An Empirical Investigation of Price Impact: An Agent-based Modelling Approach

By

Wei Cui

B.Sc. Hon (Wuhan University)

M.Eng. Hon (Dublin City University)

Thesis presented for the degree of Ph.D.

to the

Michael Smurfit School of Business

University College Dublin

Research Supervisor:

Prof. Anthony Brabazon

November 2012

Abstract

Understanding price impact is a fundamental task in finance. Many execution algorithms, used to execute a large order by dividing and spreading it over time, are based on price effects and in particular on the way how volume affects prices. Moreover, the analysis of price impact is helpful for understanding how financial markets function as price impact is one of the mechanisms determining price formation.

The thesis is motivated by the recent emergence of algorithmic trading which requires a good understanding of price impact. This thesis addresses three questions concerning price impact in order to gain a better understanding on the intraday behaviour of price impact, and the factors affecting price impact.

The first study examines the intraday behaviours of price impact and market liquidity. The data is drawn from the NYSE-Euronext TAQ database and the LSE ROB database. Six stocks from the US markets and six stocks from the UK markets are analysed. The intraday patterns on price volatility, bid-ask spread, trading volume and market depth are documented and generally confirm findings in prior studies on intraday phenomena. In particular, a reverse S-shaped intraday pattern on price impact is found for both US and UK stocks for the first time.

The second study investigates whether agent intelligence plays an important role in determining the magnitude of price impact. This chapter constructs an artificial stock market composed of zero-intelligence agents, and calibrates it using the LSE ROB data. The result shows that the price impact in the artificial market is generally larger than that in the real market. This is consistent with the hypothesis that agent intelligence plays an important role in determining the magnitude of price impact. It supports the selective liquidity argument in Farmer et al. (2004) & Hopman (2007).

The third study addresses whether order choice affects the price impact of trading a large order. A typical approach in trading a large order is to devise a strategy which divides it into numerous pieces and spreads it over time (usually one trading day). In this study, several execution strategies with various order types, and a number of simple strategies with one order type as benchmarks are constructed and evaluated by their effects on prices. Novelly, these strategies are evolved and evaluated in simulated artificial markets. The results show that the combined strategies outperform the simple strategies significantly, suggesting that order choice plays an important role in determining the price impact of trading large orders.

The results in this thesis suggest that time-of-the-day, agent intelligence and order choice are important factors affecting price impact, and need to be considered in the theoretical microstructure models and in the design of trading strategies.

Acknowledgements

I wish to express my greatest appreciation to Prof. Anthony Brabazon, my academic supervisor, for his strong support and insightful suggestions. He provided me with valuable guidance and was always accessible and generous with his comments whenever I had questions. Without his consistent encouragement and help, this work would be impossible.

A special note of my sincere gratitude goes to Dr. Michael O'Neill, for his contribution, encouragement and support. A warm thank to the wonderful FMC^2 and NCRA group members for their constant support and all their fruitful discussions.

I want to thank also the two reviewers of my dissertation: Dr. Mark Hutchinson and Dr. Conall O'Sullivan, for their involvement in reviewing the manuscript.

I dedicate this thesis to my family and friends who all came along for the ride and without whom none of this would have been possible. Special thanks go to Dad and Mom.

I would also like to thank Science Foundation Ireland. This research is based upon works supported by the Science Foundation Ireland under Grant Number 08/SRC/FM1389.

Last but not least, very special thanks to my dear wife for her unflagging love and support.

Contents

List of Figures

List of Tables

Publications

Parts of this thesis have featured in various publications. These include:

Journal Article:

- Cui, W., Brabazon, A. and O'Neill, M. (2011). Adaptive Trade Execution: A Grammatical Evolution Approach. *International Journal of Financial Markets and Derivatives*. Vol. 2, No. 1/2, pp. 4-31. (Cui et al., 2011)

Book Chapter:

- Cui, W., Brabazon, A. and O'Neill, M. (2010). Evolutionary Computation and Trade Execution. *Natural Computing in Computational Finance (volume III)*, pp. 51-69, Springer-Verlag, Berlin. (Cui et al., 2010b)

Conference Papers:

- Cui, W. and Brabazon, A. (2012). An Analysis of Liquidity in London Stock Exchange. *Proceedings of the 25th Irish Accounting and Finance Association Conference*, pp. 1-45, Galway, Ireland. (Cui & Brabazon, 2012b)

- Cui, W. and Brabazon, A. (2012). An Agent-based Modelling Approach to Study Price Impact. *Proceedings of the 2012 IEEE Conference on Computational Intelligence for Financial Engineering & Economics*, pp. 1-8, New York City, USA. (Cui & Brabazon, 2012a)

- Cui, W. and Brabazon, A. (2011). Ultra High-frequency Financial Data: Analysing NYSE Euronext TAQ Data. *Proceedings of the 24th Irish Accounting and Finance Association Conference*, pp. 1-38, Cork, Ireland. (Cui & Brabazon, 2011)

