren't adjusted to duration of the peak to trough price trend. Duration adjusted models are expected to perform better than non-adjusted ones in explaining the cancellations effect on the market. All analysis is based on ultra-high-frequency data from the order book which investigates the periods where the market was experiencing downtrend price moves (from peak to trough). Robust regressions[21], which control for efficiency under more realistic conditions, are run to check the validity of the regression results obtained.

The rest of the chapter is structured as follows. Section 2 presents a background and a literature survey. The methodology and the research hypotheses are discussed in section 3 while a description of the data used is given in section 4. Section 5 presents the data analysis and the empirical results of the study along with discussing the implications of the study. The conclusion and some suggestions for future research are given in the last section.

[21] See Andrews (1974).

2 Literature Review

2.1 Background on Price Impact, Order Submissions and Cancellations

In recent years, the wide evolution in electronic trading in both Europe and the U.S. has led to introduce new trading platforms (e.g. Chi-X and EdgeX) in stock exchanges (Colliard and Foucault, 2012). These platforms allow traders to utilize technology to swiftly buy and sell their shares in the market. Briefly, in line with the notion of fast automated execution and high technology, market participants complete their transactions fully electronically with identified counterparties who want to transact/trade in a matter of seconds. This ultimately sets the tone for emerging a new style of trading known as high frequency trading, where automated systems electronically buy and sell. HFT strategies update their orders very quickly (a few seconds or less) and have no over-night positions. To realize small profits per trade, rapid submission of cancellations and deletions is necessary (Gomber et al., 2011).

It is traditionally argued that there are, in fact, two types of traders coexisting in the financial markets: (i) "informed" traders and (ii) uninformed (or less informed) market makers. "Informed" traders place market orders for immediate execution where their cost is at half the bid-ask spread level. On the other hand, uninformed[22] (or less informed) market makers provide liquidity in a shape of placing limit orders on both sides of the order book in a hope of earning part of the bid-ask spread. In the current electronic markets, there exists a conceptual problem such that it is difficult to distinguish between an informed trader and a market maker. This is due to the fact that each participant in the market can place both limit and market orders depending on the current state of the order book or their own strategies (Eisler et al., 2012; Kyle, 1985).

[22] An asymmetry in information between a buyer and the corresponding seller is present.

Electronic trading platforms and automated trading have led to less execution costs[23] as well as higher execution efficiency[24]. Nevertheless, this automated trading follows different strategies compared to non-automated trading. The trading strategy decisions have to be made much faster as the speed of trading has significantly increased. A recent and growing body of literature suggests that not only trades are informative but also other order book events. The existing literature focuses more on the decision of whether to place a market order or a limit order. Yet, decisions after an investor places a limit order, i.e. orders cancellations, are not widely considered by this literature. Put differently, there is growing literature tends to classify all activities of electronic traders as being strategic (see Fong and Liu, 2008; Lhabitant and Gregorius, 2015; Weber and Rosenow, 2005).

On the other hand, Höschler (2011) suggests that the main risk of a limit order arises from the execution time or time-to-fill. Its execution price is fixed by the price tick when the limit order is placed though. Thus the execution price (traded VWAP) depends on its size and the liquidity available in the order book. Traders may choose to place limit orders to decrease execution costs. Limit orders also have an impact on price in that submitting a buy limit order sets an extra upwards pressure on price, while cancelling a buy limit order decreases this pressure. Furthermore, in the context of optimal order execution, understanding the price impact of orders from a theoretical perspective is important. Huberman and Stanzl (2004) suggested that arbitrage opportunities exist if the effect of trades on prices is permanent and that effect is non-linear. Gatheral (2010) also studied the effect of trades on prices and added that if the price impact function is non-linear, arbitrage could be excluded if the trades impact decays in a particular way.

---

[23] That is due to lower transaction costs and less need for human intervention.
[24] That is due to quicker reaction to incoming orders and market changes.

Bouchaud et al. (2004) carried out a detailed study of the statistics of price changes at the trade by trade level, and analysed the interplay between the impact of each trade on price and volatility. They associated limit orders with the decay of price impact of trades. They argued that there is a "delicate interplay" between two opposite tendencies: first, strongly correlated market orders which lead to super-diffusion (or persistence), and second, mean reverting limit orders which lead to sub-diffusion (or anti-persistence). Their insight implies that focusing solely on trades without considering the effect of limit orders volumes leads to ignoring an important part of the price formation mechanism.

Furthermore, the main focus of the empirical literature on price impact has been primarily on trades. Different models for price impact have been recommended in the literature; however, little agreement on how to model that impact is in place. Price impact has been described by various authors in the empirical literature as non-linear, linear, square root, mechanical, virtual, permanent, instantaneous, temporary or transient (see Cont et al., 2014; Bouchaud, 2010). The intuitive notion that prices move due to the imbalance between demand and supply seems to be the only consensus.

Eisler et al. (2012) review the dynamics of prices and state and suggest three processes. These processes are "instantaneous jumps due to events, events inducing further events and thereby affecting the future jump probabilities (described by the correlation between events), and events exerting pressure on the gaps behind the best price and thereby affecting the future jump sizes". Moreover, large orders have an impact on the subsequent behaviour of the order book; market (buy) orders drive up prices and widen the spread (i.e. the gap between best limit buy and sell order). Large orders, which are visible to all market participants, change the supply and demand balance of the asset and therefore also have an impact on the price, even without being executed (Höschler, 2011).

Traditionally, cancellations have been treated as part of order book imbalance measures. Cancellations are part of traders' trading strategies. For a limit order book market, Obizhaeva and Wang (2005) have constructed a simple dynamic model to capture the intertemporal nature of supply/demand in the market and solved for the optimal execution strategy. They have pointed out that when trading times are chosen optimally, the dynamics of the supply/demand play a key role in determining the optimal execution strategy.

Additionally, Machain and Dufour (2013) have questioned the price impact of limit order cancellations. They have explicitly modelled cancellations as a separate event using a vector autoregressive (VAR) specification. They have found limit order cancellations are negatively correlated with the cumulative quote changes. Typically, a cancellation on the buy (sell) side permanently decreases (increases) stock prices. Nevertheless, Machain and Dufour (2013) have suggested that HFTs are feasibly acting to move quotes by submitting and quickly cancelling limit orders to rapidly profit from trading the induced transitory mispricing. They have, also, made the point that non-HFTs' limit order cancellations convey some information while HFTs' cancellation activity is not informative.

Interestingly, at the best price, a cancellation of the last limit order widens the spread by either increasing the best ask or decreasing the best bid. Eisler et al. (2012) suggested that the impact of market orders was the main focus of previous studies which found that this impact decays in a way that offsets the correlation of the sign of the trades. They described the underlying mechanism as that, on one side of the book, market orders attract compensating limit orders. They added that "these limit orders do not necessarily change the best limits, but are such that the conditional impact of a buy trade following other buy trades is smaller than the conditional impact of a sell trade following buy trades." In addition, Farmer et al. (2004) studied changes in the mid-price (returns) at the level of individual events. They concluded that the returns triggered by market orders, limit orders, and cancellations are very similar and are statistically indistinguishable from those caused by market orders. They emphasised that their results are quite consistent across the cancellation of buy and sell limit orders. They added that any removal of volume (cancellation) at the best price is equivalent to removal by a market order. Both cases create a price change equivalent to the first gap. However, the effect[25] on best prices for a limit order remains unexplained.

Eisler et al. (2012) suggested that there is a long-range relationship between events happening on the same side of the book. However, the signed correlation function[26] is short ranged. This demonstrates the compensating effect such that on one side of the book, market orders attract compensating limit orders. They emphasised that market orders, all limit orders and cancellations at the bid/ask lead to a constant jump size that matches the average price change they cause. This simple picture works very well for large tick stocks, but not for small tick stocks. Also, as a generalization of the theory of the market order price impact in the associated literature, Eisler et al. (2012) pointed out that an event can induce further events[27] that amplify or dampen its own effect.

2.2 The Market Price Manipulation through Order Cancellations

Montgomery (2016) emphasises that while there exists an imbalance in the market between the buy side and the sell side, logically, participants ("unsuspecting investors") would act upon which direction they expect the price to change into. To manipulate the market price, perpetrators (or manipulators) send a huge order on one side of the order book, which creates imbalance in the market and thus misleading interpretations. These interpretations, very likely, are creating a "chain effect" where many participants actually do not wait till the large orders

[25] A limit order that falls inside the spread decreases the spread, and causes a price change in the opposite direction from market orders and cancellations.
[26] It is the function which assigns an opposite sign to the limit order and market order on the same side of the book.
[27] E.g. the arrival of excess buy market orders is shortly followed by additional sell limit orders.

are executed. Rather, they enter the market for themselves. This increasing activity by those participants ("unsuspecting investors") sets pressure on the price in one direction. This creates an opportunity (improved prices) to the perpetrators (owners of those large orders) which they use to enter the market on the opposite direction and cancel their original large orders. They, then, leave the stage with nearly risk free profits.

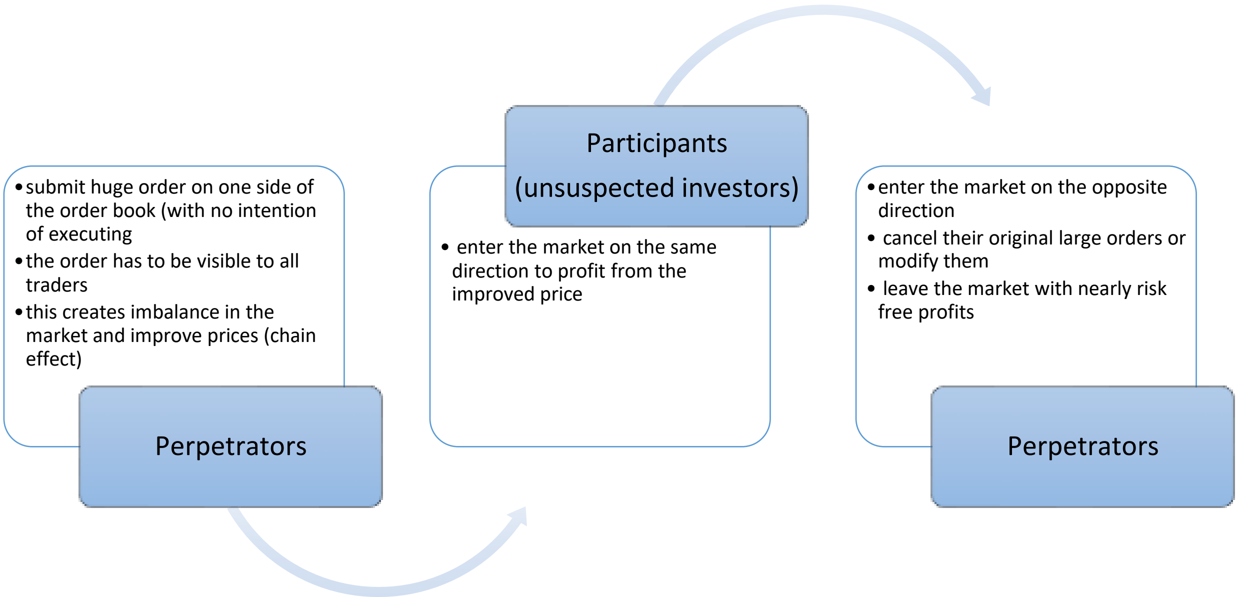


Figure 17: How manipulators could use cancellations to manipulate the market

Kyle and Viswanathan (2008) demonstrate that "'t'he term "illegal price manipulation" is difficult to define. Current US law does not explicitly define it. The finance and economics literature uses the term manipulation in an imprecise manner." They suggest that if a trader intends to hit the economic efficiency by pursuing an order which makes prices less accurate as signals for efficient resource allocation and makes markets less liquid for risk transfer, then this trading strategy is classified as "illegal price manipulation". They consider "price manipulation" and "market manipulation" as synonyms as price effects are market wide.

There are different strategies/techniques to manipulate the market price. One strategy is called Fictitious Orders, which is entering orders without an intention to execute them; this strategy is one of the simplest. Another strategy is Spoofing which is very similar to the former strategy. The third strategy is called Layering, which is the most complex one (see Montgomery, 2016).

Fictitious Large Orders: This scenario requires several assumptions to work:

1. The stock[28] should be followed by number of participants and the order book should be available to all of them.

2. Initial Large Order should be large enough[29] as compared to the average daily volume on the market of the security.

3. If short selling restrictions apply like in some countries within EU, then this account should have a position to be able to unwind.

This strategy follows three major steps: Build-Up, Un-Winding and Cancellation. During Build-Up stage of Fictitious Orders Scenario, a perpetrator[30] sends a very large limit buy order at a price below the best bid (below spread). Then, the top level of the book is instantly updated with the new order. Many participants who see this misbalance may misinterpret this misbalance (increase in demand) as probably a hedge fund having some upgrade or earnings news on this stock, so they buy in. This sets pressure on price as more existing liquidity will be consumed. Therefore, price will eventually move up (chain effect). The manipulator, then, will sell[31] this stock or in other words "un-wind" the original direction. Lastly, he either cancels

---

[28] Logically, manipulators with large trading accounts would aim at thinly traded securities and low caps. Theoretically, such an account has all means to manipulate this kind of relatively illiquid markets.
[29] At least 25% of this average daily volume should be enough to have some effect (at least short term intra-day) on the market price. Obviously, the higher the original order, the higher the odds.
[30] A perpetrator or manipulator is a trader who has a relatively large trading power.
[31] The manipulator's strategy is either selling because he had a position previously or he sells short.

the initial large order completely or modifies[32] it so that it is insignificant or not relevant, see Figure 17 above.

Spoofing: This scenario is very similar to the scenario described above with one difference in speed. However, the effect of the spoofing is smaller in terms of basis points and more short term oriented than the Fictitious Large Orders. Hence why, the activity of spoofing means sending and cancelling/modifying the large orders more frequently than in the Fictitious Large Orders scenario. Typically, this scenario is done by a "fast" trader (an ALGO or some other automated trading system). This trader sends a very large order within the spread and, then, takes it down before spread manages to adjust itself so that the large order is executed. Yet, this scenario can entice not only other ALGOs but also other carefully monitoring traders of order book.

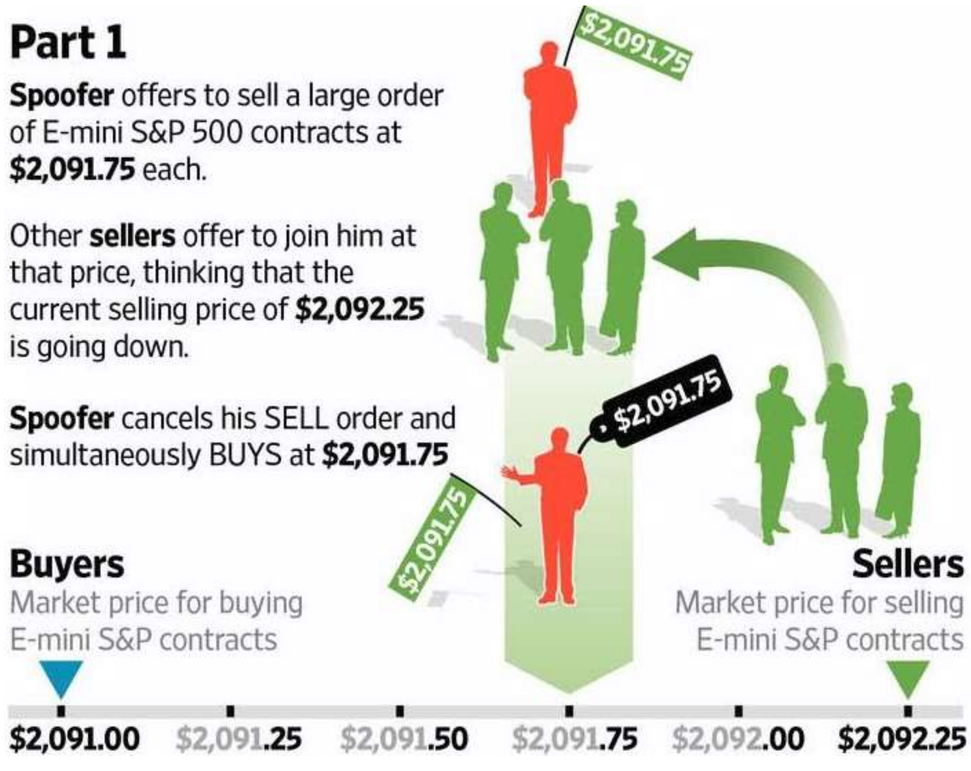


Figure 18: Illustration of Spoofing scenario. Source: Silicon Investor website[33].

[32] He either modifies the quantity or the price. He sets the price far below that level which market may reach.
[33] See: https://www.siliconinvestor.com/readmsgs.aspx?subjectid=23993&msgnum=120006&batchsize=10&batchtype=Next.

Layering: Similar to spoofing but smarter, layering is more likely to be executed by "fast" traders. In this case, the manipulator decides to mask his activity by splitting the large order into smaller sized but still significant layers[34] of multiple orders. After these layered lots, the order book would, obviously, look busier with built up demand. Other participants interpret the change in the order book as positive, so they act on it, which in turn leads price to increase. Lastly, the originator of the Layered Orders cancels all or some of the layers and sells the security for improved prices.

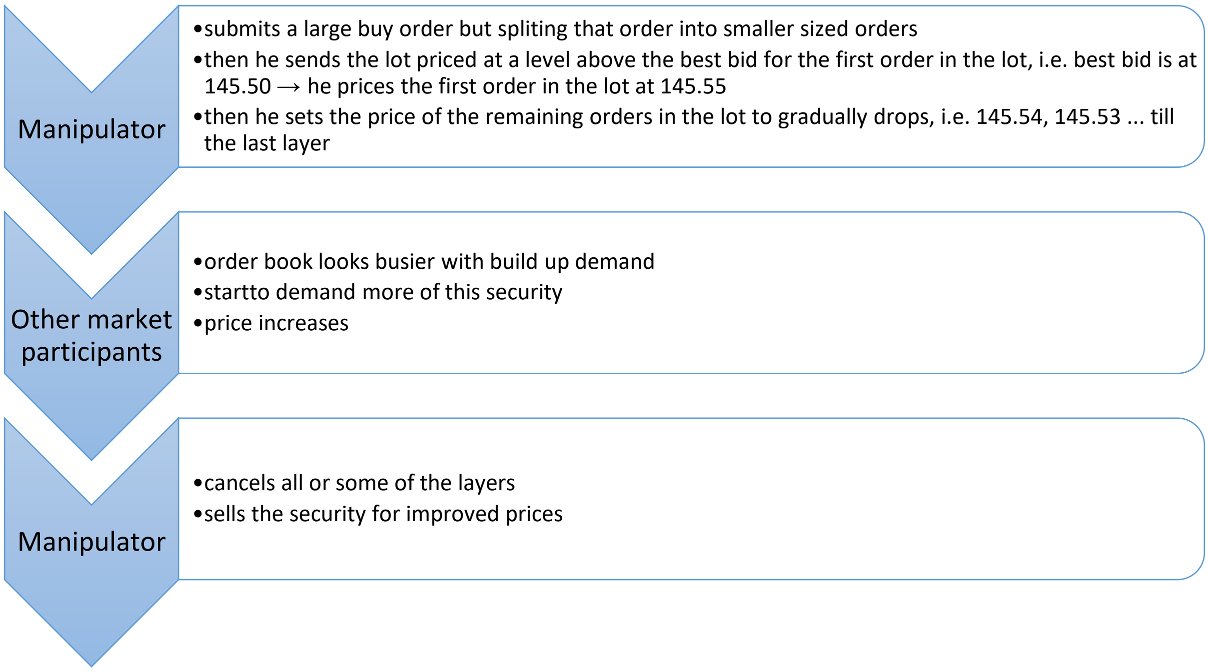


Figure 19: Market price manipulation through layering

"If the proprietary firm is layering the book with multiple bids and offers at different prices and sizes, this strategy can generate an enormous volume of orders and high cancellation rates of 90% or more", Security and Exchange Commission (2010).