- Cui, W., Brabazon, A. and O'Neill, M. (2010). Evolving Efficient Limit Order Strategy Using Grammatical Evolution. *Proceedings of the 2010 IEEE Congress on Evolutionary Computation*, pp. 1-6, Barcelona, Spain. (Cui et al., 2010d)

- Cui, W., Brabazon, A. and O'Neill, M. (2010). Evolving Dynamic Trade Execution Strategies Using Grammatical Evolution. *Proceedings of the 2010 International Conference on Applications of Evolutionary Computation, Lecture Notes in Computer Science 6025*, pp. 192-201. (nominated as best paper) (Cui et al., 2010c)

- Cui, W., Brabazon, A. and O'Neill, M. (2009). Efficient Trade Execution Using a Genetic Algorithm in an Order Book Based Artificial Stock Market. *Proceedings of the 11th Annual Conference Companion on Genetic and Evolutionary Computation Conference*, pp. 2023-2028. (Cui et al., 2009a)

- Cui, W., Brabazon, A. and O'Neill, M. (2009). Evolutionary Computation in Trade Execution. *Proceedings of the 22nd Irish Accounting and Finance Association Conference*, pp. 1-10, Dublin, Ireland. (Cui et al., 2009b)

Conference Abstract:

- Cui, W., Brabazon, A. and O'Neill, M. (2010). An Analysis of Price Impact. *Proceedings of the 3rd International Conference of European Research Consortium on Informatics and Mathematics*, London, UK. (Cui et al., 2010a)

Chapter 1

Introduction

This thesis is concerned with the effect of trading on the price of a financial security, that is price impact. Understanding price impact is a fundamental task in finance, as price impact is one of the basic mechanisms determining price formation. In one of the seminal papers on price impact, Black (1971) stated that large order execution would always exert an impact on price, irrespective of the method of execution or technological advances in market structure. Due to the availability of transaction-level financial data, an increasing number of studies have been devoted to the analysis of price impact in the last two decades (Bouchaud et al., 2009).

Recent years have seen the emergence of an important technological product of electronic trading, namely *Algorithmic Trading* (AT), which has received considerable public attention and is commonly defined as the use of computer programs to automate the trading process which encompasses making trading decisions, submitting orders and managing orders after their submissions (Hendershott et al., 2011). AT now makes up a large portion of the market activity in multiple financial markets ranging from equities to foreign exchange (Chaboud et al., 2009), to derivative (Mishra et al., 2012) markets. It accounted for more than 70% of equity trades in the U.S. by 2009 (Hendershott et al., 2011). Significant volumes traded using AT are also found in Asian markets (Decovny, 2008; Grant, 2012) and European markets (Hendershott & Riordan, 2012).

The widespread use of AT has stressed the importance of studying price impact as understanding price impact is crucial for the design of algorithmic trading strategies. Moreover, the significance of the study of price impact is also illustrated by the "Flash Crash" of May 6, 2010. An important contribution in triggering this event was the extremely rapid execution of a large order of futures contracts. From this point of view, understanding price impact is helpful for financial regulators to monitor markets.

The research objective of this thesis is to empirically investigate price impact, aiming to gain a better understanding of the intraday behaviours of price impact and the factors affecting price impact. This thesis contributes to the literature on market microstructure.

The rest of this chapter is organised as follows. The next section offers some background information relevant to this thesis, including a description of market mechanism and a brief introduction to trade execution. The research aims and objectives of this thesis are then outlined. Following from this is a short overview of the research framework, along with the contribution of the thesis and its scope limitations. An overview of the structure of the remainder of this thesis then completes the chapter.

1.1 Limit Order Market and Trade Execution

Traditional exchanges are organised as dealer markets, where designated market makers set the quotes and directly trade with other market participants. With the developments of information technology over the last few decades, electronic trading systems have gradually replaced the role of designated market makers. Many securities exchanges have set up new limit order markets, or converted traditional dealer markets to hybrid ones, which allow trading via a blend of an electronic trading platform and a traditional floor broker system.

- Examples of pure limit order markets include Tokyo Stock Exchange, Hong Kong Stock Exchange, Paris Bourse, Toronto Stock Exchange, Australian Stock Exchange, Shanghai Stock Exchange, and Singapore Exchange.

- Examples of newly set-up limit order markets in traditional exchanges include the Stock Exchange Electronic Trading Service (SETS) system in the London Stock Exchange (LSE), the Globex trading system in the Chicago Mercantile Exchange (CME), and the electronic option trading platform in the International Securities Exchange.

- Examples of hybrid markets include the New York Stock Exchange (NYSE) and NASDAQ.

The electronic revolution has also brought many off-exchange trading venues, including electronic communication networks (ECNs) and dark pools. Examples of ECNs include ArcaEdge, BATS Chi-X, and Turquoise. In contrast to ECNs and limit order markets where order and quote information are visible, participants in dark pools cannot see the orders placed by the other participants. Examples of dark pools include Instinet, Liquidnet, and POSIT.

This thesis focusses attention on limit order markets which are the prevalent trading mechanism in the world's major stock exchanges. A brief introduction to the limit order market mechanism is provided in the next section followed by a description of the trade execution problem.

1.1.1 Limit Order Book

Today more than half of the stock exchanges around the world are limit order markets which operate a limit order book to match buyers and sellers (Jain, 2003). One advantage of an electronic limit-order-book is the greater transparency offered by these systems when compared with dealer market settings. Price quotes and transactions are visible to all participants which generally improves the efficiency of price discovery, thus promoting market confidence (Gould et al., 2011; Parlour & Seppi, 2008). It also promotes competition as dealers/market makers are encouraged to post the best prices to attract order flow (CFA, 2009).

In a limit order market, traders can either submit *a limit order* or *a market order*. A market order is an order to buy or to sell a specified number of shares. It guarantees immediate execution but provides no control over its execution price. In contrast, a limit order is an order to buy or to sell a specified number of shares at a specified price. It provides control over its execution price but does not guarantee its execution.

Table 1.1: Order Book 1

Bid		**Ask**	
Shares	Prices	Prices	Shares
300	50.19	50.22	200
200	50.18	50.23	300
400	50.17	50.24	100
500	50.16	50.25	300
300	50.15	50.26	200
100	50.14	50.27	400

Table 1.2: Order Book 2

Bid		**Ask**	
Shares	Prices	Prices	Shares
300	50.19	50.22	**200**
500	50.18	50.23	300
400	50.17	50.24	100
500	50.16	50.25	300
300	50.15	50.26	200
100	50.14	50.27	400

Table 1.1 shows a sample order book, where all the buy and sell orders are visible/transparent to traders in the market. It consists of two queues which store buy and sell limit orders, respectively. Buy limit orders are called *bids*, and sell limit orders are called *offers* or *asks*. The highest bid price on the order book is called *best bid*, and the lowest ask price on the order book is called *best ask*. The

Table 1.3: Order Book 3

Bid		**Ask**	
Shares	Prices	Prices	Shares
300	50.19	50.22	**100**
500	50.18	50.23	300
400	50.17	50.24	100
500	50.16	50.25	300
300	50.15	50.26	200
100	50.14	50.27	400

Table 1.4: Order Book 4

Bid		**Ask**	
Shares	Prices	Prices	Shares
300	50.19	**50.23**	**100**
500	50.18	50.24	100
400	50.17	50.25	300
500	50.16	50.26	200
300	50.15	50.27	400
100	50.14	50.28	300

difference between best bid and best ask is called *bid-ask spread*. Prices on the order book are not continuous, but rather change in discrete quanta called *ticks*.

Orders in a limit order market arrive stochastically in time. The price limit of a newly arrived order is compared to those of orders already held in the system to ascertain whether there is a match. If so, the trade occurs at the price set by the first order. The set of unexecuted limit orders held by the system constitutes the dynamic order book, where limit orders can be cancelled or modified at any time prior to their execution. Limit orders on the order book are typically (depending on market rules) executed strictly according to (1) price priority and (2) time priority. Bid (ask) orders with higher (lower) prices get executed first with time of placement being used to break ties. A buy (sell) market order is executed at the best ask (bid) price. The limit order book is highly dynamic, because new limit orders will be added into the order book, and current limit orders will get executed or cancelled from the order book throughout the trading day. Table 1.2 shows the order book after a trader submits a buy limit order with 300 shares placed at price 50.18. Table 1.3 shows the order book after a trader submits a buy market order with 100 shares. Table 1.4 shows the order book after a trader submits a buy market order with 300 shares.

Apart from market and limit orders, some stock exchanges also offer *hidden/iceberg orders* to allow traders to conceal the total size of a large limit order.

Such orders consist of two components, a small component whose size is visible in the order book and a larger hidden component with a size known only to the order submitter. The hidden component is exposed to the market gradually through execution of the visible part of the order (Aitken et al., 2001; Biais et al., 1995). Many electronic trading platforms have introduced this kind of order, including Euronext, the Toronto Stock Exchange, the London Stock Exchange, and XETRA. Hidden limit orders are often used by large liquidity traders to hide their intent to trade (Bongiovanni et al., 2006). However, iceberg orders typically exhibit a less favorable time priority compared with pure limit orders (Bessembinder et al., 2009; Esser & Monch, 2007). After the visible portion of an iceberg order is completely matched, other visible limit orders at the same limit price that were entered before the new portion is displayed take priority.

1.1.2 Trade Execution

> "The investment process has been described as a three-legged stool supported equally by securities research, portfolio management, and securities trading. Of the three, trading is often the least understood and the least appreciated function."
>
> A. Madhavan, J.L. Treynor and W.H. Wagner (2007)

Trade execution is the implementation of trading decisions. The costs that arise during the execution process are called *trading costs*. Understanding trading costs is important for the design of efficient and effective trade execution strategy, as well as for the assessment of performance of traders, brokers and trading algorithms.

Trading costs involve commissions paid to brokers, fees paid to exchanges, spread, market impact, and opportunity cost. The first three items are visible, which can be known before any trade is made. The last two items are not directly measurable a priori. These five costs are explained as follows.

1. *Commission* is the payment made to brokers, which varies from broker to broker. Commissions have been greatly reduced over recent years. Goldstein et al. (2009) analysing institutional commissions using the data provided by Greenwich Associates find that the average commission per share in 2004 is under 5 cents, about one third of that in 1977.