[34] The manipulator would split an order of 1000 shares limit buy into 10 layers of 100 lots. He, then, sends the first 100 lot at a level above the best bid and then gradually drops (i.e. first lot is sent at 145.50 when the best bid is at 145.46 then the next at 145.49, etc.. till the tenth layer at 145.40).

2.3 Flash Crashes, High Frequency Trading and Order Book Cancellations

2.3.1 Overview

Several studies, in financial markets literature, have suggested that the rising importance of HFT has amplified the frequency and severity of flash crashes besides the high volatility of prices. Nevertheless, the debate about the benefits and costs of HFT has not been settled in the literature yet. Many empirical and theoretical studies have questioned the threatening effects of HFT which, possibly, may lead to the emergence of flash crashes (Kirilenko et al., 2011; Sornette and Von der Becke, 2011).

Security and Exchange Commission (SEC) (2010), demonstrates that "HFTs are proprietary trading firms that use high speed systems to monitor market data and submit large numbers of orders to the markets. HFTs utilize quantitative and algorithmic methodologies to maximize the speed of their market access and trading strategies." As proprietary firms are engaged in HFT, the SEC (2010) report also notes the other characteristics often attributed to those firms to be: "(1) the use of high-speed and sophisticated computer programs for generating, routing, and executing orders; (2) use of co-location services and individual data feeds offered by exchanges and others to minimize network and other types of latencies; (3) very short time-frames for establishing and liquidating positions; (4) the submission of numerous orders that are cancelled shortly after submission; and (5) ending the trading day in as close to a flat position as possible (that is, not carrying significant, unhedged positions overnight)."

Moreover, Leal et al. (2016) have associated the high price volatility and the observation of flash crashes to the presence of HFTs in the market. They pointed out two salient characteristics of HFT which explain the emergence of flash crashes; namely the ability of HFTs (i) to raise bid-ask spreads in the limit order book (LOB) by grasping market liquidity, and (ii) to synchronize on the sell-side of the LOB. However, they reported that sharp drops in prices were observed with the contemporaneous concentration of Low Frequency Traders' (LFTs) orders on the buy-side of the book.

Generally, contrary to LFTs who base their trading strategies on chronological time, HFTs apply event-based strategies. SEC (2010) describes HFTs behaviour/strategies in the market as they "now take advantage of low-latency systems and liquidity rebates by submitting large numbers of non-marketable orders[35] (often cancelling a very high percentage of them), which provide liquidity to the market electronically." Nevertheless, Leal et al. (2016) developed an agent-based model of a limit order book market in which heterogeneous HFTs interact with LFTs. They studied two different scenarios to assess the contribution of HFTs; only LFTs interact with each other in the first scenario, whereas both HFTs and LFTs interact with each other in the second scenario. They concluded that order cancellations by HFTs have an uncertain effect on price dynamics and fluctuations. They reported that on one occasion, high rates of order cancellation lead to higher volatility and more frequent flash crashes. However, on another occasion, they suggested that order cancellations also imply faster price-recoveries, which lessen the duration of flash crashes.

Furthermore, SEC (2010, 48) emphasis that "except for securities in the lowest 20% of typical buy-side depth, increasingly more severe drops in price is associated with ever-larger drops in liquidity". Yet, the report adds that the "securities that had both the worst decline in that depth and the greatest intraday average of buy-side depth had the most severe average price drop".

2.3.2 6th of May 2010, HFTs and Cancellations

In 2010, all it took was a matter of minutes and nearly $1 trillion in market value was wiped off US share values in Wall Street's infamous Flash Crash. The SEC (2010) report points out

[35] The submission of numerous orders is followed by cancellation shortly after submission, SEC (2010) Report.

that there was no uniformity in response to market conditions on May 6 amongst HFTs. Some of those traders have stopped their trading activity[36], others have reduced it while the remaining HFTs have continued to trade actively. Different motivations were behind the decision of HFTs to continue trading according to the SEC (2010). Some of them thought that algorithms under the extreme market conditions observed that afternoon would still be able to operate successfully (profitably). Later on, their aggregate trading activity picked up and increased significantly during the period in which the broad indices were rapidly declining.

After almost five years, the US regulators believe that a British high-frequency trader was largely to blame. A spotlight on trading practices after the arrest of a futures trader operating out of a suburban house in London in connection with the US equity market "Flash Crash" in 2010. Navinder Singh Sarao, named The Hound of Hounslow, has been accused of manipulating the market. He is also charged with wire and commodities fraud. In a criminal complaint, the Department of Justice in the US alleges that Navinder Singh Sarao manipulated the market for S&P 500 futures contracts, known as E-minis[37], using an automated trading program. It is alleged that he used this automated program to generate large sell orders. He then cancelled the trades after the price had fallen giving himself a chance[38] to buy new contracts before a rebound. Designed to mimic the market and push prices sharply in one direction, it is said that Mr Sarao used a variety of trading techniques. These techniques include spoofing and layering.

On the day of the Flash Crash itself, The United States Department of Justice stated that the Dow Jones Industrial Average fell by approximately 600 points in a five-minute span,

[36] This is for reasons such as the triggering of their internal risk parameters due to rapid price moves and subsequent data-integrity concerns.
[37] E-minis contracts are traded on the Chicago Mercantile Exchange, the largest US futures market.
[38] He has profited from other investors following the pattern or exiting the market.

following a drop in the price of E-Minis[39]. US authorities claim that he made over $40 million in total (≈£27 million) from his trades where they assume that the manipulations continued for five years right up until he was caught. Navinder Singh Sarao was arrested at his parents' home in Hounslow west London where the suburban surroundings bolide a scale of the accusations against him (see DOJ criminal complaint named "United States of America v. Navinder Singh Sarao", 2015[40]; Financial Times[41]).

Interestingly, the focus here is on equations that are written into computer code and acting with a bunch of a very expensive infrastructure autonomously from the intervention of a human. Besides, lawyers have argued that Sarao's conduct, as it is described in US court filings, does not amount to a criminal offense in the UK. Furthermore, Davis et al. (2013, p. 9) point out that "(c)onceptually, spoofing is easily understood, but it is far harder to enforce regulations against it". This suggests that there should be a serious review for the laws and regulations that control trading activities in the market. There have been many cases in the history of blaming individuals of a market crash but Sarao is the most recent one. He was applying the aforementioned strategy of HFTs which is explained in SEC (2010) report.

3 Motivation and Methodology

3.1 Motivation

Machain and Dufour (2013) suggested that HFTs are feasibly acting to move quotes by submitting and quickly cancelling limit orders to rapidly profit from trading the induced transitory mispricing. They, also, made the point that non-HFTs' limit order cancellations convey information while HFTs' cancellation activity is not informative. Furthermore, at the

[39] See The United States Department of Justice's report at https://www.justice.gov/opa/pr/futures-trader-charged-illegally-manipulating-stock-market-contributing-may-2010-market-flash.
[40] Available online at: https://www.justice.gov/sites/default/files/opa/press-releases/attachments/2015/04/21/sarao_criminal_complaint.pdf.
[41] Available online at: https://www.ft.com/content/a1114d60-e8d0-11e4-87fe-00144feab7de.

best price, a cancellation of the last limit order widens the spread by either increasing the best ask or decreasing the best bid. Additionally, Farmer et al. (2004) have suggested that liquidity fluctuations drive the large price fluctuations in any market. On the other hand, Leal et al. (2016) suggested that order cancellations by HFTs have an uncertain effect on price dynamics and fluctuations. They, also, reported that on one occasion, high rates of order cancellation lead to higher volatility and more frequent flash crashes.

In light of HFT and flash crashes, the previous chapter has investigated the shape of bid and ask sides of the order book using the NVWAP concept. It suggested that within an algorithmic trading system, as market state changes, the change in the shape of bid and ask sides of the order book can be used to switch between different trading strategies. Thus, understanding the order submissions and cancellations on each side of the order book will help to reveal more information about price changes in the market.

The information conveyed by the NVWAP curves is greatly useful as an indicator of prevailing market trends. Further, the information delivered by the NVWAP curves suggests that the market activity on each side of the limit order book directly affects the shape of liquidity supply and demand curves. These curves, in turn, depict the expectations of market participants regarding current and future price movements. Hence, the cumulative impact of market events is expected to have observable changes on these curves. More specifically, the change in the slope of the NVWAP curves, and the change in the total volume on each side of the limit order book can be used within an algorithmic trading strategy to exploit trading opportunities.

Closely linked to the aforementioned trading strategies, market participants[42] act upon which direction they expect the price to change into. For instance, when they observe that buy orders exceed[43] sell orders for a stock, those market participants think that the price should go up, thus, they would very likely buy this stock not waiting till large orders actually get executed. Rather, they enter the market for themselves. This increasing activity sets pressure on the price in one direction. This creates an opportunity (improved prices) to the manipulators which they use to enter the market on the opposite direction and cancel their initial large orders. As the price starts rising, investors will compete to consume liquidity quickly to take advantage of the potential higher returns. This competition will attract higher volume to the demand side. At the same time, volume on the supply side will decrease as fewer traders become willing to sell.

Leal et al. (2016) concluded that order cancellations by HFTs have an uncertain effect on price dynamics and fluctuations. They reported that on one occasion, high rates of order cancellation lead to higher volatility and more frequent flash crashes. However, on another occasion, they suggested that order cancellations also imply faster price-recoveries, which lessen the duration of flash crashes. Therefore, on one hand, this study hypothesises that cancellations have an impact on the market (price dynamics measured by the market impact) to the extent that this phenomenon may lead to market crashes. Particularly, this study inspects whether cancellations affect the size of price drops. Generally, it is expected that when the market price falls (downtrend), the bid side gradually decreases. The bid side NVWAP curve slowly contracts and becomes steeper. The NVWAP curve on ask (liquidity supply) side curve

[42] These participants are referred to as unsuspecting investors (Montgomery, 2016).

[43] A manipulator sends a huge order on one side of the order book, which creates imbalance in the market and thus misleading interpretations.

then slowly expands and becomes flatter. This effect can be seen in Figure 20 in the shape of the NVWAP curves on the bid side where the curves slowly contract and become steeper.

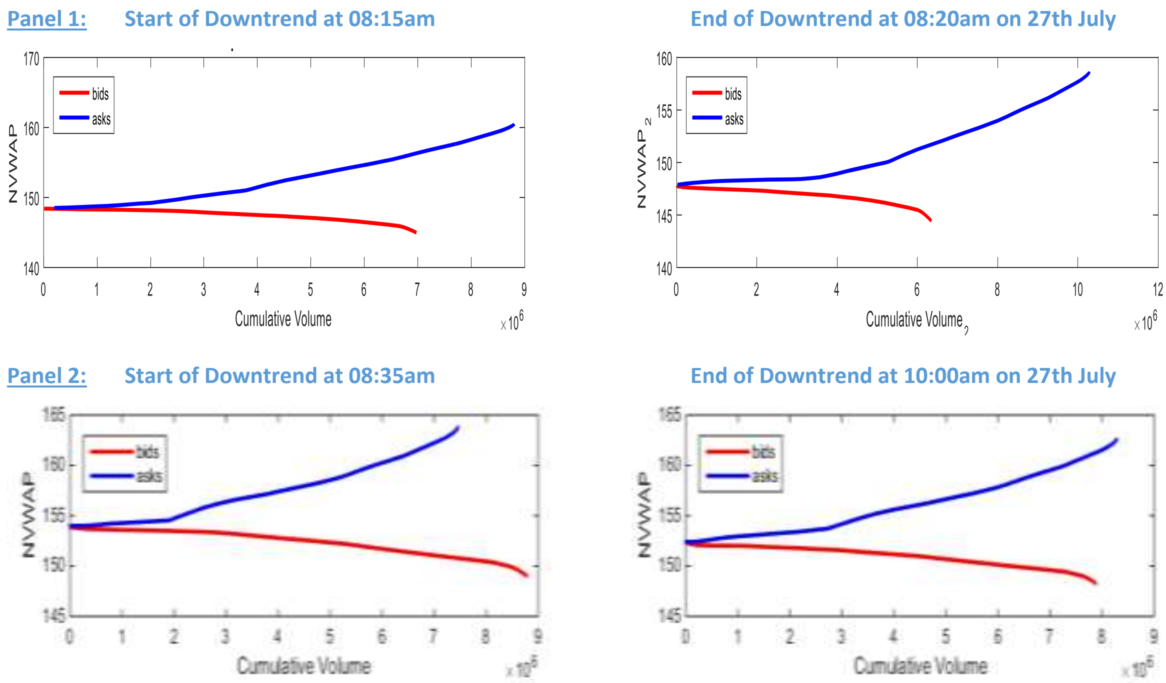


Figure 20: The NVWAP curves show the start (8:15 AM) and end (8:20 AM) of downtrend for panel 1 and the start (8:35 AM) and end (10:00 AM) of downtrend for panel 2 on 27th of July 2007 for Vodafone stock. The ask side NVWAP curve gradually expands and becomes flatter; the bid side NVWAP curve contracts and steepens. The curves in Panel 2, however, experienced larger price change.

In other words, it is expected in the downtrend price movements that new orders submitted to the buy side of the order book as well as cancellations from the sell side of the order book would decrease. It is, also, expected that new orders submitted to the sell side of the order book and cancellations from the buy side of the order book under the same price trend would increase. On the other hand, this study, also, hypothesises that cancellations have an effect on the market volatility, explicitly the bid-ask spread. It is noteworthy that realized volatility could alternatively be used instead of market volatility to examine the effect of cancellations on the market.

As cancellations behaviour affects both the size of the price drop and volatility by broadening them, this might cause a flash crash to the market. Therefore, this study inspects whether:

1. Cancellations contribute to the size of price drops (have large market impact).
2. Cancellations increase the market volatility.

The data used in this study allows comparing between two market states, volatile and quiet, as July 2007 was the month before the Financial Crisis in 2007 (quiet) while August 2007 covers its starting (volatile). Though, due to limitations in the dataset used in this study, it is not possible to identify cancellations causes, namely manipulation or price discovery. In other words, the dataset doesn't allow testing different trading strategies.

By rebuilding the E-LOB for London SETS in July and August in 2007, this study reconstructs the demand and supply curves to analyse price impact and fluctuations with special focus on orders' cancellations. The aim is to investigate how cancellations on both sides of the order book contribute to the size of the price drop (price dynamics) explanation. It, also, aims at investigating the cancellations behaviour on both sides of the order book and their effect on market volatility (denoted by the newly introduced bid-ask spread measure).

To this end, different model specifications are defined and tested comparing scenarios where variables are/aren't adjusted to duration from the peak to trough price trend. Duration adjusted models are expected to perform better than non-adjusted ones in explaining the cancellations effect on the market. All analysis is based on ultra-high-frequency data from London SETS which investigates the periods where the market was experiencing a downtrend price move (from peak to trough). Uptrend price move's effects are supposed to be similar to the downtrend price movements but opposite in signs. Robust regressions[44], which control for efficiency under more realistic conditions, are run to check the validity of the regression results obtained.

3.2 Methodology

This study aims at investigating the effect of cancellations on the market (price dynamics measured by the market impact), to the extent that this phenomenon may lead to market crashes. Particularly, this study inspects whether cancellations affect the size of price drops. It, also, examines whether cancellations have an effect on the market volatility, explicitly bid-ask spread measure. This section presents the econometric methodology used to examine the dynamics of cancellations in the limit order book of London SETS platform.

Various models to study the price impact have been proposed in the literature; however, little agreement on how to model that impact is in place (see Bouchaud, 2010; Cont et al., 2014). The intuitive notion that prices move due to the imbalance between demand and supply seems to be the only consensus. Traditionally, cancellations have been treated as part of order book imbalance measures. For a limit order book market, Obizhaeva and Wang (2005) have pointed out that when trading times are chosen optimally, the dynamics of the supply/demand play a key role in determining the optimal execution strategy. Moreover, Lo and Hall (2015) have suggested that studies based on calendar time suffer from many shortcomings. Thus, choosing an interval length for time aggregation becomes necessity when using calendar time. Choosing the appropriate interval length depends on the level of limit order activity and would likely vary among stocks in the sample. On the other hand, a potential loss of information results when aggregating data containing all order events occurring within an interval.

[44] See Andrews (1974).

Therefore, the methodological choice of sampling frequency in this paper will deviate from the typical choice in the related literature where event time or calendar time is used. This study will base the sampling frequency for the data set of order events on the directional changes (DC) approach[45] where peaks and troughs in cumulative return are identified. This methodology is followed in a previous work where time duration from peak to trough is defined as the time that the market takes to change its state[46]. Peaks and troughs in cumulative return are located by setting a minimum cumulative return change threshold of 25 basis points (bp) to identify local peaks and troughs. Peak to Trough represents the period where the market return is experiencing a downward price movement, see Table 18.

Stock / Month	Number of Peaks and Troughs Identified during the month				
	HSBA	BP	TSCO	AZN	VOD
July	206	135	175	171	184
August	352	250	215	217	259

Table 18: This table shows how many peaks and troughs were identified when the cumulative return changed by at least 25 basis points

Furthermore, this study differs from the definition of resiliency variable in Mayston et al. (2008) and Tóth et al. (2011) who have studied the resiliency of the order book market following a liquidity shock. Mayston et al. (2008) define resiliency as "the speed with which the temporary order-flow related changes induced in depth and in spreads by an order-flow shock are corrected or neutralized by the flow of new orders into the market through the competitive actions of value traders, liquidity suppliers and others." Relatively, the resiliency

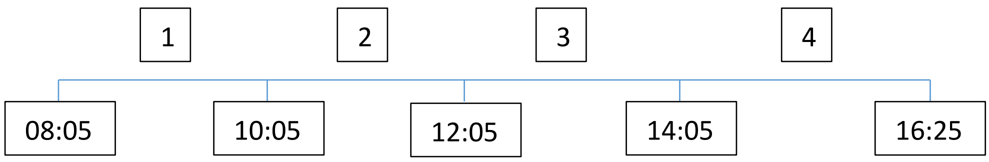
[45] See Tsang (2017).
[46] Peaks and troughs methodology is explained in the previous chapter.

variable is redefined in this study to be the time taken (duration) for the change in cumulative return from peak to trough.

With that in mind, this study follows the literature by investigating the volume, change in price and resiliency (time/duration) variables as well as the volatility and the change in cumulative return (price impact) (see Gabaix, 2003; Obizhaeva and Wang, 2005; Lo and Sapp, 2008; Mayston et al., 2008; Tóth et al., 2011; Eisler et al., 2012; Farmer et al., 2013; Lo and Hall, 2015). Therefore, this paper investigates how order cancellations, the change in spread, the change in cumulative returns, the duration from peak to trough and whether an extreme price change impact the market. All analysis is based on ultra-high-frequency data for 5 stocks from the E-LOB of London SETS which covers the periods where the market was experiencing a downtrend price change (from peak to trough).

To establish some stylised facts about the data this study considers, a new dummy variable, which is the $TimeOfTheDay_i$ variable. This variable will account for the time of the day in which the order has been submitted to the order book or cancelled to identify whether cancellations trend varies with trading hours. It is expected that more traders join the trading session from the US (around Mid-day UK time when the trading day in the US starts). In other words, we split the day into 4 periods as follows:

Figure 21: Time line to show the definition of the Time of The Day variable.



Dividing the trading day into four periods extends Mayston et al. (2008) work. They have used a time dummy which accounts only for the end of the trading day in their study. The reason behind defining this variable is to account for the new traders who trade stocks listed in the London SETS from outside the UK, namely from the US or any other country. As these traders join the trading session, they may affect the number of submissions of new orders as well as orders cancellations. The $TimeofTheDay_i$ variable takes the values from 1 to 4. Each value covers a period of at least two consecutive hours starting from 08:05 and ending at 16:25.

This study hypothesizes that order cancellations affect the market or the price dynamics (measured by the market impact) to the extent that this phenomenon may lead to market crashes. Particularly, this study inspects whether cancellations affect the size of price drops during downtrend price movements. Market impact is defined as the change in the stock cumulative return from peak to trough ($\Delta\text{ReturnFromPeakToTrough}_\text{i}$). The stock cumulative return, r, is calculated following the formula:

$$r = \log(P_{m,t+1}) - \log(P_{m,t}) \qquad (3.2.1)$$

where P_m is the mid price, and is given as $\text{P}_\text{m} = \frac{\text{P}_\text{bid}+\text{P}_\text{ask}}{2}$. Then the change in return is the first difference between the cumulative return at the peak corrected for the first minute and at the trough corrected for the last minute, such that:

$$\Delta\text{ReturnPeakToTrough}_\text{i} = r_{P+1} - r_{T-1} \qquad (3.2.2)$$

where r_{P+1} is the cumulative return at the peak time plus one minute, and r_{T-1} is the cumulative return at the trough time minus one minute. Adding and deducting one minute aims at discarding changes in cumulative returns at the turning points.