2. *Fees* are charged by exchanges during the execution process. These are often bundled into the total commission paid to brokers.

3. *Bid-ask spread* is the difference between the best available bid and ask price in markets.

4. *Market impact*, also known as *price impact*, is the effect of a trade on the price of the asset. The most obvious illustration of price impact is that a buying event always drives the price up and a selling action pushes the price down. For example, the mid-quote (the average of the best bid and the best ask) of a stock drops from $200 to $199.6 (supposed) after a market sell order is executed, where the price decline is the effect caused by trading the sell order.

5. *Opportunity cost* arises when an order fails to be fully executed. For example, a trading task of selling 120,000 shares at $200 per share is supposed to be finished by the end of the trading day, but only 100,000 shares are sold by the end of the day. At that time, the price drops to $198. The opportunity cost of failing to execute this entire order is $40,000 which is calculated as 20,000 multiplied by the $2 price decline per share.

Many empirical studies (Chan & Lakonishok, 1995, 1997; Collins, 1991; Kraus & Stoll, 1972) indicate that the visible components of trading costs are much smaller than the invisible components and price impact always dominates total trading cost. As markets have become electronic, commissions and fees have been driven down. In the optimal trading literature (Alfonsi et al., 2010; Almgren & Chriss, 1999, 2000; Bertsimas & Lo, 1998; Hora, 2006; Obizhaeva & Wang, 2012), price impact cost is equated with trading cost.

A practical problem arising in limit order markets is how to buy or sell large quantities of an asset with minimum execution costs. Executing a large order at once will cause significant price impact cost. The most obvious way to reduce this cost is to break the large order into a number of smaller pieces and spread them over time. A trade execution strategy is devised to minimise the price impact cost.

1.2 Research Aims

In this thesis, three broad questions are addressed in order to gain a better understanding of price impact:

1. *Does price impact on the LSE and the NYSE exhibit an intraday pattern?* The literature on market microstructure has discovered many intraday patterns of a number of interesting variables in financial markets. For example, bid-ask spread on the NYSE exhibits a U-shaped intraday pattern (Lee et al., 1993): the spread is relatively higher at market open and market close than during the middle of the trading day. However, only a few studies pay attention to the intraday behaviour of price impact. Chan (2000) finds that price impact displays a U-shaped pattern over the trading day on the Hong Kong Stock Exchange. This thesis aims to examines the intraday behaviour of price impact in US and UK markets, using the data drawn from the NYSE-Euronext TAQ database and the LSE ROB database.

2. *Does agent intelligence affect the magnitude of price impact?* Previous studies

(e.g. Farmer et al. (2004); Hasbrouck (1991); Lillo et al. (2003)) concerning price impact find that price impact is influenced by trade size and the state of the limit order book, and is a concave function of trade size. Some of these studies (Farmer et al., 2004; Hopman, 2007; Weber & Rosenow, 2006) argue that this concavity is due to agent's selective liquidity taking, namely agents conditioning their trade size on market liquidity. The second aim of this thesis is to investigate whether agent intelligence plays an important role in determining price impact.

3. *Does order choice affect price impact when trading large orders?* Controlling price impact when trading a large order is an important issue in financial markets. Many studies have been devoted to this problem. However, only market orders are examined in these studies. The third aim of this thesis is to examine the influence of order choice on the price impact of trading a large order.

These questions have practical significance. Without an understanding of the influences of these three factors on price impact, our understanding of price impact is incomplete. Answering these questions will help us gain a better understanding of price formation, thus will help us to better understand how securities markets function, and help practitioners to devise better trading strategies leading to reduced execution costs and therefore improved investment returns.

In the first experiment, an intraday pattern on price impact is expected to be observed in US and UK markets. In the second experiment, the price impact generated from the artificial market is expected to be higher than the price impact produced from the real market, which will support the hypothesis that agent intelligence is an important determinant of the magnitude of price impact. In the third experiment, it is expected that the trade execution strategies adopting multiple order types outperform those using single order type, which indicates that order choice between market order and limit order plays an important role in determining price impact when trading a large order.

1.2.1 Objectives of Research

The following objectives need to be achieved in order to fulfill the research aims of this thesis.

1. Review relevant literature on market microstructure and price impact.
2. Develop an algorithm for preprocessing TAQ data.
3. Develop an algorithm for rebuilding the limit order book using the LSE ROB data.
4. Develop an agent-based model of a limit order market.
5. Design trade execution strategies.
6. Analyse price impact.
7. Draw conclusions from the results.

1.3 Framework of Research

To address the first research aim, an empirical analysis of intraday market liquidity and price impact is undertaken using ultra high-frequency datasets from the London Stock Exchange (LSE) and the NYSE-Euronext. The datasets record trade and quote information on each trading day. This enables us to identify the price before and after each single trade, and thus measure the price impact of each trade. Before its use for analysis, the data has to be preprocessed.