This study, also, theorises that order cancellations have an effect on the market volatility, explicitly the new measure of bid-ask spread. Farmer et al. (2004) have suggested that the granularity of fluctuations in supply and demand remains the key factor underlying extreme price fluctuations. Therefore, we will account for liquidity fluctuations using the $\Delta\text{NVWAP Slope}_\text{i}$ (discussed in the previous chapter) as a proxy for the Bid-Ask spread. Therefore, to test the second hypothesis, the Bid-Ask spread is defined as the Bid-Ask difference between the changes in the logarithms of the slopes of the NVWAP curves on the buy and sell sides from peak to trough ($\Delta NVWAPBuySellPeakToTrough_i$). The slope of the NVWAP curve (β_1) in each side of the E-LOB is estimated by the regression:

$$\log(Y_i) = \log\beta_0 + \beta_1 X_i \qquad (3.2.3)$$

where Y_i is the market impact and X_i is the cumulative volume normalized to the i-th order at which market side that a trader wants to trade. X_i is calculated as follows:

$$X_i = \frac{\sum_{j=1}^{i} Vol_j}{ADV} \times 100 \qquad (3.2.4)$$

The exponential transformation $Y = \beta_0 e^{\beta_1 X}$ is performed to ensure that the supply curve is strictly increasing for $\beta_0, \beta_1 > 0$, as assumed by definition. The market impact Y_i or the premium a trader pays for executing a volume larger than the amount available at best price is defined as:

$$Y_i = \left(\frac{NVWAP_i - P_m}{P_b}\right) \times 10000 \qquad (3.2.5)$$

where P_m and P_b are mid- and best-price respectively. The NVWAP is the average expected transaction price for a trade size equal to the cumulative volume at the i-th level, and is calculated as:

$$NVWAP_i = \frac{\sum_{j=1}^{i} P_j \times Vol_j}{\sum_{j=1}^{i} Vol_j} \qquad (3.2.6)$$

where P_j is the price and Vol_j is the volume at order j in the book.

All regressions run are non-linear and are linearized using logs. Also, differences' variables are used to minimise the estimation bias and may be omitting it. Following a downtrend move in the price, it is expected that cancellations on the buy (sell) side of the market will increase (decrease)[47]. Hence, as shown in Table 19, the buy (sell) cancellations coefficient will be positive (negative). On the other hand, it is expected that submission of new orders on the buy (sell) side of the market will decrease (increase). This means that the coefficient of new orders on the buy (sell) side is negative (positive), see Table 19 below.

Event \ Side	Buy	Sell
Cancellations	+	-
New Orders	-	+

Table 19: This table emphasises the expected signs of cancellations and new orders' coefficients in each side of the market, namely buy and sell, following a downtrend move in the price.

All regressions performed account for two levels of cumulative return change in basis points as dependant variable, normal and extreme. The stock cumulative return change is considered to be normal if it does not exceed 100bp while it is considered as extreme change when it exceeds the 100bp level as demonstrated in (3.2.7) below, unless stated otherwise.

$$\Delta\text{ReturnPeakToTrough} \stackrel{\text{def}}{=} \begin{cases} \text{Normal } \forall \ \Delta\text{ReturnPeakToTrough} < 100\text{bp.} \\ \text{Extreme } \forall \ \Delta\text{ReturnPeakToTrough} \geq 100\text{bp.} \end{cases} \quad (3.2.7)$$

The dependant variable is then adjusted to duration to define other model specifications to analyse the minutely market impact, i.e. the minutely change in cumulative return from peak to trough in basis points ($\Delta ReturnPeakToTroughPerMinute_i$). To adjust the ΔReturnPeakToTrough defined in 3.2.7, the stock cumulative return, r, is calculated following equations 3.2.1 to 3.2.6. Then the change in return is calculated by taking first difference

[47] See Höschler (2011).

between the cumulative return at the peak and at the trough. Lastly, this change in return is normalised to duration as follows:

$$\Delta ReturnPeakToTroughPerMinute_i = \left(\frac{\Delta ReturnPeakToTroguh_i}{Duration_i}\right) \times 10000 \qquad (3.2.8)$$

where $Duration_i$ is the time duration in minutes from peak to trough, i.e. the difference between the start time and end time of a downtrend price move.

$$\mathrm{Duration_i} = t_e - t_s \quad (3.2.9)$$

t_e and t_s are the end time and start time of the i^{th} downtrend price move, respectively.

After that, $\Delta ReturnPeakToTroughPerMinute_i$ is defined as normal and extreme as well. To set a threshold after which change in cumulative return from peak to trough per minute (higher market impact) is recognised as extreme, the $\Delta ReturnPeakToTrough_i$ reporting criterion demonstrated in equation (3.2.7) is restructured following more strict criteria. To decide the cut-off point on which the reporting of $\Delta ReturnPeakToTroughPerMinute_i$ is based, the descriptive statistics for the new variable[48] are stated in Table 34. As Table 34 below best illustrates, the mean value of the $|\Delta ReturnPeakToTroughPerMinute_i|$ is 1.64 in July compared to 2.26 in August. Using August's statistics is more convenient due to the fact that the financial crisis has already kicked off in that month. Moreover, the top 10% values of the distribution are beyond or equal to |3.64|. This value, consequently, will be used to restructure the minutely changes in cumulative return from peak to trough ($\Delta ReturnPeakToTroughPerMinute_i$).

[48] See equation (3.2.8).

Thus, the minutely changes in cumulative return from peak to trough ($\Delta ReturnPeakToTroughPerMinute_i$) is split into two categories[49] to shed the light on the market impact of large price drops which happen in a shorter time span, i.e. have larger impact, such that:

$$\Delta ReturnPeakToTroughPerMinute_i \overset{\text{def}}{=} \begin{cases} \text{Normal } \forall\ \Delta ReturnPeakToTroughPerMinute_i < |3.64|. \\ \text{Extreme } \forall\ \Delta ReturnPeakToTroughPerMinute_i \geq |3.64|. \end{cases} \tag{3.2.10}$$

Moreover, Chordia et al. (2008) make the point that over the recent years, algorithmic trading was behind the increase in trading volume as well as the drop in the average trade size. Thus, the increase in trading volume may explain the concerns about the impact of algorithmic trading on price discovery process and volatility. With that in mind, different independent variables are defined and used to build different model specifications. These variables are defined as follows:

1- $TotalBuyCancellationToADV_i$ (equation 3.2.11) and $TotalSellCancellationToADV_i$ (equation 3.2.12) are the total volume of all cancellations occurred on the buy and sell sides, respectively, during the span from peak to trough, then normalising (dividing) that volume to the average daily volume.

$$TotalBuyCancellationToADV_i = \sum BuyCancellation_i / ADV \tag{3.2.11}$$

$$TotalSellCancellationToADV_i = \sum SellCancellation_i / ADV \tag{3.2.12}$$

[49] Generally, for any change in cumulative return to be classified as extreme, it should be no less than 100 basis points and happen in a time span shorter or equal 30 minutes. Although 50 basis points change is not considered as extreme change in cumulative return, it is believed to be high enough to hit the market. Additionally, the reason why 30 minutes is considered here is that at the time of the Flash Crash in 2010, the market took 20 minutes to crash.

2- $TotalBuyNewToADV_i$ (equation 3.2.13) and $TotalSellNewToADV_i$ (equation 3.2.14) are the total volume of all new orders submitted to the buy and sell sides, respectively, during the span from peak to trough, then normalising (dividing) that volume to the average daily volume.

$$TotalBuyNewToADV_i = \sum BuyNew_i / ADV \quad (3.2.13)$$

$$TotalSellNewToADV_i = \sum SellNew_i / ADV \quad (3.2.14)$$

3- ln $TotalBuyCancellation_i$ and ln $TotalSellCancellation_i$ are the logarithm of the total volume of all cancellations occurred on the buy and sell sides, respectively, during the span from peak to trough.

$$\ln TotalBuyCancellation_i = \ln(\sum BuyCancellation_i) \qquad (3.2.15)$$

$$\ln TotalSellCancellation_i = \ln(\sum SellCancellation_i) \qquad (3.2.16)$$

4- ln $TotalBuyNew_i$ and ln $TotalSellNew_i$ are the logarithm of the total volume of all new orders submitted to the buy and sell sides, respectively, during the span from peak to trough.

$$\ln TotalBuyNew_i = \ln(\sum BuyNew_i / ADV) \quad (3.2.17)$$

$$\ln TotalSellNew_i = \ln(\sum SellNew_i / ADV) \quad (3.2.18)$$

5- $\ln AverageBuyCancellation_i$ (equation 3.2.19) and $\ln AverageSellCancellation_i$ (equation 3.2.20) are the logarithm of the average buy and sell cancellations, i.e. the total volume of all cancellations occurred on the buy and sell sides, respectively, during the span from peak to trough divided by the number of cancellation events/orders during the same time span.

$$\ln AverageBuyCancellation_i = \ln(\sum BuyCancellation_i / \#CancellationEvents) \quad (3.2.19)$$

$$\ln AverageSellCancellation_i = \ln(\sum SellCancellation_i / \#CancellationEvents) \quad (3.2.20)$$

6- $\ln AverageBuyNew_i$ (equation 3.2.21) and $\ln AverageSellNew_i$ (equation 3.2.22) are the logarithm of the average new buy and sell orders, i.e. total volume of all new orders submitted to the buy and sell sides, respectively, during the span from peak to trough divided by the number of new submission events/orders during the same time span.

$$\ln AverageBuyNew_i = \ln(\sum BuyNew_i / \#NewEvents) \qquad (3.2.21)$$

$$\ln AverageSellNew_i = \ln(\sum SellNew_i / \#NewEvents) \qquad (3.2.22)$$

Additionally, to answer the second research question of this study, i.e. how cancellations affect the market volatility (the bid-ask spread), the change in bid-ask spread from peak to trough (i.e. when the market experiences a downtrend price move) is used as a proxy to measure the market volatility, and is set as the dependent variable. The ordinary bid-ask spread definition is not informative in the dataset used because it changes either by 0.5 or 1. Therefore, the dependant variable in this set of models (the bid-ask spread) is defined as the Bid-Ask difference between the change in the logarithm of slope of the NVWAP curves from peak to trough $\Delta \mathrm{NVWAPBuySellPeakToTrough}_i$.

To calculate this variable, first, the slope of the NVWAP curve in each side of the E-LOB is estimated by equations (3.2.3) to (3.2.6). Then, $\Delta \mathrm{NVWAPBuySellPeakToTrough}_i$ is calculated by estimating the difference between the logarithms of slopes of the NVWAP curves in the bid and ask sides of the E-LOB as follows:

$$\Delta \mathrm{NVWAPBuySellPeakToTrough}_i = \Delta \mathrm{NVWAPBuyPeakToTrough}_i - \Delta \mathrm{NVWAPSellPeakToTrough}_i \qquad (3.2.23)$$

where:

- $\Delta \mathrm{NVWAPBuyPeakToTrough}_i$ is the change in the logarithm of slope of the NVWAP curves in the bid side of the E-LOB from peak to trough, which is calculated as:

$$\Delta \text{NVWAPBuyPeakToTrough}_i = \text{SlopeNVWAPBuyPeak}_i - \text{SlopeNVWAPBuyTrough}_i \quad (3.2.24)$$

- $\Delta \text{NVWAPSellPeakToTrough}_i$ is the change in the logarithm of slope of the NVWAP curves on the sell side of the E-LOB from peak to trough, which is calculated as:

$$\Delta \text{NVWAPSellPeakToTrough}_i = \text{SlopeNVWAPSellPeak}_i - \text{SlopeNVWAPSellTrough}_i \quad (3.2.25)$$

Finally, to check the validity of the previous regressions, robust regressions will be used to down-weight the effect of any outliers or standard errors in the estimations, and to generalise the findings. Therefore, robust regressions control for efficiency under more realistic conditions. Different sets of robust regressions are run; namely, standard errors and clustered robustness checks are used. Below are the models defined to test this paper's hypotheses.

3.2.1 MARKET IMPACT ANALYSIS (The Effect of Orders Cancellations on the Size of Cumulative Return Drops from Peak to Trough)

To answer the first research question of this study, i.e. how cancellations affect the market (market impact or price dynamics), different model specifications of cancellations are defined and run. In particular, the effect of cancellations on the size of price drops is studied under various provisions. The size of drop in cumulative return from peak to trough (when the market experiences a downtrend price move) is used as a proxy to denote the market impact, and is used as the dependent variable. Below are the models' specifications used to study how cancellations impact the market.

MODEL SPEC (1): The Effect of Total Volume of Cancelled and New Orders Normalised by the Average Daily Volume (ADV) on the Size of Drop in Cumulative Return from Peak to Trough Measured in Basis points

In this model specification, using the total volumes of cancellations and submissions (new orders) on both sides of the E-LOB after normalising them to the average daily volume, the effect of cancellations on the size of price drop is investigated. This normalisation is performed to comply with the fairly small number of the change in cumulative return from peak to trough. The size of the price drop is used as a proxy of market impact, and is defined as the change in cumulative return from peak to trough in basis points. This model is as follows:

$$\Delta ReturnPeakToTrough_i = \beta_0 + \beta_1 TotalBuyCancellationToADV_i + \beta_2 TotalBuyNewToADV_i + \beta_3 TotalSellCancellationTOADV_i + \beta_4 TotalSellNewToADV_i + e_i \qquad (3.2.1.1)$$

where:

- $\Delta ReturnPeakToTrough_i$ is the change in the stock cumulative return from peak to trough defined earlier in equations (3.2.1 to 3.2.6).
- $TotalBuyCancellationToADV_i$ (equation 3.2.11) and $TotalSellCancellationToADV_i$ (equation 3.2.12) are the total volume of all cancellations occurred on the buy and sell sides, respectively, during the span from peak to trough, then normalising (dividing) that volume to the average daily volume.
- $TotalBuyNewToADV_i$ (equation 3.2.13) and $TotalSellNewToADV_i$ (equation 3.2.14) are the total volume of all new orders submitted to the buy and sell sides, respectively, during the span from peak to trough, then normalising (dividing) that volume to the average daily volume.

MODEL SPEC (2): The Effect of Logged Total Volume of Cancelled and New Orders on the Size of Drop in Cumulative Return from Peak to Trough Measured in Real Points (Basis Points×10^4)[50]

This model specification uses the logarithms of total volumes of cancellations and submissions on both sides of the E-LOB to study the effect of cancellations on the size of price drop. The right hand side of the equation (3.2.1.2) is logged to conform the left hand side variable which is reported in basis points and is a difference of logged values. The size of the price drop is defined as the change in cumulative return from peak to trough which describes the market impact. This model is of the shape:

$$\Delta RPReturnPeakToTrough_i = \beta_0 + \beta_1 \ln TotalBuyCancellation_i + \beta_2 \ln TotalSellCancellation_i + \beta_3 \ln TotalBuyNew_i + \beta_4 \ln TotalSellNew_i + e_i \quad (3.2.1.2)$$

where:

- $\Delta RPReturnPeakToTrough_i$ is the change in the stock cumulative return from peak to trough in real points. This change in return is calculated following the formula:

$$\Delta RPReturnPeakToTrough_i = \Delta ReturnPeakToTrough_i \times 10000 \quad (3.2.1.3)$$

where $\Delta RPReturnPeakToTrough_i$ is the change in the stock cumulative return from peak to trough defined earlier in equations (3.2.1 to 3.2.6).

- $\ln TotalBuyCancellation_i$ and $\ln TotalSellCancellation_i$ are the logarithm of the total volume of all cancellations occurred on the buy and sell sides, respectively, during the span from peak to trough.

[50] Explain why this case

- $\ln TotalBuyNew_i$ and $\ln TotalSellNew_i$ are the logarithm of the total volume of all new orders submitted to the buy and sell sides, respectively, during the span from peak to trough.

MODEL SPEC (3): The Effect of Logged Average Volume of Cancelled and New Orders on the Size of Drop in Cumulative Return from Peak to Trough Measured in Real Points (Basis Points×10^4)

To study the effect of cancellations on the size of price drop, this model specification uses the logarithms of the average volumes of cancellations and submissions on both sides of the E-LOB. The size of the price drop (the market impact) is defined as the change in cumulative return from peak to trough. Using the average values may mitigate the effect of any extreme values in the number of shares traded in each order. Therefore, this model is form as:

$$\Delta RPReturnPeakToTrough_i=\beta_0+\beta_1 \ln AverageBuyCancellation_i+\beta_2 \ln AverageSellCancellation_i$$

$$+\beta_3 \ln AverageBuyNew_i+\beta_4 \ln AverageSellNew_i+e_i \qquad (3.2.1.4)$$

where:

- $\Delta RPReturnPeakToTrough_i$ is the change in the stock cumulative return from peak to trough. This change in return is calculated as real points following formula (3.2.1.3) and (3.2.1 to 3.2.6) explained above.

- $\ln AverageBuyCancellation_i$ (equation 3.2.19) and $\ln AverageSellCancellation_i$ (equation 3.2.20) are the logarithm of the average buy and sell cancellations, i.e. the total volume of all cancellations occurred on the buy and sell sides, respectively, during the span from peak to trough divided by the number of cancellation events/orders during the same time span.

- $\ln AverageBuyNew_i$ (equation 3.2.21) and $\ln AverageSellNew_i$ (equation 3.2.22) are the logarithm of the average new buy and sell orders, i.e. total volume of all new orders submitted to the buy and sell sides, respectively, during the span from peak to trough divided by the number of new submission events/orders during the same time span.

3.2.2 MINUTELY MARKET IMPACT ANALYSIS (The Effect of Orders Cancellations on the Size of Cumulative Return Drops per Minute from Peak to Trough)

To appropriately analyse the minutely market impact, the minutely change in cumulative return from peak to trough ($\Delta ReturnPeakToTroughPerMinute_i$) is defined in 3.2.8, calculated and analysed. Models' specifications 4 (equation 3.2.2.1), 5 (equation 3.2.2.2) and 6 (equation 3.2.2.3) are run where the new structure in (3.2.10) is applied.

MODEL SPEC (4): The Effect of Total Volume of Cancelled and New Orders Normalised by the Average Daily Volume (ADV) on the Size of Cumulative Return Drop from Peak to Trough Measured in Real Points (Basis points×10^4) and Normalised to Duration (Minutely Change in Cumulative Return)

The total volumes of cancellations and submissions on both sides of the E-LOB after normalising them to the average daily volume are used in this specification to assess the effect of cancellations on the size of price drop from peak to trough. The size of the price drop is defined as the change in cumulative return from peak to trough in real points after normalising to the resiliency (duration from peak to trough). The model applied here is of the formula:

$$\Delta ReturnPeakToTroughPerMinute_i=\beta_0+\beta_1 TotalBuyCancellationToADV_i+\beta_2 TotalBuyNewToADV_i$$

$$+\beta_3 TotalSellCancellationTOADV_i+\beta_4 TotalSellNewToADV_i+e_i \qquad (3.2.2.1)$$

where:

- $\Delta ReturnPeakToTroughPerMinute_i$ is the resiliency variable. In other words, it is the change in the stock cumulative return from peak to trough normalised to the duration of the i^{th} downtrend price movement (equation 3.2.8).

- $TotalBuyCancellationToADV_i$ (equation 3.2.11) and $TotalSellCancellationToADV_i$ (equation 3.2.12) are the total volume of all cancellations occurred on the buy and sell sides, respectively, during the span from peak to trough, then normalising (dividing) that volume to the average daily volume.
- $TotalBuyNewToADV_i$ (equation 3.2.13) and $TotalSellNewToADV_i$ (equation 3.2.14) are the total volume of all new orders submitted to the buy and sell sides, respectively, during the span from peak to trough, then normalising (dividing) that volume to the average daily volume.

MODEL SPEC (5): The Effect of the Logged Total Volume of Cancelled and New Orders on The Size of Cumulative Return Drop from Peak to Trough Measured in Real Points (Basis points×10^4) and Normalised to Duration (Minutely Change in Cumulative Return)

In this model specification, the total volumes of cancellations and submissions on both sides of the E-LOB are logged (normalised) to examine the effect of cancellations on the size of price drop from peak to trough. The size of the price drop is defined as the change in cumulative return from peak to trough in real points after normalising to the resiliency. The model formula applied here is:

$$\Delta ReturnPeakToTroughPerMinute_i = \beta_0 + \beta_1 \ln TotalBuyCancellation_i + \beta_2 \ln TotalSellCancellation_i$$
$$+ \beta_3 \ln TotalBuyNew_i + \beta_4 \ln TotalSellNew_i + e_i \qquad (3.2.2.2)$$

where:

- $\Delta ReturnPeakToTroughPerMinute_i$ is the change in the stock cumulative return from peak to trough normalised to the duration of the ith downtrend price movement as previously mentioned (equations 3.2.1 to 3.2.6 and 3.2.8).

- ln $TotalBuyCancellation_i$ and ln $TotalSellCancellation_i$ are the logarithm of the total volume of all cancellations occurred on the buy and sell sides, respectively, during the span from peak to trough (equations 3.2.15 and 3.2.16).

- ln $TotalBuyNew_i$ and ln $TotalSellNew_i$ are the logarithm of the total volume of all new orders submitted to the buy and sell sides, respectively, during the span from peak to trough (equations 3.2.17 and 3.2.18).

MODEL SPEC (6): The Effect of the Logged Average Volume of Cancelled and New Orders on the Size of Cumulative Return Drop from Peak to Trough Measured in Real Points (Basis Points×10^4)

The effect of cancellations on the size of price drop is investigated in this model specification where the logarithms of the average volumes of cancellations and submissions on both sides of the E-LOB are used. The size of the price drop is defined as the change in cumulative return from peak to trough after normalising to the duration (minutely change in cumulative return). This model is of the shape:

$$\Delta ReturnPeakToTroughPerMinute_i = \beta_0 + \beta_1 \ln AverageBuyCancellation_i + \beta_2 \ln AverageSellCancellation_i + \beta_3 \ln AverageBuyNew_i + \beta_4 \ln AverageSellNew_i + e_i \qquad (3.2.2.3)$$

where:

- $\Delta ReturnPeakToTroughPerMinute_i$ is the change in the stock cumulative return from peak to trough normalised to the duration of the ith downtrend price movement as previously mentioned (equations 3.2.1 to 3.2.6 and 3.2.8).

- ln $AverageBuyCancellation_i$ (equation 3.2.19) and ln $AverageSellCancellation_i$ (equation 3.2.20) are the logarithm of the average buy and sell cancellations, i.e. the total volume of all cancellations occurred on the buy and sell sides, respectively, during the span from peak to trough divided by the quantity of cancellation events/orders during the same time span.

- ln $AverageBuyNew_i$ (equation 3.2.21) and ln $AverageSellNew_i$ (equation 3.2.22) are the logarithm of the average new buy and sell orders, i.e. total volume of all new orders submitted to the buy and sell sides, respectively, during the span from peak to trough divided by the quantity of new submission events/orders during the same time span.

3.2.3 MARKET VOLATILITY ANALYSIS (The Effect of Order Cancellations on the Spread)

To answer the second research question of this study, i.e. how cancellations affect the market volatility (the bid-ask spread), different model specifications of market events are defined. In particular, the effect of cancellations on the market volatility is analysed under various settings. The change in bid-ask spread from peak to trough (i.e. when the market experiences a downtrend price move) is used as a proxy to measure the market volatility, and is set as the dependent variable. Thus, model specifications to study how cancellations impact market volatility are described below.

The ordinary bid-ask spread definition is not informative in the dataset used because it changes either by 0.5 or 1. Therefore, the dependant variable in this set of models (the bid-ask spread) is defined as the Bid-Ask difference between the change in the logarithm of slope of the NVWAP curves from peak to trough $\Delta \mathrm{NVWAPBuySellPeakToTrough}_i$. This variableis calculated is estimated following equations 3.2.3 to 3.2.6 and 3.2.23 to 3.2.25.

On the other hand, the right hand side of the models are either normalised to the ADV or logged to maintain consistency measures. Below are the models proposed to investigate the second research question.

MODEL SPEC (1): The Effect of Logged Total Volume of Cancelled and New Orders on The $\Delta NVWAP_i$ from Peak to Trough

This model specification uses the logarithms of total volumes of market events, namely cancellations and submissions, on both sides of the E-LOB to study the effect of cancellations on the spread. The right hand side of the equation (3.2.3.1) below is logged to conform the left hand side variable, which is a difference of logged values. This model studied here is the following:

$$\Delta NVWAPBuySellPeakToTrough_i = \beta_0 + \beta_1 Duration_i + \beta_2 \ln TotalBuyCancellation_i + \beta_3 \ln TotalSellCancellation_i + \beta_4 \ln TotalBuyNew_i + \beta_5 \ln TotalSellNew_i + e_i \qquad (3.2.3.1)$$

where:

- $\Delta NVWAPBuySellPeakToTrough_i$ is the Bid-Ask difference between the change in the logarithm of slope of the NVWAP curves from peak to trough (equation 3.2.23).
- $Duration_i$ is the time duration in minutes from peak to trough, i.e. the difference between the start time and end time of a downtrend price move (equation 3.2.9).
- $\ln TotalBuyCancellation_i$ and $\ln TotalSellCancellation_i$ are the logarithm of the total volume of all cancellations occurred on the buy and sell sides, respectively, during the span from peak to trough (equations 3.2.15 and 3.2.16).
- $\ln TotalBuyNew_i$ and $\ln TotalSellNew_i$ are the logarithm of the total volume of all new orders submitted to the buy and sell sides, respectively, during the span from peak to trough (equations 3.2.17 and 3.2.18).

MODEL SPEC (2): The Effect of Logged Average Volume of Cancelled and New Orders on The $\Delta NVWAP_i$ from Peak to Trough

To study the effect of cancellations on market volatility (the spread), this model specification uses the logarithms of the average volumes of cancellations and submissions on both sides of the E-LOB. The reason behind using the average values is that they may mitigate the effect of any extreme values in the number of shares traded in each order. Therefore, this model takes the form:

$$\Delta NVWAPBuySellPeakToTrough_i = \beta_0 + \beta_1 Duration_i + \beta_2 \ln AverageBuyCancellation_i + \beta_3 \ln AverageSellCancellation_i + \beta_4 \ln AverageBuyNew_i + \beta_5 \ln AverageSellNew_i + e_i \quad (3.2.3.2)$$

where:

- $\Delta NVWAPBuySellPeakToTrough_i$ is the Bid-Ask difference between the change in the logarithm of slope of the NVWAP curves from peak to trough (equation 3.2.23).
- $Duration_i$ is the time duration in minutes from peak to trough, i.e. the difference between the start time and end time of a downtrend price move (equation 3.2.9).
- $\ln AverageBuyCancellation_i$ (equation 3.2.19) and $\ln AverageSellCancellation_i$ (equation 3.2.20) are the logarithm of the average buy and sell cancellations, i.e. the total volume of all cancellations occurred on the buy and sell sides, respectively, during the span from peak to trough divided by the quantity of cancellation events/orders during the same time span.
- $\ln AverageBuyNew_i$ (equation 3.2.21) and $\ln AverageSellNew_i$ (equation 3.2.22) are the logarithm of the average new buy and sell orders, i.e. total volume of all new orders submitted to the buy and sell sides, respectively, during the span from peak to trough divided by the quantity of new submission events/orders during the same time span.

MODEL SPEC (3): The Effect of Total Volume of Cancelled and New Orders Normalised by the Average Daily Volume (ADV) on The $\Delta NVWAP_i$ from Peak to Trough

The total volumes of cancellations and submissions on both sides of the E-LOB after normalising them to the average daily volume are used in this specification to evaluate the effect of cancellations on the market volatility. The formula for the model applied here is:

$$\Delta NVWAPBuySellPeakToTrough_i = \beta_0 + \beta_1 Duration_i + \beta_2 TotalBuyCancellationToADV_i + \beta_3 TotalBuyNewToADV_i + \beta_4 TotalSellCancellationTOADV_i + \beta_5 TotalSellNewToADV_i + e_i \tag{3.2.3.3}$$

where:

- $\Delta NVWAPBuySellPeakToTrough_i$ is the Bid-Ask difference between the change in the logarithm of slope of the NVWAP curves from peak to trough (equation 3.2.23).
- $Duration_i$ is the time duration in minutes from peak to trough, i.e. the difference between the start time and end time of a downtrend price move (equation 3.2.9).
- $TotalBuyCancellationToADV_i$ (equation 3.2.11) and $TotalSellCancellationToADV_i$ (equation 3.2.12) are the total volume of all cancellations occurred on the buy and sell sides, respectively, during the span from peak to trough, then normalising (dividing) that volume to the average daily volume.
- $TotalBuyNewToADV_i$ (equation 3.2.13) and $TotalSellNewToADV_i$ (equation 3.2.14) are the total volume of all new orders submitted to the buy and sell sides, respectively, during the span from peak to trough, then normalising (dividing) that volume to the average daily volume.

Type of Variable	Variable Name	Model Specification								
		1	2	3	4	5	6	7	8	9
Dependent	$\Delta ReturnPeakToTrough_i$	×								
Dependent	$\Delta RPReturnPeakToTrough_i$		×	×						
Dependent	$\Delta ReturnPeakToTroughPerMinute$				×	×	×			
Dependent	$\Delta NVWAPBuySellPeakToTrough_i$							×	×	×
Independent	$Total\ Buy\ Cancellation/ADV$	×			×					×
Independent	$Total\ Buy\ New/ADV$	×			×					×
Independent	$Total\ Sell\ Cancellation/ADV$	×			×					×
Independent	$Total\ Sell\ New/ADV$	×			×					×
Independent	$\ln Total\ Buy\ Cancellation$		×			×		×		
Independent	$\ln Total\ Buy\ New$		×			×		×		
Independent	$\ln Total\ Sell\ Cancellation$		×			×		×		
Independent	$\ln Total\ Sell\ New$		×			×		×		
Independent	$\ln Average\ Buy\ Cancellation$			×			×		×	
Independent	$\ln Average\ Buy\ New$			×			×		×	
Independent	$\ln Average\ Sell\ Cancellation$			×			×		×	
Independent	$\ln Average\ Sell\ New$			×			×		×	

Table 20: This table describes what variables are used in each model specification to test both hypotheses of this study

4 The Data

In this study, we analyse data on five carefully selected stocks to represent different market sectors. These stocks are traded at LSE where data is derived from London SETS. The data covers the period of two consecutive months, July and August, in 2007; a total of 44 trading days is used. I reconstruct the historic limit order book of London SETS platform over the period of study. I based the analysis on these two months because the month of July was a quiet month where the market was not volatile. However, the market in the month of August was volatile due to the Financial Crisis which started August 2007-9. Eigner and Umlauft (2015: pp. 5) suggest that "The Great Financial Crisis of 2007-2009 was the most severe economic crisis since the Second World War. Only the Great Depression was similar in severity and length."

The following table (Table 21) provides a basic description of the stocks presented and analysed in this study. The impact of daily opening and closing auctions is excluded as we only evaluate return and market events between 8:01 AM and 4:29 PM. This approach filters out both auctions and overnight trading effect on one hand, and removes the impact of corporate actions (e.g. dividends, stock splits, etc.) on the other hand.

Stock	Ticker	Industry Sector
HSBC Holdings	HSBA	Banks
British Petroleum Plc	BP	Petroleum
Tesco Plc	TSCO	Retail
Astra Zeneca Plc	AZN	Pharmaceutical Products
Vodafone Group Plc	VOD	Telecommunications

Table 21: List of stocks used in this study

5 Analysis and Results

This study inspects whether cancellations affect the size of price drops. It, also, studies whether cancellations have an effect on the market volatility, namely the bid-ask spread measure. To this end, this section demonstrates the analyses carried out to answer the hypotheses of this study.

5.1 Preliminary Analysis and Descriptive Statistics

This section describes the statistical properties of the new orders, cancelled orders, spread, change in price and change in stock return for HSBC stock data. These statistics are based on ultra-high-frequency basis covering 2 consecutive months (July and August 2007) before and after the financial crisis.

Stock \ Month	percentage of cancellations to new orders for JULY%		percentage of cancellations to new orders for AUGUST%	
	Buy	Sell	Buy	Sell
HSBA	69.43	70.09	71.64	92.94
BP	71.96	71.55	75.61	75.58
TSCO	75.07	75.47	72.42	72.96
AZN	87.44	89.42	88.18	88.90
VOD	72.86	72.85	79.79	82.69
Average	75.35	75.91	77.39	78.22

Table 22: This table presents the percentage of order cancellations to new order submissions for all peak to trough trends for both July and August 2007 and for all five stocks studied

The statistics (Table 22 above) on the percentages of cancellations to new orders for all five stocks in July and August suggest that around 77% of the new orders submitted to the market are being cancelled. For TSCO stock, the cancellations percentages were as high as 75% of the new orders on both sides of the market in July and August. In July, the cancellations on the sell side of the market were more than their likes on the buy side for AZN[51] (AstraZeneca plc) and HSBA. However, it is the opposite scenario where cancellations were more on the buy side of the market for BP in July. Moreover, cancellations in August indicate that they occurred more on the sell side of the LOB for VOD compared to TSCO, BP, HSBA and AZN where cancellations were almost at the same level for both sides of the order book in that month. Generally, contrary to the large difference in cancellations' percentages between buy and sell sides of the LOB in July, the percentages of cancellations on both sides of the market in August were not significantly different from each other. Cancellations on both sides in August for VOD were 7-10% higher from their counters in July. This might have been due to more volatile market in the month of August 2007.

[51] AstraZeneca plc is an Anglo–Swedish multinational pharmaceutical and biopharmaceutical company.

Stock \ Month	New Buy to New Sell Ratio		Cancelled Buy to Cancelled Sell Ratio	
	July 2007	August 2007	July 2007	August 2007
HSBA	0.8582	0.8725	0.8501	0.8798
BP	0.8922	0.8103	0.8973	0.8105
TSCO	0.8460	0.9080	0.8415	0.9013
AZN	0.8422	1.0035	0.8235	0.9954
VOD	0.8980	0.6789	0.8981	0.6550
Average	0.8668	0.8896	0.8611	0.8819

Table 23: For August and July 2007 and for all five stocks studied, this table presents the Peak-Trough ratios of: (1) new orders submitted to the buy side of the LOB to the new orders submitted to the sell side of the LOB and (2) cancelled orders from the buy side of the LOB to the cancelled orders from the sell side of the LOB.

Moreover, Table 23 above highlights which side of the market (buy or sell) new and cancelled orders have occurred most for the five stocks analysed in this study (see Figure below). It suggests that, in August, NEW orders submitted to the BUY side of the market were higher from their counterparts on the sell side for HSBA and AZN. Nonetheless, in July, the pressure were more on the SELL side where NEW orders submitted to the SELL side of the market were higher for all stocks. Furthermore, cancelled orders from the BUY side of the market were higher from their matches on the sell side for HSBA only in August. Interestingly, new orders (submissions) and cancellations (deletions) density were noticeably larger on the sell side of the market in both months.

The descriptive statistics (Table 24 below) on cancellations and new orders for the HSBC stock suggest that the time of the day (in trading hours) where most cancellations occur was between 09:10-13:25 on both sides of the market in August and between 09:30-13:45 in July, when return changes by less than 50bp[52]. However, these cancellations have occurred between 09:40-13:55, when return changes by at least 50bp but less than 100bp, in August while they

[52] This is noted as low change in Table 24 and measured by basis points.

occurred between 09:40-15:15 in July. When return changes by more than 100bp, these cancellations have occurred between 11:20-15:40 in August compared to between 08:50-13:50 in July. These times indicate that cancellation are most likely to occur almost an hour after trading starts and an hour before trading ends in both months.

In August, the percentage of total cancellations to the new orders on the buy side was almost 84% compared to 73% in July. This percentage was at nearly 70% and 68% on the sell side in August and July respectively. Also, the average volumes of cancelled orders on the buy side in both months noticeably increase with price fluctuations[53]. The same attitude (increase) is observed on the sell side of the market in both months, but is less noticeable when the change in return is less than 100bp. Interestingly, the percentage of cancellations, in August, has dropped from 83% to 72% when return changes by more than 50bp on the buy side. Rather, this percentage has slightly increased on the sell side when return changes by more than 50bp in August. Similarly, the same attitude can be observed in July on both sides in the market.

Likewise, when return changes by less than 50bp, the average volume of new orders submitted to the buy side of the market in August increase as price fluctuates, but this increase was more severe in July, where the increase was almost 3000 shares on average. On the other hand, the average volume of new orders, in July, submitted to the buy side of the market when price change was between 50 and 100bp decreased more in comparison with August. Conversely, this volume has increased more in July than in August when price change exceeded 100bp. Generally, cancellations are higher on the sell side than their counter on the buy side in both months. Yet, cancellations' volume in July was higher on average than their likes in August. Lastly, the spread was consistent in both months regardless of the price change.

[53] Returns have increased by more than 50 basis points.