To address research aims two and three, an *agent-based modelling* (ABM) approach is adopted. This is a computerised simulation consisting of a number of agents. The emergent properties of an agent-based model are the result of "bottom-up" processes, where the decisions of individual and interacting agents at a microscopic level determine the macroscopic behavior of the system (LeBaron, 2006; Samanidou et al., 2007; Tesfatsion, 2006). Unlike classical micro-economic models, the concept of equilibrium has little meaning in ABM-simulated systems, as the interactions of the agents result in continual movement between system states. The collective emergent property of the complex system makes the whole become more than the sum of its parts.

Agent-based modelling approach has been applied in the fields of economics and finance for many years. One of the earliest agent-based models in concepts was developed by the Nobel laureate Thomas Schelling (1971). It is now widely used in economic and finance as an alternative research methodology and offers a number of advantages. The first is that most agent-based stock markets exhibit some stylised facts, like volatility clustering and fat tails of price returns, which can not be explained by traditional theoretical models (Chen et al., 2012). Another advantage is that it is very easy to analyse one specific factor by isolating it in the simulation. On the contrary, it is often difficult to isolate the effects resulting from a specific factor from others with econometric approaches.

In order to investigate whether agent intelligence and order choice decision affect price impact, agent-based modelling is used to simulate artificial markets which isolate these two factors. Chapter 6 simulates an artificial market consisting of intelligence-isolated agents, namely zero-intelligence agents, in order to investigate the price impact generated by zero-intelligence agents. In Chapter 7, artificial markets considering agent's order choice decisions are simulated, as well as artificial markets isolating agent's order choice decisions, in order to evaluate the effect of order choice on the price impact when trading a large order. In this thesis, the artificial market settings possess the salient features of electronic limit order markets: continuous trading, a visible book of orders, price-time priority rules, instantaneous trade reporting rules, order cancellation capabilities, and both limit order and market order functionality.

1.4 Contributions of Thesis

The work presented in this thesis has given rise to a number of contributions as described below.

- **Up-to-date review of relevant literature**

  This thesis presents an up-to-date survey of relevant literature on market microstructure, price impact, and agent-based modellings of financial markets. The review in Chapter 2 covers up-to-date studies on the relationship between price impact and trade size, studies on large price impact are covered, studies of intraday phenomena on price volatility, bid-ask spread, trading volume and market depth, studies on optimal trading strategies for a single asset, and studies on order submission strategies. The literature review of agent-based financial markets in Chapter 3 covers various applications of agent-based modelling to financial markets.

- **Analysis of intraday behaviours of price impact in UK and US markets**

  This thesis analyses intraday behaviours of price impact, using the recent data drawn from the NYSE-Euronext TAQ database and the LSE ROB database. While many studies analyse intraday behaviours of price volatility, bid-ask spread, trading volume and market depth, little attention has been paid to price impact in UK and US markets. This thesis aims to shed light on this, and for the first time documents intraday patterns of price impact in UK and US markets.

- **Up-to-date evidence on intraday behaviours of market liquidity in UK and US markets**

  This thesis analyses intraday behaviours of market liquidity, using the data drawn from the NYSE-Euronext TAQ database and the LSE ROB database. Unlike previous studies on intraday phenomena which only examined the market depth at the best level of the limit order book, this thesis for the first time analyses the intraday behaviour of market depth for up to ten levels of the limit order book in the UK market. Moreover, since the previous studies on intraday phenomena are dated, this thesis aims to provide up-to-date evidence on intraday behaviours of price volatility, bid-ask spread, market depth, trading volume, trading frequency and trade size in UK and US markets.

- **Investigation of the role of agent intelligence in determining the magnitude of price impact**

  Several studies argue that the concavity of price impact is due to selective liquidity taking. However, no studies have been conducted to investigate this. This thesis for the first time investigates whether agent intelligence plays an important role in determining the magnitude of price impact, and provides evidence supporting this hypothesis.

- **Examination of the effect of order choice on the price impact when trading a large order**

  Many studies examine how to control price impact when trading a large order. However, these studies only consider market orders. This thesis for the first time analyses whether order choice affects the price impact of trading a large order, and demonstrates that it does.

- **Use of ABM for investigating price impact**

  Various applications of agent-based modelling to economics and finance can be found in prior literature, but there has been little attempt to use ABM for investigating price impact. This thesis makes the first attempt to investigate price impact using ABM in its two experimental studies. In Chapter 6, agent-based modelling is used to simulate an artificial limit order market consisting of zero-intelligence agents, in order to examine the price impact produced by zero-intelligence agents. In Chapter 7, agent-based artificial markets are also simulated in order to analyse the price impact of trading a large order.

1.5 Scope Limitations

This thesis makes a useful contribution to our understanding of price impact. However, it does not conduct an exhaustive analysis of all possible avenues. There are a number of areas worthy of further research that fall outside the scope of this work. The following parts of this section list some of these areas.

When examining the intraday behaviours of price impact, the focus of the present work is on the price impact caused by market orders. The analysis can be extended to price impact of other types of orders. For example, Eisler et al. (2012); Hautsch & Huang (2012) examine the price impact caused by limit orders, and Gabaix et al. (2006); Keim & Madhavan (1996) investigate the price impact of upstairs orders.