Month	Variable	Change in Cumulative Return in basis points for HSBC on the Buy side			Change in Cumulative Return in basis points for HSBC on the sell side		
		Low (<50bp)	Medium (<100 and >50bp)	High (>100bp)	Low (<50bp)	Medium (<100 and >50bp)	High (>100bp)
August	Time of The Day	2.216216 (1.108932)	2.71875 (1.113969)	3.384615 (1.192928)	2.216216 (1.113969)	2.71875 (1.113969)	3.384615 (1.192928)
	Average Cancellations per order	7723.272 (1845.368)	8277.699 (2000.066)	9244.942 (2274.748)	8243.519 (2222.337)	8226.206 (1731.032)	9553.733 (1008.321)
	Average New submissions per order	7586.675 (1592.617)	8391.374 (1976.896)	9284.978 (1992.878)	9027.287 (2176.723)	8833.018 (1915.816)	9877.508 (1159.892)
	Total Cancellations	3061501 (4732361)	6445218 (4732361)	1.03e+07 (7060071)	3652428 (3599949)	6940305 (4452188)	1.21e+07 (7765693)
	Total New Orders	4256603 (4652654)	8938266 (6306044)	1.45e+07 (9970520)	5256980 (5396799)	9760131 (6181134)	1.68e+07 (1.07e+07)
	Spread	0.5135135 (0.0821995)	0.578125 (0.184451)	0.5384615 (0.138675)	0.5135135 (0.0821995)	0.578125 (0.184451)	0.5384615 (0.138675)
July	Time of The Day	2.52 (1.159023)	2.85 (1.308877)	2.25 (1.488048)	2.52 (1.159023)	2.85 (1.308877)	2.25 (1.488048)
	Average Cancellations per order	8765.742 (2320.515)	9630.541 (2982.46)	10582.74 (3401.486)	9152.537 (2264.373)	9244.174 (3183.682)	9139.266 (3192.953)
	Average New submissions per order	8383.002 (1697.482)	9454.428 (2188.659)	12171.41 (4335.696)	9872.725 (2299.963)	9644.853 (3123.662)	10852.56 (6094.291)
	Total Cancellations	4830310 (5136356)	6823367 (5760829)	1.31e+07 (8334784)	5613898 (5329542)	8033376 (6125016)	1.55e+07 (1.06e+07)
	Total New Orders	6765385 (6943768)	9836540 (8189457)	1.99e+07 (1.26e+07)	8432472 (7565189)	1.13e+07 (8263501)	2.17e+07 (1.39e+07)
	Spread	0.56 (0.1658312)	0.55 (0.1538968)	0.5 (0)	0.56 (0.1658312)	0.55 (0.1538968)	0.5 (0)

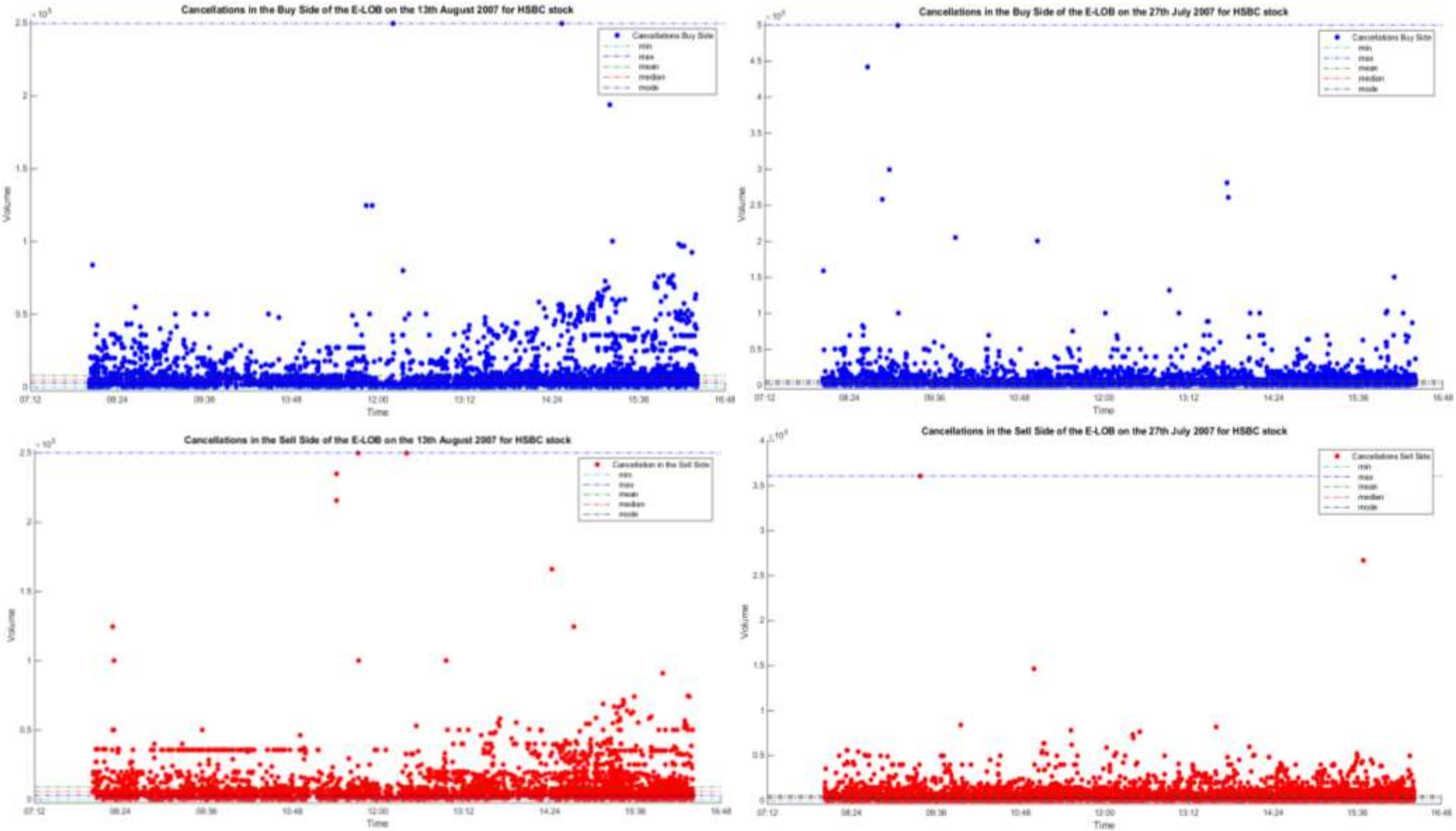
Table 24: Peak – Trough descriptive statistics for the cancelled and new orders for the HSBC stock in August and July 2007 under different return changes (from normal to extreme changes) for both sides of the market where price experiences a downtrend movement is presented in this table. Numbers in brackets are the standard deviations.

On the other hand, Table 25 below illustrates the percentage of cancelled orders to new orders submitted to the market on both sides of the LOB. It is expected, in general, that market price is noticeably affected by large orders' new submission, cancellations or executions.

Remarkably, this table classifies the *cancelled to new orders*' percentages according to the order's size to three categories: firstly, orders below one hundred thousand shares (small), secondly, orders which are between 100 and 300 thousand shares (medium), and lastly, orders which are above or equal 300 thousand shares (large). Notably, in July and August, the share of orders which are 100 thousand or more from the total orders is relatively small, i.e. around 1.5% of the total orders for BP and less than 0.5% for the remaining stocks other than VOD for both buy and sell sides of the LOB. For VOD, the ratios were 14.88% and 17.64% for the buy and sell sides in July respectively. However, these ratio were 19.83% and 31.54% for the buy and sell sides in August respectively.

For instance, the share of medium orders (between 100k-300k) for VOD stock was about 19.24% and 13.53% of the total orders on the buy side and 30.93% and 16.16% on the sell side, for August and July respectively. Yet, the share of large orders (300k or more) for the same stock was 0.59% in August in comparison to 1.35% in July on the buy side while they were 0.61% and 1.48% on the sell side in August and July respectively. This suggests that VOD stock was heavily traded in larger volumes in comparison to the other stocks sampled in this study.

Figure 22: This figure depicts the cancellation pressure in each side of the E-LOB for the stock of HSBC. The left hand side panel shows cancellations on the 13th of August 2007 on both sides, namely buy and sell. Yet, the right hand side panel presents the cancellation pressure on the buy and sell sides on the 27th of July 2007.



Stock \ Ratio		Ratio of cancellations to new for orders below 100k shares (Small)		Ratio of cancellations to new for orders between 100k-300k shares (Medium)		Ratio of cancellations to new for orders above or equal 300k shares (Large)	
		Buy	Sell	Buy	Sell	Buy	Sell
HSBA	July	0.7018	0.7216	0.4972	0.5332	0.9113	0.1199
	August	0.7171	0.7102	0.5507	0.5721	0.3363	0.4319
BP	July	0.7349	0.7492	0.6041	0.6365	0.9067	0.7261
	August	0.7608	0.7703	0.7198	0.6793	0.7868	0.7480
TSCO	July	0.7205	0.7462	0.8598	0.8483	1.0000	1.0000
	August	0.7189	0.7342	0.8729	0.6392	0.7261	0.6757
AZN	July	0.8740	0.8949	0.6470	0.4854	1.0000	0.2556
	August	0.8816	0.8927	0.9348	0.7212	N/A	N/A
VOD	July	0.6901	0.7193	0.7454	0.7740	0.7731	0.6399
	August	0.7185	0.7409	0.8804	0.9073	0.6783	0.5193

Table 25: For July and August 2007 and for all five stocks studied, this table presents the percentages of: (1) cancellations to new for orders below 100k shares (small size orders), (2) cancellations to new for orders between 100k-300k shares (medium size orders), and (3) cancellations to new for orders above or equal 300k shares (large size orders) for all peak to trough periods.

Table 25 above, also, indicates that, in July, cancellation rates were higher for the orders *below 100 thousand shares* and *300 thousand shares or more* than the ones *between 100 and 300 thousand shares* in general for all stocks but TSCO and AZN. Nevertheless, cancellations' ratios were significantly higher on the buy side than their likes on the sell side for HSBA, TSCO, BP and AZN in July 2007. This goes in line with what is expected as per Table 19. This continues to be the case with cancellations ratios on the buy side in August 2007 for BP, TSCO and VOD. For HSBA stock, cancellations' ratio on the sell side was higher than the buy side, which contradicts Table 19 expectations.

In August, the highest cancellation rates were for orders of *300 thousand shares or more* on the buy side for BP, TSCO, AZN and HSBA where cancellations were at 90%+. The least cancellation rates were on the sell side in August for orders of the size *300 thousand shares or more* for HSBA.

Moreover, cancellation rates on both sides in August were higher than their counterparts in July for the medium size orders for all stocks. The case is the opposite for the large size orders where cancellations' rates in July were higher than their counterparts in August for all stocks but HSBA. Interestingly, for medium size orders in August, more than 93% of submitted orders were cancelled on the buy side compared to almost 72% on the sell side of the market in August for AZN. This percentage was more than 88% for VOD on both sides of the market. For all stocks, the cancellation pressure was higher on the buy side for orders of the large size in July. Relatively, this pressure was higher in July on the sell side for the medium size orders for HSBA, BP and VOD while it was higher on the buy side for BP, TSCO and AZN in August.

Month	Time of The Day	Change in Return on the buy side in basis points			Change in Return on the sell side in basis points		
		Low (<50bps)	Medium (<100 and >50bps)	High (>100bps)	Low (<50bps)	Medium (<100 and >50bps)	High (>100bps)
August	1	4917255.6	6800468	16176581	6161962.76	8746708.62	19394737
	2	3263167.8	4569219	12077889	4323374.59	6048706.5	17494975
	3	4412686.4	6927079	N/A	4727427.83	6593423	N/A
	4	8268608.8	9328584	8919486	8732176.04	10109245.7	10151483
July	1	9909460	13790122	17690667.25	11208338.56	15617902	20918604
	2	5170844	5466144	6906405	6959112.54	9275965.5	9575949
	3	4903452	4039595	N/A	6246726.43	5413290	N/A
	4	5961927	6727081	9123497.33	6721606	7370233.4	10128902

Table 26: These descriptive statistics describe cancellations for the HSBC stock in August and July 2007 at different times of the trading day and under different return changes (normal and extreme) for both sides of the market where price experiences a downtrend movement. The change in return is deemed to be low when the change is by less than 50 bp, while it is classified as medium when the change in return is between 50bp and 100 bp. The change in return is classified as high when the change is by 100 bp or more.

The descriptive statistics (Table 26 above) on cancellations for the HSBC stock suggest that under low change in return, most of the cancellations occur at the beginning (between 08:05

and 10:05) and the end of the day (between 14:05 and 16:25) on both sides of the market in August. Yet, the least cancellations on both sides of the market occur between 10:05 and 12:05 in August, when return changes by less than 50bp. The same scenario applies on cancellations on both sides of the market in August under the medium change in return, where the changes at least by 50bp but less than 100bp. However, cancellations on both sides of the market have occurred during the first two quarters of the trading day, i.e. between 08:05-14:05, when return changes by at least 100bp, in August. Interestingly, between 12:05 and 14:05, the return change remained under 100bp on both sides of the market in August and July. This may suggest that the market is a little bit less vulnerable at that time of the trading day.

In July, under the low change in return condition, most cancellations on the buy and sell sides occur at the first two hours of the trading day. The rest of cancellations in the market after the first two hours of the trading day till the end of it add up almost to similar volumes. Likewise, most cancellations on the buy and sell sides occur at the first two hours of the trading day when returns change by more than 50bp but less than 100bp. However, cancellation pressure was significant on the second two hours of the trading day on the sell side compared to the rest of the day under the same price change condition. Cancellation pressure was higher in the last quarter of the trading day than the third quarter though on both sides of the market (buy and sell).

Notably, cancellation volumes increase as the change in return gets larger (more volatile market). This is true for both sides of the market in July and August. This, also, goes in line with the higher volume of orders submitted to both sides of the market as shown in Table 24 below. The highest cancellation volumes on both sides of the market in July occurred under the extreme change in return where return changes by at least 100bp. This may clearly suggest that traders try to actively discover the market price. Moreover, cancellations were substantial in the last two hours of the trading day on the buy side in July, and in the second and last quarters of the trading day on the sell side.

Variable	Time of The Day (Buy side)				Time of The Day (Sell side)			
	1	2	3	4	1	2	3	4
Average Cancellations	7862.346 (2431.876)	6946.275 (1336.248)	7853.118 (1742.298)	9472.292 (1544.927)	8811.936 (2713.465)	7255.155 (1232.025)	8181.823 (1522.501)	9107.161 (1434.019)
Average New Orders	7695.201 (2137.899)	7389.775 (1628.992)	8134.083 (1692.164)	9088.337 (1683.569)	9344.162 (2313.849)	8174.438 (1999.725)	8966.786 (1630.036)	9556.877 (1709.286)
Total Cancellations	4917256 (5596488)	3263168 (2738831)	4412686 (3546827)	8268609 (6115154)	6161963 (6551147)	4323375 (3911165)	4727428 (2477927)	8732176 (6473587)
Total New Orders	7223897 (8122179)	4729437 (4215753)	6064551 (4678651)	1.12e+07 (8491953)	8846098 (9134337)	6058409 (5612751)	6737343 (3733176)	1.21e+07 (8987815)
Spread	0.547619 (0.150396)	0.5 (0)	0.5277778 (0.117851)	0.5769231 (0.183973)	0.547619 (0.150396)	0.5 (0)	0.5277778 (0.117851)	0.5769231 (0.183973)

Table 27: Some descriptive statistics are presented here to describe the cancelled and new orders for the HSBC stock in August, defined by the time of the day they occurred. Time 1 is between 08:05 and 10:05. Time 2 is between 10:05 and 12:05. Time 3 is between 12:05 and 14:05. Time 4 is between 14:05 and 16:25. Numbers in brackets are the standard deviations.

Variable	Time of The Day (Buy side)				Time of The Day (Sell side)			
	1	2	3	4	1	2	3	4
Average Cancellations	10091.51 (3500.017)	8931.52 (3256.226)	7897.177 (836.389)	9511.613 (2074.806)	9269.495 (3000.796)	9983.552 (3753.345)	8256.354 (1317.641)	9119.695 (2211.195)
Average New Orders	10080.23 (3841.429)	8850.96 (2531.757)	8431.577 (1183.057)	9337.631 (1944.334)	10276.55 (4481.851)	10557.43 (3569.954)	8662.25 (1130.917)	9760.809 (2632.168)
Total Cancellations	9909460 (8812000)	5170844 (6568360)	4903452 (2830110)	5961927 (3974786)	1.12e+07 (1.03e+07)	6959113 (6868360)	6246726 (3106856)	6721606 (4767113)
Total New Orders	1.44e+07 (1.30e+07)	7570098 (9214153)	6776737 (3552947)	8544537 (6094496)	1.61e+07 (1.35e+07)	1.03e+07 (9853541)	8461879 (4096968)	9504210 (6661926)
Spread	0.5625 (0.170782)	0.5909091 (0.20226)	0.5 (0)	0.525 (0.111803)	0.5625 (0.170782)	0.5909091 (0.20226)	0.5 (0)	0.525 (0.111803)

Table 28: Some descriptive statistics are presented here to describe the cancelled and new orders for the HSBC stock in July, defined by the time of the day they occurred. Numbers in brackets are the standard deviations.

Overall, Table 27 and Table 28 show that on both sides of the market, the percentage of total volume of the cancelled orders to the total volume of new orders, in August, increases with the time of the day. The increase is more rapid on the sell side though. This suggests that traders are considering their positions in the market before the end of the trading day. Remarkably, cancellations on both sides in the market are at most between 12:05 and 14:05, in July, while they are at their minimum between 10:05 and 12:05. Also, cancellations on the sell side of the market were slightly higher than their rivals on the buy side.

Variable	HSBC August				HSBC July			
	Mean	Std. Dev.	Min	Max	Mean	Std. Dev.	Min	Max
Duration in minutes	51.09756	47.89407	0	340	92.77778	103.3115	5	470
Avg. (Buy) Cancellations	8180.874	2022.902	5121.78	13476.94	9355.989	2753.351	4272.62	17812.48
Tot. (Buy) Cancellation	5526200	5220897	392739	2.37e+07	6833210	6395019	529123	2.87e+07
Avg. (Sell) Cancellations	8444.479	1930.592	5347.53	17893.05	9228.137	2721.033	4901.42	16555.51
Tot. (Sell) Cancellation	6280840	5578040	735969	2.66e+07	8037831	7227596	456667	3.37e+07
Avg. (Buy) New Orders	8169.947	1893.593	4849.38	13935.7	9341.072	2687.662	4889.46	21995.24
Tot.Vol (Buy) New Orders	7713305	7286301	811715	3.25e+07	9842162	9355846	357692	4.17e+07
Tot. Vol (Sell) New Orders	8840440	7792441	1245813	3.45e+07	1.15e+07	9855012	811245	4.45e+07
Avg. (Sell) New Orders	9086.266	1958.891	5575.16	15701.34	9933.489	3325.534	5789.79	25545.82
Spread at best	0.5426829	0.1405725	.5	1	0.5462963	0.1462912	.5	1
Δ price	-1.021341	8.78544	-25.75	32	-5.796296	3.160813	-15.25	-2.5
Δ return	-0.007076	0.005384	-0.028745	-0.002732	-.006432	0.003565	-0.017075	-0.002741

Table 29: The descriptive statistics for all downtrend intervals for the HSBC stock during the months of July and August 2007 for all orders in the order book for both buy and sell sides.

The descriptive statistics on HSBC stock (Table 29) shows that the average duration from peak to trough (downtrend interval) is 51 minutes in August compared to nearly 93 minutes in July. A downtrend interval in August can be as small as one minute or may last up till 340 minutes, which accounts for almost one trading day. In comparison, this interval ranged between 5 to 470 minutes in July. These numbers indicate how volatile the market was in August compared to July 2007, where the price trend changes more frequently in August. Also, the average volumes of cancelled orders on the buy side and sell side were 8181 in August and 9356 in July, respectively, with a higher standard deviation in July. The same attitude can be observed on the sell side as well with 8444 in August compared to 9228 in July. Moreover, the average volume of new orders submitted to the buy side of the market in August was 8170 where it was 9086 on the sell side. On the other hand, the average volume of new orders, in July, submitted to the buy side of the market was 9341 in comparison with 9933 on the sell side.

Table 30 below presents the Peak-Trough descriptive statistics for all variables used to build the model specifications which are used to test the hypotheses of stated earlier. These statistics are reported for both July and August 2007 for HSBC stock.

Variable \ Descriptive Stats	July		August		July		August	
	Mean	Std. Dev.	Min	Max	Mean	Std. Dev.	Min	Max
$\Delta ReturnPeakToTrough_i$	-0.0064	0.0035	-0.1707	-0.0027	-0.0070	0.0053	-0.0287	-0.0027
$\Delta RPReturnPeakToTrough_i$	-64.3231	35.6501	-170.7529	-27.4123	-70.7622	53.8457	-287.4483	-27.3224
$\Delta ReturnPeakToTroughPerMinute$	-1.6370	2.0093	-9.9559	-0.1222	-2.2638	2.7967	-19.7585	-0.0816
$\Delta NVWAPBuySellPeakToTrough_i$	0.5389	0.3918	-0.0531	1.7189	0.5407	0.3667	-0.1563	2.0407
$Total\ Buy\ Cancellation/ADV$	0.1613	0.1509	0.0124	0.6785	0.1304	0.1232	0.0092	0.5587
$Total\ Buy\ New/ADV$	0.2323	0.2208	0.0084	0.9834	0.1820	0.1719	0.0191	0.7669
$Total\ Sell\ Cancellation/ADV$	0.1897	0.1401	0.0107	0.7951	0.1482	0.1316	0.0173	0.6289
$Total\ Sell\ New/ADV$	0.2706	0.2326	0.0191	1.0514	0.2086	0.1839	0.0294	0.8136
$\ln Total\ Buy\ Cancellation$	15.3583	0.9199	13.1789	17.1740	15.1289	0.9080	12.8809	16.9796
$\ln Total\ Buy\ New$	15.7051	0.9616	12.7874	17.5450	15.4696	0.8923	13.6069	17.2964
$\ln Total\ Sell\ Cancellation$	15.5337	0.9163	13.0317	17.3325	15.3112	0.8407	13.5089	17.0981
$\ln Total\ Sell\ New$	15.9225	0.8606	13.6063	17.6119	15.6736	0.7963	14.0353	17.3555
$\ln Average\ Buy\ Cancellation$	9.1016	0.2952	8.8359	9.7876	8.9798	0.2448	8.5412	9.5087
$\ln Average\ Buy\ New$	9.1075	0.2601	8.4948	9.9985	8.9818	0.2321	8.4866	9.5422
$\ln Average\ Sell\ Cancellation$	9.0901	0.2821	8.4972	9.7144	9.0175	0.2170	8.5843	9.7921
$\ln Average\ Sell\ New$	9.1616	0.2780	8.6638	10.1482	9.0929	0.2069	8.6260	9.6615

Table 30: Descriptive statistics for all variables used to define the model specifications to test the hypotheses of this study are reported in this table. All observations are for Peak-Trough trends for HSBC stock in July and August 2007.

Overall, the average volume of the cancelled orders in August on the sell side was larger than its like on the buy side. The case was different in July though where the average volume of the cancelled orders on the buy side was larger than its counterpart on the sell side. Furthermore, the average volume of new orders on the sell side of the market was larger than the one on the buy side in both months. However, the difference between the average volume of new orders between the two sides in August was larger than the difference in July. This suggests that traders try to minimise the price change effect and maximise returns. Besides,

Table 29 suggests that changes in stock price and the stock return in August were more volatile than the ones in July as suggested by the higher standard deviations.

5.2 The effect of Changes in the size of Cumulative Return on Orders' Cancellations Analysis

This section reports the results from running the regressions in model specifications defined in sections 3.2.1 and 3.2.2 to answer the first research question/hypothesis of this study. This study hypothesises that cancellations have an effect on the market, i.e. price dynamics measured by the market impact, to the extent that this phenomenon may lead to market crashes. Mainly, cancellations' effect on the size of price drops is investigated.

In July, Table 31 below shows that under normal cumulative return changes, model specification (2)[54] has performed best in terms of explaining those changes. Nevertheless, the best model which has explained these changes in August has been model specification (3)[55] with R^2 equals to 55%. This study's focus is on the explanations of the extreme[56] changes in cumulative return, as a proxy for market impact, i.e. potential market crash. Hence, cancellations on the buy side as well as the new orders on the buy side are highly significant determinants of the change in cumulative return. The signs of this effect are consistent with the expectations emphasised in Table 19. This suggests that in a downward price move, higher cancellations associated with lower orders' submissions on the buy side would lead to higher market impact. This recommendation, additionally, goes in line with Machain and Dufour (2013) findings. Notably, cancellations effect on the buy side in models (1)[57] and (2) is significant in determining the change in cumulative return under extreme price conditions.

[54] This model is represented in equation (3.2.1.2).
[55] This model is represented in equation (3.2.1.4).
[56] Extreme price conditions definition in (3.2.7) is used unless stated otherwise.
[57] This model is represented in equation (3.2.1.1).

Neither cancellations nor submissions on the sell side significantly affect the change in cumulative return under extreme change in returns in both months, July and August.

Further, cancellations as well as orders' submissions under normal conditions in August are significant determinants of the change in return on the sell side for all three models. The effects on the change in returns match the expected effect in signs. On the other hand, that effect is not significant in July under normal return change under all models but model (3) where orders' submissions effect, only, was significant. These effects' directions, also, match the expectations in demonstrated in Table 19. These findings conform to Eisler et al. (2012) and Mayston et al. (2008) suggestions.

Further, models (5)[58] and (6)[59] suggest that cancellations effect on the sell side on the change in cumulative return under normal price conditions is highly significant in July (p-value<1%) and less significant (p-value<20%) in August. The sign of this effect is opposite to what is expected in Table 19 though. Besides, under model (5), cancellations effect was significant in August on the change in cumulative return on the buy side under normal return change. This effect's direction is as expected in Table 19.

To sum up, it can be concluded that when the market is volatile and the change in cumulative return is not extreme, cancellations on the buy side play significant role in determining the change in cumulative return. Yet, cancellations on the sell side of the market are responsible for the change in cumulative return when the market is not volatile and the change in cumulative return is not extreme. On the other hand, when the market is volatile and the change in cumulative return from peak to trough is extreme, the change in the cumulative return is significantly affected by the cancellations occur on the buy side of the market.

[58] This model is represented in equation (3.2.2.2) under (3.2.10) definition.
[59] This model is represented in equation (3.2.2.3) under (3.2.10) definition.

Cancellations, however, on both sides of the market have no clear effect on the change in cumulative return. It can be argued that, generally, the effect of cancellations on the sell side of the market has a larger effect on the cumulative return than cancellations on the buy side of the market when the change in cumulative return is extreme and the market is not volatile. This effect is not significant though. Therefore, if manipulations are to occur in the market, they are more likely to be on the sell side when the market is not volatile and on the buy side when the market is volatile.

Table 31: This table reports the analysis results of the model specifications which are used to analyse the effect of orders' cancellations on the size of price drops from peak to trough for HSBC stock in August and July 2007 for both buy and sell sides. *** means that the coefficient is significant at 1% while ** means that it is significant at 5%. * means that the coefficient is significant at 10%. These models are:

(1) $\Delta Return\ from\ peak\ to\ trough_i = \beta_0 + \beta_1 TotalBuyCancellationToADV_i + \beta_2 TotalBuyNewToADV_i + \beta_3 TotalSellCancellationTOADV_i + \beta_4 TotalSellNewToADV_i + e_i$,

(2) $\Delta RPReturn\ from\ peak\ to\ trough_i = \beta_0 + \beta_1 \ln TotalBuyCancellation_i + \beta_2 \ln TotalSellCancellation_i + \beta_3 \ln TotalBuyNew_i + \beta_4 \ln TotalSellNew_i + e_i$ and

(3) $\Delta RPReturn\ from\ peak\ to\ trough_i = \beta_0 + \beta_1 \ln AverageBuyCancellation_i + \beta_2 \ln AverageSellCancellation_i + \beta_3 \ln AverageBuyNew_i + \beta_4 \ln AverageSellNew_i + e_i$.

Dependant variable is $\Delta ReturnPeakToTrough_i$												
Model Spec / Variable	Model Spec(1)				Model Spec(2) [60]				Model Spec(3)[61]			
	July		August		July		August		July		August	
	Normal	Extreme	Normal	Extreme	Normal	Extreme	Normal	Extreme	Normal	Extreme	Normal	Extreme
Constant	-0.005***	-0.013	-0.004***	-0.016***	54.429	-123.250	99.334	67.084	110.396	-572.930	-162.928”	1021.814
$\ln Total\ Buy\ Cancellation$					5.851	-85.335	10.951	259.054”				
$\ln Total\ Buy\ New$					-21.243	145.058	-21.077	-198.602				
$\ln Total\ Sell\ Cancellation$					-11.580	33.758	-27.487*	15.923				
$\ln Total\ Sell\ New$					19.829	-95.119	27.428”	-85.032				
$\ln Average\ Buy\ Cancellation$									17.742	-57.404	29.419”	304.901**
$\ln Average\ Buy\ New$									-41.733*	153.579	-35.037*	-396.77**
$\ln Average\ Sell\ Cancellation$									-34.333	71.510	-33.459”	-385.159
$\ln Average\ Sell\ New$									40.068	-121.470	51.094**	344.745
$Total\ Buy\ Cancellation/ADV$	0.021	-0.035	0.015	0.126”								
$Total\ Buy\ New/ADV$	-0.015	0.038	-0.021	-0.091								
$Total\ Sell\ Cancellation/ADV$	-0.014	-0.016	-0.024”	-0.043								
$Total\ Sell\ New/ADV$	0.007	-0.003	0.022”	0.029								
R^2	0.1071	0.3333	0.1657	0.4858	0.1959	0.2270	0.2614	0.4304	0.1457	0.4019	0.1085	0.5471

[60] The dependant variable in this model specification is multiplied by 10000 to get the change in basis points as an integer number, i.e. $\Delta RPReturnPeakToTrough_i$.
[61] The dependant variable in this model specification is multiplied by 10000 to get the change in basis points as an integer number, i.e. $\Delta RPReturnPeakToTrough_i$.

Table 32: This table reports the analysis results of the model specifications which are used to analyse the effect of orders' cancellations on the size of price drops per minute from peak to trough for HSBC stock in August and July 2007 for both buy and sell sides. *** means that the coefficient is significant at 1% while ** means that it is significant at 5%. * means that the coefficient is significant at 10%. These models are:

(4) $\Delta RPReturnPeakToTroughPerMinute_i=\beta_0+\beta_1 TotalBuyCancellationToADV_i+\beta_2 TotalBuyNewToADV_i+\beta_3 TotalSellCancellationTOADV_i+\beta_4 TotalSellNewToADV_i+e_i$,

(5) $\Delta RPReturnPeakToTroughPerMinute_i=\beta_0+\beta_1 \ln TotalBuyCancellation_i+\beta_2 \ln TotalSellCancellation_i+\beta_3 \ln TotalBuyNew_i+\beta_4 \ln TotalSellNew_i+e_i$ and

(6) $\Delta RPReturnPeakToTroughPerMinute_i=\beta_0+\beta_1 \ln AverageBuyCancellation_i+\beta_2 \ln AverageSellCancellation_i+\beta_3 \ln AverageBuyNew_i+\beta_4 \ln AverageSellNew_i+e_i$.

Dependant variable is $\Delta RPReturnPeakToTroughPerMinute_i$

Model Spec / Variable	Model Spec(4)				Model Spec(5)				Model Spec(6)			
	July		August		July		August		July		August	
	Normal	High	Normal	High	Normal	High	Normal	High	Normal	High	Normal	High
Constant	-2.695***	-	-3.167***	-2.211**	-16.181***	-	-20.006**	-17.603*	18.244*	-	38.398**	7.224
$\ln Total\ Buy\ Cancellation$					-1.107	-	3.824"	-3.796				
$\ln Total\ Buy\ New$					-1.062	-	-3.710	3.641				
$\ln Total\ Sell\ Cancellation$					5.647***	-	-3.875"	-4.070				
$\ln Total\ Sell\ New$					-2.485	-	4.881"	5.089				
$\ln Average\ Buy\ Cancellation$									-0.437	-	-0.708	-3.509
$\ln Average\ Buy\ New$									-0.726	-	-2.198	2.730
$\ln Average\ Sell\ Cancellation$									8.747***	-	-2.295	-6.211
$\ln Average\ Sell\ New$									-9.688***	-	0.665	5.932
$Total\ Buy\ Cancellation/ADV$	5.148	-	27.183	-46.889								
$Total\ Buy\ New/ADV$	-18.128"	-	-21.818	50.951								
$Total\ Sell\ Cancellation/ADV$	24.410"	-	-14.546	-20.025								
$Total\ Sell\ New/ADV$	-0.986	-	16.267	3.079								
R^2	0.2297	-	0.0536	0.1935	0.4009	-	0.1086	0.4653	0.3072	-	0.0858	0.2476

5.3 The Effect of Order Cancellations on the Spread Analysis

In this section, the results from running the regressions in model specifications defined in section 3.2.3 are presented. These model specifications try to answer the second research question/hypothesis of this study where we examine the effect of cancellations on the market volatility, i.e. the bid-ask spread., to the extent that this phenomenon may lead to market crashes. In specific, the analysis on how cancellations affect the bid-ask spread is reported in this section.

As Table 33 below best illustrates, model specification (2) has the best performance in explaining the change in the spread from peak to trough. It performed best under both price change conditions, normal and extreme. Furthermore, it can be noticed under model (2) that there is remarkable evidence on a significant relationship between the change in spread from peak to trough and both cancellations and newly submitted orders on the sell side of the market in July under the extreme change in cumulative return. Interestingly, under model specification (1), cancellations and new orders on the sell side were, also, significant determinants of the change in spread at 20% significance level in August where the change in cumulative return was normal (below 100bp); model specification (1) had the worst goodness of fit at almost 70% amongst the three models though. On the other hand, cancellations' effect on the buy side was significant on the change in spread under normal cumulative return change.

Nevertheless, change in spread is not significantly affected by any of the factors included in models (1), (2) and (3) in August under the extreme change in cumulative return. These findings partially conform to the findings of Mayston et al. (2008). The duration of the downtrend price move was a significant determinant of the change in spread in July but not August under both settings of cumulative return change for model (2). Yet, this effect's direction was positive under normal change in cumulative return and negative under the extreme change in cumulative return. The significant relationship between time duration from peak to trough and the spread on the sell side of the market goes in line with Mayston et al. (2008) and Obizhaeva and Wang (2005) findings. However, Table 33 suggests that the spread is positively affected by cancellations on the buy side, which matches the expected sign in Table 19 and supports what Eisler et al. (2012) suggested.

In a nutshell, the analysis suggests that when the market is volatile but the change in cumulative return is not extreme, cancellations on the sell side significantly affect the spread. The effect direction confirms the expectation stated in Table 19. Cancellations, nevertheless, on the buy side of the market justifies the change in spread when the market is not volatile and the change in cumulative return is not extreme. On the other hand, it is not clear how cancellations on both sides of the market affect the spread under the extreme change in cumulative return from peak to trough and the market is volatile. It can be concluded, additionally, that the effect of cancellations on the sell side of the market strongly affect the spread when the change in cumulative price is extreme but the market is not volatile. This effect is of opposite direction of what is expected though.

Table 33: This table reports the analysis results of the model specifications which are used to analyse the effect of orders' cancellations on the size of price drops from peak to trough for HSBC stock in August and July 2007 for both buy and sell sides. *** means that the coefficient is significant at 1% while ** means that it is significant at 5%. * means that the coefficient is significant at 10%. These models are:

(1) $\Delta NVWAPBuySellPeakToTrough_i=\beta_0+\beta_1 Duration_i+\beta_2 \ln TotalBuyCancellation_i+\beta_3 \ln TotalSellCancellation_i+\beta_4 \ln TotalBuyNew_i+\beta_5 \ln TotalSellNew_i+e_i$,

(2) $\Delta NVWAPBuySellPeakToTrough_i=\beta_0+\beta_1 Duration_i+\beta_2 \ln AverageBuyCancellation_i+\beta_3 \ln AverageSellCancellation_i+\beta_4 \ln AverageBuyNew_i+\beta_5 \ln AverageSellNew_i+e_i$ and

(3) $\Delta NVWAPBuySellPeakToTrough_i=\beta_0+\beta_1 Duration_i+\beta_2 TotalBuyCancellationToADV_i+\beta_3 TotalBuyNewToADV_i+\beta_4 TotalBuyCancellationTOADV_i+\beta_5 TotalSellNewToADV_i+e_i$.

Dependant variable is *ΔNVWAPBuySellPeakToTrough*$_i$												
Variable \ Model Spec	Model Spec(1)				Model Spec(2)				Model Spec(3)			
	July		August		July		August		July		August	
	Normal	Extreme	Normal	Extreme	Normal	Extreme	Normal	Extreme	Normal	Extreme	Normal	Extreme
Constant	-0.175	1.925	-1.770”	-2.166	-0.886	-16.378*	-5.402**	0.453	0.431***	1.137”	0.535***	0.697**
Duration	0.001	-0.001	-0.001	-0.006	0.001”	-0.002”	0.001	-0.002	0.000	0.002	-0.001	0.001
$\ln Total\ Buy\ Cancellation$	0.399	-2.849	-0.001	-1.525								
$\ln Total\ Buy\ New$	-0.705*	4.942	-0.020	2.381								
$\ln Total\ Sell\ Cancellation$	-0.333	1.828	-0.573”[62]	1.058								
$\ln Total\ Sell\ New$	0.675	-4.002	0.733”	-1.727								
$\ln Average\ Buy\ Cancellation$					0.681”	-1.134	0.109	-0.858				
$\ln Average\ Buy\ New$					-0.380	4.036”	0.071	1.148				
$\ln Average\ Sell\ Cancellation$					-0.240	3.840*	-0.217	0.479				
$\ln Average\ Sell\ New$					0.083	-4.866*	0.689”	-0.738				
$Total\ Buy\ Cancellation/ADV$									1.606	-12.015”	-1.861	-2.308
$Total\ Buy\ New/ADV$									-2.188	14.647”	1.796	3.997
$Total\ Sell\ Cancellation/ADV$									0.673	3.446	-1.513	-1.694
$Total\ Sell\ New/ADV$									0.658	-10.122	1.029	-1.293
R^2	0.1170	0.6978	0.0674	0.2745	0.1147	0.8982	0.1008	0.0889	0.0408	0.7989	0.0434	0.1024

[62] This Coefficient is nearly significant at 10% (P-value is 0.109).

5.4 Robust Regressions

This section aims at validating the results obtained in sections 5.2 and 5.3 above. Consequently, robust regressions for model specifications discussed earlier are reported in this section. Robust regressions are used to account for any measurement errors in the estimations to allow generalisation of results.

Overall, examining Table 37, Table 38 and Table 39, it can be noted that the outcome of the robust regressions confirm what have been previously reported in section 5.2. Remarkably, the results suggest that the change in cumulative return can be attributed to cancellations on both sides of the market when the market is not volatile regardless of the change size of cumulative return from peak to trough. The effect of cancellations on the sell side on the change in cumulative return is higher though. In addition, in the extreme price conditions, new orders on the sell side also turned to be significant determinant of the size of price drop.

On the other hand, analysing Table 40, also, strengthens the results reported in section 5.3. The aforementioned table adds that under the extreme change in cumulative return, cancellations on the buy side of the market have significant effect on the spread when the market is volatile. This suggests that traders, as profit maximizers, may be able to alter their transaction costs using cancellations on the buy side of the market during price fall periods.

6 Conclusion

The present study aims at investigating and analysing how cancellations on both sides of the order book of London SETS contribute to the market impact (price dynamics) explanation. It, also, aims at investigating the cancellations behaviour in on both sides of the order book and their effect on market volatility (measured by the bid-ask spread). To this end, this paper rebuilds the E-Order Book for London SETS for July and August in 2007, and reconstructs the demand and supply curves to analyse price impact and fluctuations with special focus on cancellations. All analysis is based on ultra-high-frequency data (in micro seconds) for five different stocks from different market sectors in London SETS. This study concentrates on the periods where the market was experiencing a downtrend price movements only (from peak to trough) as microstructure effects are supposed to be similar in the uptrend price movements but in different signs. Peak to trough analysis was introduced and explained in the previous chapter and utilises the DC methodology.

Closely linked to the rise of electronic market microstructure, high frequency finance and algorithmic trading, this study attempts to uncover some of the effects for the observable multitude of order cancellations in E-LOB. Thus, it follows the literature by investigating the variables that contribute to the explanations of the market impact and volatility. However, this work deviates from the literature which uses event-time or calendar-time for sampling to using a new methodology which is Peak-Trough sampling methodology. This research, also, introduces new definitions to the spread as a measure to market volatility. Therefore, it defines and tests different model specifications to tackle the research questions.

This work emphasises that when the market is volatile and the change in cumulative return is not extreme, cancellations on the buy side is a significant determinant of the change in cumulative return. Also, cancellations on the sell side of the market are deemed to be responsible for the change in cumulative return when the market is not volatile and the change in cumulative return is not extreme. On the other hand, when the market is volatile and the change in cumulative return from peak to trough is extreme, the change in the cumulative return is significantly affected by cancellations occurred on the buy side of the market. Yet, when the change in cumulative return is extreme and the market is not volatile, cancellations on both sides of the market have no clear effect on the change in cumulative return. Particularly, the effect of cancellations on the sell side of the market has a greater effect on the cumulative return than cancellations on the buy side of the market when the change in cumulative return is extreme and the market is not volatile. This effect is not significant though.

The analysis, also, proposes that when the market is volatile but the change in cumulative return is not extreme, cancellations on the sell side play a significant role in affecting the spread. Cancellations, nonetheless, on the buy side of the market significantly affect the spread when the market is not volatile and the change in cumulative return is not extreme. Yet, it is still not clear how cancellations on both sides of the market affect the spread under the extreme change in cumulative return from peak to trough and the market is volatile. Furthermore, the effect of cancellations on the sell side of the market strongly affect the spread when the change in cumulative price is extreme but the market is not volatile. This effect is of opposite direction of what is expected though. Generally, the findings go in line with the expected/proposed sign or direction of change in cumulative return. Lastly, cancellations are found to increase on both sides of the market as the trading day approaches its end regardless of the volatility.

Given the limitations of this work, the methods proposed are rather simple and general. The data available doesn't provide traders' IDs, so strategies followed by different traders, including spoofing and layering, cannot be traced or checked. Also, cancellations have been discussed in many studies in the literature as an order book event, but the effect of cancellations is yet not quantified. A future work to extend this study's analysis could be using comparative approaches, e.g. GARCH and error correction models, to analysis cancellations effect on the market. Another extension could be building an agent-based model to enhance the inferences and test different trading scenarios and behaviours.