In this thesis, an agent-based modelling approach is used to model price impact, that is the immediate effect of trading on prices. This could be extended to consider the resilience of price impact. For example, Gatheral et al. (2011) models that the price impact decays exponentially over time.

While studying the issue of controlling price impact, attention is limited to the trading of a large order occurring in one trading venue. This could be extended to the case where a large order is being traded across multiple trading venues. For example, Foucault & Menkveld (2008) empirically examine the smart order routing systems, and Rawal (2010) discusses the importance of intelligent decision-making in order routings.

1.6 Structure of Thesis

The remainder of this thesis is structured as follows:

- **Chapter 2 - Review of Relevant Literature**

  A review of the literature on intraday patterns, price impact, trade execution strategies is provided. Some research gaps in the current state of the art are identified at the end of the chapter.

- **Chapter 3 - Agent-based Modelling in Finance**

  A brief introduction to agent-based modelling and its applications to finance are given in this chapter, along with an introduction to the Genetic Algorithm, a technique often adopted for simulating agent learning.

- **Chapter 4 - Datasets and Data Processing**

  A description of the datasets used to analyse price impact, as well as the data preprocessing steps required is provided.

- **Chapter 5 - Intraday Behaviour of Price Impact**

Ultra high-frequency data from the LSE and the NYSE-Euronext is used to analyse the intraday behaviour of price impact, as well as intraday features of market activity and market liquidity.

- **Chapter 6 - An Investigation of Price Impact and Agent Intelligence**

  An agent-based modelling approach is adopted to simulate the market for a particular stock in the LSE, where every aspect of the order flow in the LSE market is imitated in the artificial market except the trader's intelligence. Then the price impact reproduced from the artificial market is compared with that in real market, in order to understand whether agent intelligence plays an important role in determining the magnitude of price impact.

- **Chapter 7 - Dynamic Trade Execution with Limit and Market Orders**

  A number of trade execution strategies which adopt different types of orders are devised as well as some benchmark strategies which use single order type. These strategies are evaluated by their trading impact on the market, in order to examine the effect of order choice on price impact when trading a large order.

- **Chapter 8 - Conclusions**

  A summary of the thesis is provided in this chapter, along with an overview of its contributions and limitations. Opportunities for future work are also outlined.

Part I

Literature Review

In Part I the relevant literature is reviewed and the main methodology of research is described.

In Chapter 2, the theoretical background of this thesis, that is market microstructure, is introduced, and a brief overview is provided. Previous work on price impacts, intraday patterns and trade execution is reviewed, and gaps of research are revealed.

In Chapter 3, agent-based modelling approach is introduced, and framework of modelling is described. Prior work on artificial markets modelling is reviewed.

Chapter 2

Market Microstructure

This chapter provides a literature review on relevant work, including an overview of market microstructure, a review of literature on intraday phenomena, a review of studies concerning price impact, and a review of literature on trade execution strategies. At the end of the literature review, a few research gaps in the current state of the art are identified which lead to the three experimental studies of this thesis.

2.1 Introduction

The theoretical foundation for the studies in this thesis is found primarily in the field of market microstructure. Market microstructure is formally defined by O'Hara (1995) as "the study of the process and outcomes of exchanging assets under explicit trading rules".

Studies on market microstructure are helpful in understanding the returns to financial assets and how financial markets become efficient. Moreover, it has immediate application in understanding the design of trading strategies and the regulation of financial markets.

The domain of market microstructure is wide and broad. An overview of this area is provided in section 2.2. This thesis is related to several streams of market microstructure research. The following topics will form the focus of this chapter.

- Price impact.
- Intraday pattern.
- Trade execution strategies.

2.2 Overview of Market Microstructure

The study of market microstructure involves every aspect of the trading process. There has been a vast number of studies on market microstructure since its emergence. Several papers and books have provided a valuable survey of the literature in this area. O'Hara (1995) gives a comprehensive review of theoretical work on market microstructure. Madhavan (2000) provides an overview of market microstructure literature from the viewpoint of informational economics. Also, he distinctively summarises the issues examined in market microstructure and classifies them into the four following categories.

1. **Price formation**. It is also known as price discovery, is the process by which prices come to impound new information. This theme concerns every aspect of the trading process.

2. **Market structure and design**. This theme focuses on how different market structures affect the speed and quality of price discovery, liquidity, and the cost of trading, and concerns how to design a better market. The design of a market determines its market microstructure, and thus affects the quality of the market.

3. **Market transparency**. It is defined as the ability of market participants to observe information about the trading process (O'Hara, 1995). This theme focuses on how various levels of market transparency influence the process of price formation.

4. **Applications to other areas in finance**. This theme concerns the interface of market microstructure with other areas of finance including corporate finance, asset pricing and international finance.