7 Appendix

Change in Cumulative e Return in real points by duration	July		August	
Percentage	**Percentiles**	**Smallest**	**Percentiles**	**Smallest**
1%	-9.955937	-9.955937	-19.75854	-19.75854
5%	-6.9407	-8.847126	-5.586596	-14.52134
10%	-2.636648	-6.9407	-3.641602	-7.470693
25%	-1.660771	-5.48246	-2.613166	-6.372242
50%	-1.122029	**Largest**	-1.492679	**Largest**
75%	-0.4363247	-0.1801158	-0.874963	-0.4343376
90%	-0.3287552	-0.150734	-0.694445	-0.3977094
95%	-0.150734	-0.1308936	-0.5088683	-0.2152334
99%	-0.1222918	-0.1222918	-0.0816994	-0.0816994
Mean	-1.637076		-2.263851	
Std. Dev.	2.009392		2.796788	
Skewness	-2.711168		-4.252297	
Kurtosis	10.34847		24.36465	

Table 34: The descriptive statistics for the change in cumulative return from peak to trough per minute measured in real points (basis points x 10000)

Duration in Minutes From Peak to Trough	July		August	
Percentage	**Percentiles**	**Smallest**	**Percentiles**	**Smallest**
1%	5	5	0	0
5%	15	5	10	5
10%	20	15	10	5
25%	30	15	20	10
50%	55	**Largest**	40	**Largest**
75%	100	320	70	130
90%	240	375	110	140
95%	375	385	125	155
99%	470	470	340	340
Mean	92.7778		51.0976	
Std. Dev.	103.3115		47.8941	
Skewness	2.0879		3.0040	
Kurtosis	6.6991		17.4453	

Table 35: The descriptive statistics for the duration in minutes for a downtrend price move from peak to trough

Change in Return Per minute Is More Than |3.65|	July		August	
Percentage	**Percentiles**	**Smallest**	**Percentiles**	**Smallest**
1%	0	0	0	0
5%	0	0	0	0
10%	0	0	0	0
25%	0	0	0	0
50%	0	**Largest**	0	**Largest**
75%	0	1	0	1
90%	0	1	1	1
95%	1	1	1	1
99%	1	1	1	1
Mean	0.09259		0.1097	
Std. Dev.	0.2926		0.3145	
Skewness	2.8110		2.4969	
Kurtosis	8.9020		7.2344	

Table 36: The descriptive statistics for the change in cumulative return from peak to trough per minute measured in real points as a dummy variable used to identify extreme changes in returns.

Table 37: This table reports the robust regression results of the model specifications which are used to analyse the effect of orders' cancellations on the size of price drops from peak to trough for HSBC stock in August and July 2007 for both buy and sell sides. *** means that the coefficient is significant at 1% while ** means that it is significant at 5%. * means that the coefficient is significant at 10%. These models are:

(1) ΔReturn from peak to trough$_i$=β_0+β_1TotalBuyCancellationToADV$_i$+β_2TotalBuyNewToADV$_i$+β_3TotalSellCancellationTOADV$_i$+β_4TotalSellNewToADV$_i$+e_i,

(2) ΔReturn from peak to trough$_i$=β_0+β_1 ln TotalBuyCancellation$_i$ +β_2 ln TotalSellCancellation$_i$ +β_3 ln TotalBuyNew$_i$ +β_4 ln TotalSellNew$_i$ +e_i and

(3) ΔRPReturn from peak to trough$_i$=β_0+β_1 ln AverageBuyCancellation$_i$ +β_2 ln AverageSellCancellation$_i$ +β_3 ln AverageBuyNew$_i$ +β_4 ln AverageSellNew$_i$ +e_i.

Dependant variable is $\Delta ReturnPeakToTrough_i$												
Model Spec / Variable	Model Spec(1)				Model Spec(2) [63]				Model Spec(3)[64]			
	July		August		July		August		July		August	
	Normal	Extreme	Normal	Extreme	Normal	Extreme	Normal	Extreme	Normal	Extreme	Normal	Extreme
Constant	-0.005***	-0.013**	-0.004***	-0.016***	54.429	-123.250	99.334*	67.084	110.396	-572.930	-162.928	1021.814
$\ln Total\ Buy\ Cancellation$					5.851	-85.335	10.951	259.054**				
$\ln Total\ Buy\ New$					-21.243”[65]	145.058	-21.077”	-198.602”				
$\ln Total\ Sell\ Cancellation$					-11.580	33.758	-27.487*[66]	15.923				
$\ln Total\ Sell\ New$					19.829	-95.119	27.428”	-85.032				
$\ln Average\ Buy\ Cancellation$									17.742	-57.404	29.419”[67]	304.901**
$\ln Average\ Buy\ New$									-41.733**	153.579	-35.037*	-396.770**
$\ln Average\ Sell\ Cancellation$									-34.333”	71.510	-33.459*	-385.159
$\ln Average\ Sell\ New$									40.068”	-121.470	51.094**	344.745
$Total\ Buy\ Cancellation/ADV$	0.021”	-0.035	0.015	0.126*								
$Total\ Buy\ New/ADV$	-0.015”	0.038	-0.021	-0.091”								
$Total\ Sell\ Cancellation/ADV$	-0.014	-0.016	-0.024*	-0.043								
$Total\ Sell\ New/ADV$	0.007	-0.003	0.022”[68]	0.029								
R^2	0.1071	0.3333	0.1657	0.4858	0.1959	0.2270	0.2614	0.4304	0.1457	0.4019	0.1085	0.5471

[63] The dependant variable in this model specification is multiplied by 10000 to get the change in basis points as an integer number.
[64] The dependant variable in this model specification is multiplied by 10000 to get the change in basis points as an integer number.
[65] This coefficient is almost significant at 10% significance level (p-value is 0.114).
[66] This coefficient is almost significant at 5% significance level (p-value is 0.057).
[67] This coefficient is almost significant at 10% significance level (p-value is 0.103).
[68] This coefficient is almost significant at 10% significance level (p-value is 0.104).

Table 38: This table reports the robust regression results of the model specifications which are used to analyse the effect of orders' cancellations on the size of price drops per minute from peak to trough for HSBC stock in August and July 2007 for both buy and sell sides. * means that the coefficient is significant at 1% while ** means that it is significant at 5%. * means that the coefficient is significant at 10%.These models are:**

(4) $\Delta RPReturnPeakToTroughPerMinute_i=\beta_0+\beta_1 TotalBuyCancellationToADV_i+\beta_2 TotalBuyNewToADV_i+\beta_3 TotalSellCancellationTOADV_i+\beta_4 TotalSellNewToADV_i+e_i$,

(5) $\Delta RPReturnPeakToTroughPerMinute_i=\beta_0+\beta_1 \ln TotalBuyCancellation_i+\beta_2 \ln TotalSellCancellation_i+\beta_3 \ln TotalBuyNew_i+\beta_4 \ln TotalSellNew_i+e_i$ **and**

(6) $\Delta RPReturnPeakToTroughPerMinute_i=\beta_0+\beta_1 \ln AverageBuyCancellation_i+\beta_2 \ln AverageSellCancellation_i+\beta_3 \ln AverageBuyNew_i+\beta_4 \ln AverageSellNew_i+e_i$.

Dependant variable is $\Delta RPReturnPeakToTroughToDuration_i$												
Model Spec / Variable	Model Spec(4)				Model Spec(5)				Model Spec(6)			
	July		August		July		August		July		August	
	Normal	Extreme	Normal	Extreme	Normal	Extreme	Normal	Extreme	Normal	Extreme	Normal	Extreme
Constant	-2.258***	-7.286	-2.354***	-13.152***	-14.502***	-59.541**	-15.236***	-108.077***	5.315	2.777	19.669***	80.170
ln Total Buy Cancellation					-1.200*[69]	4.240	2.459**	13.812				
ln Total Buy New					0.817	-8.818	-2.208"	-10.136				
ln Total Sell Cancellation					4.135*[70]	6.491"	-1.154	-18.329				
ln Total Sell New					-2.851"	1.715	1.801"	20.880				
ln Average Buy Cancellation									-2.624*	6.243	-0.179	-8.179
ln Average Buy New									1.470	5.546	-1.710	8.762
ln Average Sell Cancellation									7.740*	0.773	0.026	-20.661
ln Average Sell New									-7.273"	-13.288"	-0.521	10.692
Total Buy Cancellation/ADV	2.737	2.066	10.965	80.256								
Total Buy New/ADV	-9.087"	-7.330	-9.231	-84.512								
Total Sell Cancellation/ADV	11.475"	1.139	-1.119	-58.299								
Total Sell New/ADV	1.734	13.179	5.443	85.732								
R^2	0.2060	0.6364	0.1421	0.5860	0.5115	0.9206	0.2968	0.8303	0.3106	0.6073	0.1224	0.8022

[69] This coefficient is almost significant at 5% significance level (p-value is 0.056).
[70] This coefficient is almost significant at 5% significance level (p-value is 0.063).

Table 39: This table reports the robust regression results of the model specifications which are used to analyse the effect of orders' cancellations on the size of price drops per minute from peak to trough for HSBC stock in August and July 2007 for both buy and sell sides. * means that the coefficient is significant at 1% while ** means that it is significant at 5%. * means that the coefficient is significant at 10%. These models are:**

(7) $\Delta RPReturnPeakToTroughPerMinute_i=\beta_0+\beta_1 TotalBuyCancellationToADV_i+\beta_2 TotalBuyNewToADV_i+\beta_3 TotalSellCancellationTOADV_i+\beta_4 TotalSellNewToADV_i+e_i$,

(8) $\Delta RPReturnPeakToTroughPerMinute_i=\beta_0+\beta_1 \ln TotalBuyCancellation_i+\beta_2 \ln TotalSellCancellation_i+\beta_3 \ln TotalBuyNew_i+\beta_4 \ln TotalSellNew_i+e_i$ and

(9) $\Delta RPReturnPeakToTroughPerMinute_i=\beta_0+\beta_1 \ln AverageBuyCancellation_i+\beta_2 \ln AverageSellCancellation_i+\beta_3 \ln AverageBuyNew_i+\beta_4 \ln AverageSellNew_i+e_i$.

Dependant variable is $\Delta RPReturnPeakToTroughToDuration_i$												
Model Spec / Variable	Model Spec(7)				Model Spec(8)				Model Spec(9)			
	July		August		July		August		July		August	
	Normal	High	Normal	High	Normal	High	Normal	High	Normal	High	Normal	High
Constant	-2.695***	-	-3.167***	-2.211**	-16.181**	-	-20.006**	-17.603**	18.244*	-	38.398**	7.224
ln Total Buy Cancellation					-1.107	-	3.824	-3.796				
ln Total Buy New					-1.062	-	-3.710	3.641				
ln Total Sell Cancellation					5.647**	-	-3.875"	-4.070				
ln Total Sell New					-2.485	-	4.881"	5.089				
ln Average Buy Cancellation									-0.437	-	-0.708	-3.509
ln Average Buy New									-0.726	-	-2.198	2.730
ln Average Sell Cancellation									8.747***	-	-2.295	-6.211
ln Average Sell New									-9.688***	-	0.665	5.932
Total Buy Cancellation/ADV	5.148	-	27.183"	-46.889								
Total Buy New/ADV	-18.128***	-	-21.818	50.951								
Total Sell Cancellation/ADV	24.410*	-	-14.546	-20.025								
Total Sell New/ADV	-0.986	-	16.267	3.079								
R^2	0.2297	-	0.0536	0.1935	0.4009	-	0.1086	0.4653	0.3072	-	0.0858	0.2476

Table 40: This table reports the robust regression results of the model specifications which are used to analyse the effect of orders' cancellations on the size of price drops from peak to trough for HSBC stock in August and July 2007 for both buy and sell sides. *** means that the coefficient is significant at 1% while ** means that it is significant at 5%. * means that the coefficient is significant at 10%. These models are:

(1) $\Delta NVWAPBuySellPeakToTrough_i=\beta_0+\beta_1 Duration_i+\beta_2 \ln TotalBuyCancellation_i+\beta_3 \ln TotalSellCancellation_i+\beta_4 \ln TotalBuyNew_i+\beta_5 \ln TotalSellNew_i+e_i$,

(2) $\Delta NVWAPBuySellPeakToTrough_i=\beta_0+\beta_1 Duration_i+\beta_2 \ln AverageBuyCancellation_i+\beta_3 \ln AverageSellCancellation_i+\beta_4 \ln AverageBuyNew_i+\beta_5 \ln AverageSellNew_i+e_i$ and

(3) $\Delta NVWAPBuySellPeakToTrough_i=\beta_0+\beta_1 Duration_i+\beta_2 TotalBuyCancellationToADV_i+\beta_3 TotalBuyNewToADV_i+\beta_4 TotalBuyCancellationTOADV_i+\beta_5 TotalSellNewToADV_i+e_i$.

Dependant variable is $\Delta NVWAPBuySellPeakToTrough_i$												
Model Spec / Variable	Model Spec(1)				Model Spec(2)				Model Spec(3)			
	July		August		July		August		July		August	
	Normal	Extreme	Normal	Extreme	Normal	Extreme	Normal	Extreme	Normal	Extreme	Normal	Extreme
Constant	-0.175	1.925	-1.770”	-2.166	-0.886	-16.378*	-5.402**	0.453	0.431***	1.137”	0.535***	0.697**
Duration	0.001	-0.001	-0.001”	-0.006”	0.001**	-0.002*	0.001[71]	-0.002	0.000	0.002	-0.001	0.001
$\ln Total\ Buy\ Cancellation$	0.399”	-2.849*	-0.001	-1.525”								
$\ln Total\ Buy\ New$	-0.705**	4.942*	-0.020	2.381”								
$\ln Total\ Sell\ Cancellation$	-0.333	1.828*	-0.573”	1.058								
$\ln Total\ Sell\ New$	0.675	-4.002*	0.733*	-1.727								
$\ln Average\ Buy\ Cancellation$					0.681*[72]	-1.134	0.109	-0.858				
$\ln Average\ Buy\ New$					-0.380	4.036”	0.071	1.148				
$\ln Average\ Sell\ Cancellation$					-0.240	3.840”[73]	-0.217	0.479				
$\ln Average\ Sell\ New$					0.083	-4.866*[74]	0.689*	-0.738				
$Total\ Buy\ Cancellation/ADV$									1.606	-12.015”	-1.861	-2.308
$Total\ Buy\ New/ADV$									-2.188	14.647”	1.796	3.997
$Total\ Sell\ Cancellation/ADV$									0.673	3.446	-1.513	-1.694
$Total\ Sell\ New/ADV$									0.658	-10.122”	1.029	-1.293
R^2	0.1170	0.6978	0.2989	0.2745	0.1147	0.8982	0.1008	0.0889	0.0408	0.7989	0.0434	0.1024

[71] This coefficient is almost significant at 20% significance level (p-value is 0.203).
[72] This coefficient is almost significant at 5% significance level (p-value is 0.057).
[73] This coefficient is almost significant at 10% significance level (p-value is 0.103).
[74] This coefficient is almost significant at 5% significance level (p-value is 0.061).

5 Chapter Five

Conclusion

The increase in trading volume has raised many concerns about the impact of algorithmic trading, which links to HFT, on the market, i.e. price discovery process and volatility. There has been considerable effort on questioning market quality and HFT in both theoretical and empirical literature. Initially, HFTs and other intermediaries played a positive role in stabilizing prices by providing liquidity. However, intermediaries step aside in many historical cases at times of extreme volatility. It, consequently, does not appear to be a problem with HFT but rather a general feature of equity markets. Most of the research which has investigated HFT has concentrated on the market impact of HFT on market quality measures. Besides, another bulk of studies has inspected the suitability of the current market regulations in light of HFT.

However, there is still a gap between the results of academic literature and the beliefs predominant or sustained in media, public and even regulatory discussions on HFT impact on markets. Additionally, it is important to highlight that evaluating the strategies used by HFTs is of overriding significance. Yet, more attention should be paid to investigating automated trading strategies. Introducing the European Commission FTT has raised many concerns about examining the impact of this tax on traders (HFTs in particular) and financial markets. The first chapter of this thesis has attempted to empirically examine the impact of levying FTT on the stock market by simulating an electronic limit order market (order-driven market) with a heterogeneous set of players/traders. This set of players includes traders who are liquidity takers and providers where they buy and sell a single asset in the market. This artificial market is populated exclusively with zero-intelligence agents who randomly place orders. Yet, the specifications of the limit order book in this artificial market has led to generate a non-random price. The novelty of the first chapter lies in the use of simulated order flow data to assess the impact of the FTT on the market quality as well as on the market participants. I found support to the regulatory policy of the European Commission regarding introducing a FTT of 0.1% on all stock transactions. In particular, traders are found to tend to submit orders and start to trade more aggressively by increasing the volume traded in order to remain in a profitable position after the tax. The market, nevertheless, ended up with higher trading volume, lower bid-ask spread and almost same price volatility. This may suggest that the proposed tax does not affect HFTs in a negative way. On the contrary, it induced them to perform more trading volume.

The second and third chapters rebuild the London Stock Exchange Electronic Order Book (SETS) in real-time for 5 different stocks for two consecutive months, July and August 2007 and utilise the Directional Changes (DC) methodology to track price trends. DC tracks peak to trough (P-T) and trough to peak (T-P) trends that have cumulative returns in excess of 25 basis points thresholds.

The second chapter of this thesis aims at analysing the intraday dynamics of liquidity in London SETS electronic limit order book. Utilising the concept of the NVWAP, the shape of bid and ask sides of the order book is modelled to predict the market trend under different time intervals. The prediction methodology consists of four statistics which are $\Delta log(slope^{ask})$, $\Delta log\left(slope^{bid}\right)$, $\Delta log\left(Q^{bid}\right)$ and $\Delta log(Q^{ask})$. These statistics represent the steepening/flattening and the expansion/contraction in the NVWAP curves. The empirical analysis confirms the that information from the NVWAP along with the resulted liquidity supply and demand curves in the limit order book reveal consistent observable market behaviour during both uptrend and downtrend intervals. In fact, examining the change in the shape of NVWAP curves for five different stocks over a period of two consecutive months, July and August 2007, suggests that the aforementioned four statistics have correctly identified prevailing market trend in 86.61% of the total cases on average. Thus, these four statistics are revealed to be robust measures to identify the prevailing market trend without prior knowledge of the price. These results go in line with Malik and Markose (2012) findings for different set of stocks. Within an algorithmic trading system, as market state changes, these four statistics can be used to switch between different trading strategies or build new strategies.

The third chapter of this thesis investigates how cancellations on both sides of the order book of London SETS contribute to the market impact (price dynamics) explanation. It, also, analyses the cancellations behaviour on both sides of the order book and their effect on market volatility (denoted by the bid-ask spread). To this end, I rebuild the E-Order Book for London SETS for July and August in 2007, and reconstruct the demand and supply curves to analyse price impact and fluctuations with special focus on cancellations. The main focus of the third chapter will be on the periods where the market was experiencing a downtrend price movements only (from peak to trough) as microstructure effects are supposed to be similar in the uptrend price movements but in different signs. This chapter tries to uncover some of the effects for the observable multitude of order cancellations on the market quality. It follows the literature by investigating the variables that contribute to the explanations of the market impact and volatility. However, this study deviates from the literature which uses event-time or calendar-time for sampling to using a new methodology which is Peak-Trough sampling methodology inspired by DC approach. This study, also, introduces new definitions to the spread as a measure of market volatility. Therefore, this paper defines and tests different model specifications to tackle the research questions.

Third chapter suggests that when the market is volatile and the change in cumulative return is not extreme, cancellations on the buy side is a significant determinant of the change in cumulative return. Also, cancellations on the sell side of the market are deemed to be responsible for the change in cumulative return when the market is not volatile and the change in cumulative return is not extreme. On the other hand, when the market is volatile and the change in cumulative price return from peak to trough is extreme, the change in the cumulative return is significantly affected by cancellations occurred on the buy side of the market. Put differently, this study, also, proposes that when the market is volatile but the change in cumulative return is not extreme, cancellations on the sell side play a significant role in affecting the spread. Cancellations, nonetheless, on the buy side of the market significantly affect the spread when the market is not volatile and the change in cumulative return is not extreme. Moreover, it is still not clear how cancellations on both sides of the market affect the spread under the extreme change in cumulative price return from peak to trough and the market is volatile. This study, also, suggests that the effect of cancellations on the sell side of the market strongly affect the spread when the change in cumulative price is extreme but the market is not volatile. This effect is of opposite direction of what is expected though.

Given the limitations of this work, the methods proposed are rather simple and general. The data available doesn't provide traders' IDs, so strategies followed by different traders, including spoofing and layering, cannot be checked. Also, cancellations have been discussed in many studies in the literature as order book events, but the effect of cancellations is yet not quantified. A future work to extend this study's analysis could be using comparative approaches to analysis cancellations effect on the market. Another possible extension is to build an agent-based model to enhance the inferences and test different trading strategies/scenarios and behaviours.

6 References

Aldridge, I. (2013). "High Frequency Trading", 2nd edition Wiley.

Almgren, R., Thum, C., Hauptmann, E. and Li, H. (2005). "Direct estimation of equity market impact", Risk, 18 (No. 7), pp. 5862.

Andrews, D. (1974). "A robust method for multiple linear regression", Technometrics, Vol. 16(4), pp. 523-531.

Aquilina, M., Ibikunle, G., Mollica, V. and Steffen, T. (2017). "Benchmark Regulation and Market Quality", (FCA Occasional Paper No. 27) (pp. 1–52). Financial Conduct Authority. Retrieved from https://www.fca.org.uk/publication/occasional-papers/op17-27.pdf.

Baltagi, B., Li, D. and Li, Q (2006). "Transaction tax and stock market behavior: evidence from an emerging market", Empirical Economics, 31(2), 393-408.

Bansal, A., Mishra, K. and Pachouri, A. (2010). "Algorithmic trading (AT)-framework for futuristic intelligent human interaction with small investors". International Journal of Computer Applications, 1(21), pp.01-05.

Baruch, S. and Glosten, L. R. (2013). "Fleeting Orders", working paper.

Berkowitz, D., Logue, D. and Noser, E.A. (1988) The Total Cost of Transactions on the NYSE. The Journal of Finance, American Finance Association, Vol. 43, No. 1, 97-112, doi:10.1111/j.1540-6261.1988.tb02591.x.

Bijlsma, M., Elbourne, A., Lever, M. and Zwart, G. (2012). "Financial transaction tax: review and assessment", Discussion paper 202, CPB Netherlands Bureau for Economic Policy Analysis.

Bouchaud, J. P. (2010). "Price impact", Encyclopedia of quantitative finance.

Bouchaud, J. P., Gefen, Y., Potters, M. and Wyart, M. (2004). "Fluctuations and response in financial markets: the subtle nature of 'random' price changes", Quantitative finance, 4 (No.2), pp. 176-190.

Bouchaud, J. P., Mézard, M. and Potters, M. (2002). "Statistical properties of stock order books: empirical results and models", Quantitative finance, Vol. 2(4), pp. 251-256.

Brogaard, J. (2010). "High frequency trading and its impact on market quality", Northwestern University Kellogg School of Management, Working Paper.

Brownlees, C. T., Cipollini, F., and Gallo, G. M. (2011). "Intra-daily volume modeling and prediction for algorithmic trading". Journal of Financial Econometrics, 9(3); 489-518.

Campbell, J., and Froot, K. (1994), "International experiences with securities transaction taxes", In The internationalization of equity markets, National Bureau of Economic Research (pp. 277-308), University of Chicago Press.

Cao, C., Hansch, O. and Wang, X (2003). "The Information Content of an Open Limit-Order Book". Journal of Future Markets, Vol. 29, No. 1, 16-41.

Carroll, S. M. (2012). "HIGH FREQUENCY TRADING: A SURVEY OF MARKET IMPACT", Doctoral dissertation, The University of Utah.

Chaboud, A., Hjalmarsson, E., Vega, C., and Chiquoine, B. (2009). "Rise of the machines: Algorithmic trading in the foreign exchange market", FRB International Finance Discussion Paper, (980).

Chlistalla, M., Speyer, B., Kaiser, S., and Mayer, T. (2011). "High frequency trading", Deutsche Bank Research.

Chordia, T., Roll, R. and Subrahmanyam, A. (2008). "Why has trading volume increased", Working Paper, Anderson School, UCLA.

Cliff, D., Brown, D., and Treleaven, P. (2011). "Technology trends in the financial markets: A 2020 vision", Foreshight Driver Review, DR3, preprint, 22(03).

Colliard, J. and Foucault, T. (2012). "Trading fees and efficiency in limit order markets", Review of Financial Studies, 25(11), 3389-3421.

Cont, R., Kukanov, A. and Stoikov, S. (2014). "The price impact of order book events", Journal of financial econometrics, 12(1), pp. 47-88.

Cvitanic, J., and Kirilenko, A. (2010). "High Frequency Traders and Asset Prices", Available at SSRN 1569075.

Daniel, G., and Cappellini, A. (2006). "Extracting exploitable information from a simulated limit order book", Working Paper.

Davis, M., Kumiega, A. and Van Vliet, B. (2013). "Ethics, finance, and automation: a preliminary survey of problems in high frequency trading", Science and engineering ethics, Vol. 19(3), pp. 851-874.

Degryse, H., De Jong, F., Van Ravenswaaij, M., and Wuyts, G. (2005). "Aggressive orders and the resiliency of a limit order market". Review of Finance, 9(2), 201-242.

Domowitz, I., and Yegerman, H. (2005). "The Cost of Algorithmic Trading", available online at: http://www.globalposit.org/news_events/papers/AlgorithmicTrading_2.24.2005.pdf.

Easley, D., López de Prado, M., and O'Hara, M. (2012). "The volume clock: Insights into the high frequency paradigm", Available at SSRN 2034858.

Eigner, P. and Umlauft, T. (2015). "The Great Depression(s) of 1929-1933 and 2007_2009? Parallels, Differences and Policy Lessons", Hungarian Academy of Science MTA-ELTE Crisis History Working Paper No. 2. Available at SSRN-id2612243.

Eisler, Z., Bouchaud, J. P. and Kockelkoren, J. (2011). "Models for the impact of all order book events".

Eisler, Z., Bouchaud, J. P. and Kockelkoren, J. (2012). "The price impact of order book events: market orders, limit orders and cancellations", Quantitative Finance, 12 (No. 9), pp. 1395-1419.

Engle, R. and Lunde, A. (2003). "Trades and Quotes: A Bivariant Point Process", Journal of Financial Econometrics, 11, pp. 159–188.

Engle, R., Ferstenberg, R. and Russell, J. (2012). "Measuring and modelling execution cost and risk", The Journal of Portfolio Management, 38(2), 14-28.

European Commission (2011). "European Commission Executive summary of The Impact Assessment", archived from the original (PDF) on 15 November 2011. Retrieved 26 February 2012. Available online at: https://web.archive.org/web/20111115010948/http:/ec.europa.eu/taxation_customs/resources/documents/taxation/other_taxes/financial_sector/summ_impact_assesmt_en.pdf.

European Commission official website, http://europa.eu/legislation_summaries/internal_market/single_market_services/financial_services_general_framework/mi0087_en.htm.

Fabozzi, F.; Focardi, S. and Jonas; C. (2011). "High-Frequency Trading: Methodologies and market impact". Review of Futures Markets, 19, 7-38.

Farmer, J., Gerig, A., Lillo, F. and Waelbroeck, H. (2013). "How efficiency shapes market impact". Quantitative Finance Vol. 13, no. 11, pp. 1743-1758.

Farmer, J., Gillemot, L., Lillo, F., Mike, S. and Sen, A. (2004). "What really causes large price changes?". Quantitative finance 4, No. 4 (2004), pp. 383-397.

Finer, B. M. H., Dull, T. J., Qin, L., Farson, D. and Roberts, M. R. (2016). "The Flash Crash a New Deconstruction", Draft, no. 1.

Fong, K. and Liu, W. (2010). "Limit order revisions". Journal of Banking & Finance, Vol. 34, No. 8, pp. 1873-1885.

Foucault, T. (1999). "Order flow composition and trading costs in a dynamic limit order market", Journal of Financial markets, 2(2), 99-134.

Foucault, T. (2016). "Where are the risks in high frequency trading?". FSR FINANCIAL, p.53.

Frei, C., and Westray, N. (2015). "Optimal execution of a VWAP order: a stochastic control approach". Mathematical Finance, 25(3), 612-639.

Gabaix, X., Gopikrishnan, P., Plerou, V. and Stanley, H. (2003). "A theory of power-law distributions in financial market fluctuations", Nature 423, No. 6937, pp. 267-270.

Gatheral, J. (2010). "No-dynamic-arbitrage and market impact", Quantitative finance, 10 (No. 7), pp. 749-759.

Gomber, P., Arndt, B., Lutat, M., and Uhle, T. (2011). "High frequency trading", Working paper, Available at SSRN 1858626.

Groth, S. (2011). "Does Algorithmic Trading Increase Volatility? Empirical Evidence from the Fully-Electronic Trading Platform Xetra", Working Paper.

Grover, V., Ramanlal, P. and Segars, A. (1999). "Electronic Commerce and Market Transformation", International Journal of Electronic Commerce, Vol. 3, No. 4, pp. 89-102.

Guo, X. (2012). "High frequency Trading", University of California at Berkeley, Technical Report No. UCB/EECS-2012-130.

Habermeier, K., and Kirilenko, A (2003). "Securities transaction taxes and financial markets", IMF Staff Papers, 165-180.

Hagströmer, B. and Nordén, L. (2012). "The diversity of high frequency traders", School of Business at Stockholm University, Available at SSRN 2153272.

Hanson, T. (2011). "The Effects of High Frequency Traders in a Simulated Market", In Midwest Finance Association 2012 Annual Meetings Paper, Available at SSRN 1918570.

Harris, L.E. and Panchapagesan (2005). "The Information Content of The Limit Order Book: Evidence from NYSE specialist". Journal of Financial Markets, 8, 25-67.

Hasbrouck, J. (1991). "Measuring the information content of stock trades", Journal of Finance Vol. 46, pp. 179-208.

Hasbrouck, J. and Saar, G. (2009). "Technology and liquidity provision: The blurring of traditional definitions", Journal of financial Markets, 12(2), pp. 143-172.

Hasbrouck, J. and Saar, G. (2011). “Low-latency trading”, Johnson school research paper series, (35-2010), Available at SSRN 1695460.

Hasbrouck, J., Sofianos, G., Sosebee, D. and Director, F (1993). “New York stock exchange systems and trading procedures”, NYSE Working paper #93-01.

Hautsch, N. and Huang, R. (2012). “The market impact of a limit order”. Journal of Economic Dynamics and Control 36, no. 4, pp. 501-522.

Hendershott, T. and Riordan, R. (2009). “Algorithmic Trading and Information”, NET Institute. Working Paper No. 09-08.

Hendershott, T., and Riordan, R. (2011). “High frequency trading and price discovery”, Manuscript, University of California, Berkeley, 3.

Hendershott, T., Jones, C. M., and Menkveld, A. J. (2011). “Does algorithmic trading improve liquidity?”. The Journal of Finance, 66(1), 1-33.

Höschler, M. (2011). “Limit order book models and optimal trading strategies”, PhD thesis.

Huang, R. and Hautsch, N. (2011). “Limit Order Flow, Market Impact and Optimal Order Sizes: Evidence from NASDAQ Total View - ITCH Data”, discussing paper, vol. 056.

Huberman, G. and Stanzl, W. (2004). “Price Manipulation and Quasi-Arbitrage”, Econometrica, 72(4), pp. 1247-1275.

Jones, C. (2013). “What Do We Know About High frequency Trading?”, (March 20, 2013), Columbia Business School Research Paper No. 13-11, Available at SSRN 2236201.

Kane, D., A. Liu and K., Nguyen (2011). “Analysing an Electronic Limit Order Book”. The R Journal Vol. 3/1, 64-68.

Kato, T. (2015). “VWAP execution as an optimal strategy”. Long version of the paper" VWAP execution as an optimal strategy" in JSIAM Letters, 7, 33-36.

Keim, D. and Madhavan, A. (1991). “The upstairs market for large-block transactions: Analysis and measurement of price effects”, working paper, University of Pennsylvania.

Kim, A. (2008). "An order flow model and a liquidity measure of financial markets", Doctoral dissertation, Massachusetts Institute of Technology.

Kirilenko, A., Kyle, A., Samadi, M. and Tuzun, T. (2011). "The Flash Crash: The impact of high frequency trading on an electronic market", Available at SSRN 1686004.

Kozhan, R. and Salmon, M. (2012). "The information content of a limit order book: the case of an FX market". Journal of Financial Markets, Vol. 15 (No.1). pp. 1-28. ISSN 1386-4181.

Kyle, A. S. (1985). "Continuous auctions and insider trading", Econometrica: Journal of the Econometric Society, pp. 1315-1335.

Kyle, A. S. and Viswanathan, S. (2008). "How to define illegal price manipulation", American Economic Review, Vol. 98(2), pp. 274-79.

Leal, S., Napoletano, M., Roventini, A. and Fagiolo, G. (2016). "Rock around the clock: an agent-based model of low-and high-frequency trading". Journal of Evolutionary Economics, Vol. 26(1), pp. 49-76.

Leinweber, D. (2009). "Nerds on Wall Street: Math, Machines and Wired Markets", John Wiley & Sons.

Leis, D. (2012). "High Frequency Trading Market Manipulation and Systemic Risks from an EU Perspective", Available at SSRN 2108344.

Lhabitant, F. S. and Gregoriou, G. N. (2015). "High-Frequency Trading: Past, Present, and Future", Handbook of High Frequency Trading, 1st ed., G. N. Gregoriou, Ed. Elsevier Inc., pp. 155-166.

Lo, D. and Hall, A. (2015). "Resiliency of the limit order book", Journal of Economic Dynamics and Control, Vol. 61, pp. 222-244.

London Stock Exchange (LSE) website, http://www.londonstockexchange.com/products-and-services.

Machain, L. and Dufour, A. (2013). “The Price Impact of Limit Order Cancellations”, Available at SSRN.

Madhavan, A. (2000). “Market microstructure: A survey”, Journal of financial markets 3, no. 3: 205-258.

Malik, A. and Markose S. (2012). “Price Trends and Liquidity Dynamics in London SETS Limit Order Book: An Application of Real Time Rebuild of SETS”, working paper.

Malinova, K. and Park, A. (2011). “Subsidizing liquidity: the impact of make / take fees on market quality”, mimeo, University of Toronto.

Mayston, D., Kempf, A. and Yadav, P. (2008). “Resiliency in Limit Order Book Markets: A Dynamic View of Liquidity”. AFA 2009 San Francisco Meetings Paper. Available at SSRN: 1108411.

Mills, T., and Markellos, R. (2008). “The Econometric Modelling of Financial Time Series”, Cambridge University Press.

Moinas, S. (2006). “Hidden limit orders and liquidity in limit order markets”, Unpublished working paper, Toulouse Business School, Toulouse, France, available online at: http://www.efmaefm.org/efma2006/papers/306635_full.pdf.

Montgomery, J. (2016). “Spoofing, Market Manipulation, and the Limit-Order Book”, Available at SSRN: https://ssrn.com/abstract=2780579.

Morck, R., Yeung, B. and Yu, W. (2000). “The information content of stock markets: why do emerging markets have synchronous stock price movements?”, Journal of financial economics, Vol. 58(1-2), pp. 215-260.

Nerudová, D., and Dvořáková, V. (2014). “Financial Transaction Tax: Can it be sufficient resource of EU budget when introduced through enhanced cooperation?”, *Procedia Economics and Finance*, *12*, 453-461.

Obizhaeva, A. and Wang, J. (2005). “Optimal trading strategy and supply/demand dynamics”. NBER Working Papers, No 11444.

Paddrik, M., Hayes, R., Todd, A., Yang, S., Beling, P. and Scherer, W. (2012). “An agent based model of the E-Mini S&P 500 applied to Flash Crash analysis”. In Computational Intelligence for Financial Engineering & Economics (CIFEr), 2012 IEEE Conference on (pp. 1-8). IEEE.

Plerou, V., Gopikrishnan, P., Gabaix X. and Stanley, H. (2002). “Quantifying stock-price response to demand fluctuations”. Physical Review E 66, No. 2, pp. 027104.

Saporta, V. and Kan, K. (1997). “The Effects of Stamp Duty on the Level and Volatility of Equity Prices”, Bank of England 1997, ISSN 1368-5562.

Scholtus, M. and van Dijk, D. (2015). “High-Frequency Activity on NASDAQ”, Handbook of High Frequency Trading, 1st ed., G. N. Gregoriou, Ed. Elsevier Inc., pp. 3-23.

Schulmeister, S. (2009). “A General Financial Transaction Tax: A Short Cut of the Pros, the Cons and a Proposal”, WIFO Working Paper No. 344, Available at SSRN 1714336.

Schulmeister, S., Schratzenstaller, M., and Picek, O. (2008). “A General Financial Transaction Tax–Motives, Revenues, Feasibility and Effects”, Study of the Austrian Institute of Economic Research (WIFO), Vienna.

Securities and Exchange Board of India (SEBI) (2016). “Discussion paper on 'Strengthening of the Regulatory framework for Algorithmic Trading & Co-location”; available online on: https://www.sebi.gov.in/sebi_data/attachdocs/1470393485587.pdf.

Security and Exchange Commission (SEC) (2010). “Findings Regarding the Market Events of May 6, 2010: Report of the Staffs of the CFTC and SEC to the Joint Advisory Committee on Emerging Regulatory Issues”. U.S. Commodity Futures Trading Commission and U.S. Securities and Exchange Commission joint committee, September 30. Available at: https://www.sec.gov/news/studies/2010/marketevents-report.pdf.

Sornette, D. and Von der Becke, S. (2011). "Crashes and High Frequency Trading: An evaluation of risks posed by high-speed algorithmic trading", Swiss Finance Institute Research Paper Series No. 11-63, available online at: http://ssrn.com/abstract=1976249.

Tóth, B., Lemperiere, Y., Deremble, C., De Lataillade, J., Kockelkoren, J. and Bouchaud, J-P. (2011). "Anomalous price impact and the critical nature of liquidity in financial markets", Physical Review X 1, No. 2, pp. 021006.

Treleaven, P., Galas, M., and Lalchand, V. (2013). "Algorithmic trading review", Communications of the ACM, Vol. 56(11), pp. 76-85.

Tsang, E (2017). "Directional Changes: A New Way to Look At Price Dynamics". Mandal J., Dutta P., Mukhopadhyay S. (eds) Computational Intelligence, Communications, and Business Analytics (CICBA), Communications in Computer and Information Science, Vol. 775, Springer, 2017, 45-55. Accessed online at: http://www.bracil.net/finance/papers/Tsang-DCHorizon-2017.pdf.

Umlauf, S. (1993). "Transaction taxes and the behavior of the Swedish stock market", Journal of Financial Economics, 33(2), 227-240.

United Nations Development Program (1992). "1992 Human Development Report", New York: Oxford University Press.

United States Department of Justice Criminal Report (2015). "United States of America v. Navinder Singh Sarao", United States District Court Northern District of Illinois Eastern Division. Available online at: https://www.justice.gov/sites/default/files/opa/press-releases/attachments/2015/04/21/sarao_criminal_complaint.pdf.

Vacha, L., Barunik, J., and Vosvrda, M. (2012). "How do skilled traders change the structure of the market", International Review of Financial Analysis, 23, 66-71.

Vives, X. (2008). "Information and learning in markets: the impact of market microstructure", Princeton University Press.

Weber, P. and Rosenow, B. (2005). “Order book approach to price impact”, Quantitative Finance, Vol. 5(4), pp. 357-364.

Zhang, S. (2012). “Need for Speed: An Empirical Analysis of Hard and Soft Information in a High Frequency World”, School of Economics and Business Engineering, Karlsruhe Institute of Technology, Germany.

Zovko, I. (2008). “Topics in Market Microstructure”, Doctoral dissertation, University of Amsterdam.