2.2.1 Market Liquidity

Market liquidity is an important quality characteristic of a financial market and is also a significant research topic in the area of market microstructure. Understanding liquidity is beneficial for both market participants and the exchange. O'Hara (1995) interprets liquidity as a measure of the cost of waiting and a function of the scale of trading. High liquidity means the ability to trade securities quickly and at desirable prices. Market liquidity is a crucial consideration when market participants make their investment decisions and trading decisions. Moreover, a more liquid exchange can attract more participants which has a network externality effect. The more participants the exchange has, the more valuable the exchange is.

Market liquidity is a multi-dimensional variable. The following four dimensions of liquidity are distinguished by the literature (Black, 1971; Kyle, 1985).

1. *Market tightness* is the ability to execute a buy order and a sell order of an asset at the same price at the same time. The proportional quoted spread (Copeland & Galai, 1983) is often used as an approximation of liquidity costs, measuring hypothetical trading cost.

2. *Market depth* is the ability to execute a sale / purchase of a certain amount of an asset without influence on the price.

3. *Market immediacy* is the ability to execute an order immediately at the prevailing price.

4. *Market resiliency* measures the speed at which liquidity recovers from shocks.

Market liquidity is affected by market structure. Market makers are the only suppliers of liquidity in dealer markets. However, the role of market makers has largely been superceded, as most market have adopted electronic limit order books. In a limit order market, liquidity is supplied by market participants themselves, and is illustratd by the state of the limit order book. The limit order book evolves over time as matched limit orders are deleted from the order book and as new incoming limit orders are added into the order book.

2.2.2 Theoretical Market Microstructure Models

One of the crucial aims of market microstructure is to understand how price is formed. Theoretical market microstructure models of price formation fall into the three types of models. This section briefly introduce the three types of models.

1. *Inventory based model*. These models try to explain security prices from inventory imbalances of market makers. The market maker maintains the bid-ask spread in order to compensate the risk of holding inventory.

2. *Information based model*. The second class of models is based on the idea of asymmetric information. Traders can either be informed or uninformed (Bagehot, 1971). If they posses private information on the value of the asset, they are called informed traders. If they trade in order to rebalance their positions,

they are called liquidity or uninformed traders. A consequence of asymmetric information is that trading itself conveys information.

3. *Limit order market model*. The third class of models tries to explain the price formation process on the limit-order-book market. In a limit order market, the difference between informed traders and uninformed traders is not obvious. Moreover, unlike the dealer market, the limit order market has no specific market maker who maintains market liquidity. The roles of supplying and demanding liquidity are played by the market participants by using either limit orders or market orders. This category of models has a particular focus on analysing how order submission strategies and other aspects of trading affect the price of an asset in limit order markets.

Financial markets are constantly evolving and becoming more complex. From 2005 to 2009, the average trade size on the NYSE has decreased from 724 shares to 268 shares and the average daily trades have increased from 2.9 million to 22.1 million (Durbin, 2010). Thus, one challenge faced by researchers in market microstructure is the dramatically growing amount of data.

2.3 Price Impact

The effect of trading on the price of the asset is called *price/market impact*, which is considered a dynamic property of liquidity. A sell trade is always associated with a following fall in market price and a buy trade is always followed by a rise in market price. Black (1971) states that large order execution would always exert an impact on price, irrespective of the method of execution or technological advances in market structure.

The earliest research empirically examining the effects of trading large orders on a market dates back to Kraus & Stoll (1972), who found that the execution of large orders on the NYSE had a temporary price impact and a permanent price impact. This fact is later confirmed by Biais et al. (1995), Coppejans et al. (2003), Evans & Lyons (2002), Holthausen et al. (1987), and Holthausen et al. (1990). They argue that the temporary impact is related to the trade size and market liquidity, and the permanent impact is associated with information. Recently, a few studies have different views on price impact. Bouchaud et al. (2004) present the view that price impact is fixed and temporary. Farmer et al. (2006) argue that price impact is variable and permanent.

The explanation for price impact in earlier studies asserts that trade affects price because it contains information, and this effect increases with trade size. French & Roll (1986) show that price movements are due to private information conveyed through trading. Hasbrouck (1991) proposes a method to measure the information content in a trade.

2.3.1 Price Impact and Trade Size

The relationship between price impact and trade volume has been extensively studied in the financial market microstructure literature. The conclusion of previous studies is that the price impact of a market order is a concave function of order size, and this has been validated using transaction-level data in Lillo et al. (2003); Farmer & Lillo (2004); Farmer et al. (2005b); Lim & Coggins (2005a); Hopman (2007).

Price impact is typically measured as the difference of mid-quote price before the order arrives and the mid-quote price after it has been executed. The functional form is

$$p = \gamma * v^{\mu}$$

Table 2.1: Price Impact Function of Trade Size

Authors	Functions	Parameters	Market	Years
Hasbrouck (1991)	nonlinear, positive, and increasing, but concave		NYSE	1989
Hausman et al. (1992)	strongly concave		NYSE	1988
Keim & Madhavan (1996)	concave		upstairs market	1985-1992
Torre (1997)	kx^β	$\beta = 0.5$	NYSE	1994
Dufour & Engle (2000)	nonlinear		NYSE (TORQ)	1-Nov-1990 to 31-Jan-1991
Lillo et al. (2003)	$sign(x)\|x\|^\beta / C^\lambda$	$\beta = 0.2 - 0.5; \lambda = 0.4$	NYSE	1995-1998
Potters & Bouchaud (2003)	$\Re(\tau) ln(x)$		Paris Bourse; NASDAQ	2002
Gabaix et al. (2003)	kx^β	$\beta = 0.5$	NYSE	1994-1995
~~Farmer & Lillo (2004)~~	kx^β	$\beta = 0.26$	~~LSE~~	May 2000 to Dec. 2002
Lim & Coggins (2005a)	kx^β / C^λ		ASX	2001-2004
Farmer et al. (2005b)	kx^β	$\beta = 0.25$	LSE (SETS)	1-Aug-1998 to 30-Apr-2000
Hopman (2007)	$kx^\beta + \epsilon$	$\beta = 0.37$	Paris Bourse	4-Jan-1995 to 19-Sep-1999
Bouchaud et al. (2009)	kx^β	$\beta = 0.3$	LSE	2002

where p is the price impact of a market order with order size v, γ and μ are constants of the function. The specific constants of the function vary in different markets and at different time periods. Although previous literature has agreement on the concavity functional form of price impact, it is inconclusive in the specific functional form. A comparison of the empirical studies on immediate price impact is shown in Table 2.1.

Previous empirical studies show that large trades have less marginal price impact than small trades. Three explanations for the concavity result from previous literature are summarised in Bouchaud et al. (2009).

The first one is from the information context as postulated by Barclay & Warner (1993). The price impact reflects the informativeness of trades, which increase with their private information content. The concavity is due to informed traders' "stealthy trading", keeping their orders small in order to avoid revealing their superior knowledge. Hasbrouck (1991) propose a VAR approach to measure the information content in each trade.

The second explanation claims that the concavity is due to the shape of the limit order book. The depth in the order book as a function of the price will determine the market impact for a market order as a function of its size (Daniels et al., 2003). Bouchaud et al. (2002) have observed that the average order book on the Paris Bourse has a maximum away from the best bid/ask. Smaller market orders are faced with less liquidity than larger ones in such limit order markets. This explanation is ruled out by Weber & Rosenow (2006) and Farmer & Zamani (2007). In Bouchaud et al. (2004), the concave shape of the impact as a function of the volume is understood as an order book effect, where the average size of the queue increases with depth.

The third explanation is called selective liquidity taking (Farmer et al., 2004). Traders with large orders to trade are more patient while smaller order traders are less patient. Large order traders may wait for periods of higher liquidity (Weber & Rosenow, 2006; Hopman, 2007).

2.3.2 Large Price Fluctuations and the Order Book

Market liquidity plays a significant role in the price formation process. The role of real information appears to be rather thin when there are large price fluctuations. Price impact is affected by the market liquidity. The price responses to trades of the same size is highly variable. Many studies show that the random walk nature of price changes on short timescales is not due to the unpredictable nature of incoming news, but appears as a dynamic consequence of the competition between market participants (Bouchaud et al., 2004). Recent work suggest that liquidity fluctuations play an essential role in price formation (Farmer & Lillo, 2004; Farmer et al., 2004; Gillemot et al., 2005; Lillo et al., 2005; Weber & Rosenow, 2006). In price formation, liquidity fluctuations dominate over fluctuations in transaction size. Joulin et al. (2008) provide direct evidence that large transaction volumes are not responsible for large price jumps, and conjecture that most price jumps are induced by order flow fluctuations close to the point of vanishing liquidity.

Recently, several papers cast doubts on concavity, and provide evidence that the concave shape cannot be a property of the true price impact a trader in an actual market would observe. Weber & Rosenow (2005) found a strong anticorrelation between price changes and order flow. Weber & Rosenow (2006) found little evidence that price changes larger than five standard deviations can be explained by an extreme order flow alone. They found that a low density of limit orders in the order book, i.e. limited liquidity, was a necessary prerequisite for the occurrence of extreme price fluctuations. Taking into account both order flow and liquidity, large stock price changes can be explained quantitatively. There are two reasons to explain why this concave shape (of volume imbalance) cannot be a property of the true price impact a trader in an actual market would observe (Weber & Rosenow, 2005). Firstly, such a concave price impact would be an incentive to do large trades in one step instead of breaking them up into many smaller ones as is done in practice. Secondly, in the presence of bluffers in the market, the enforcement of a strictly linear price impact is the only way in which a market maker or liquidity trader can protect herself against suffering losses.

The above review of literature on price impact has shown that price impact is determined by trade size as well as the state of the limit order book when the trade occurs. However, the two following questions concerning price impact are still unanswered.

1. It is still unknown about the up-to-date intraday behaviour of price impact. Chan (2000) showed that price impact displayed a U-shaped pattern over the trading day on the HKSE. The review of literature on intraday phenomena will show that many financial variables exhibit intraday patterns which differ across markets and time periods.

2. No study has been conducted to investigate whether agent intelligence plays an important role in determining the magnitude of price impact. Farmer et al. (2004); Hopman (2007); Weber & Rosenow (2006) all argue that trader's choice of trading time affect price impact.

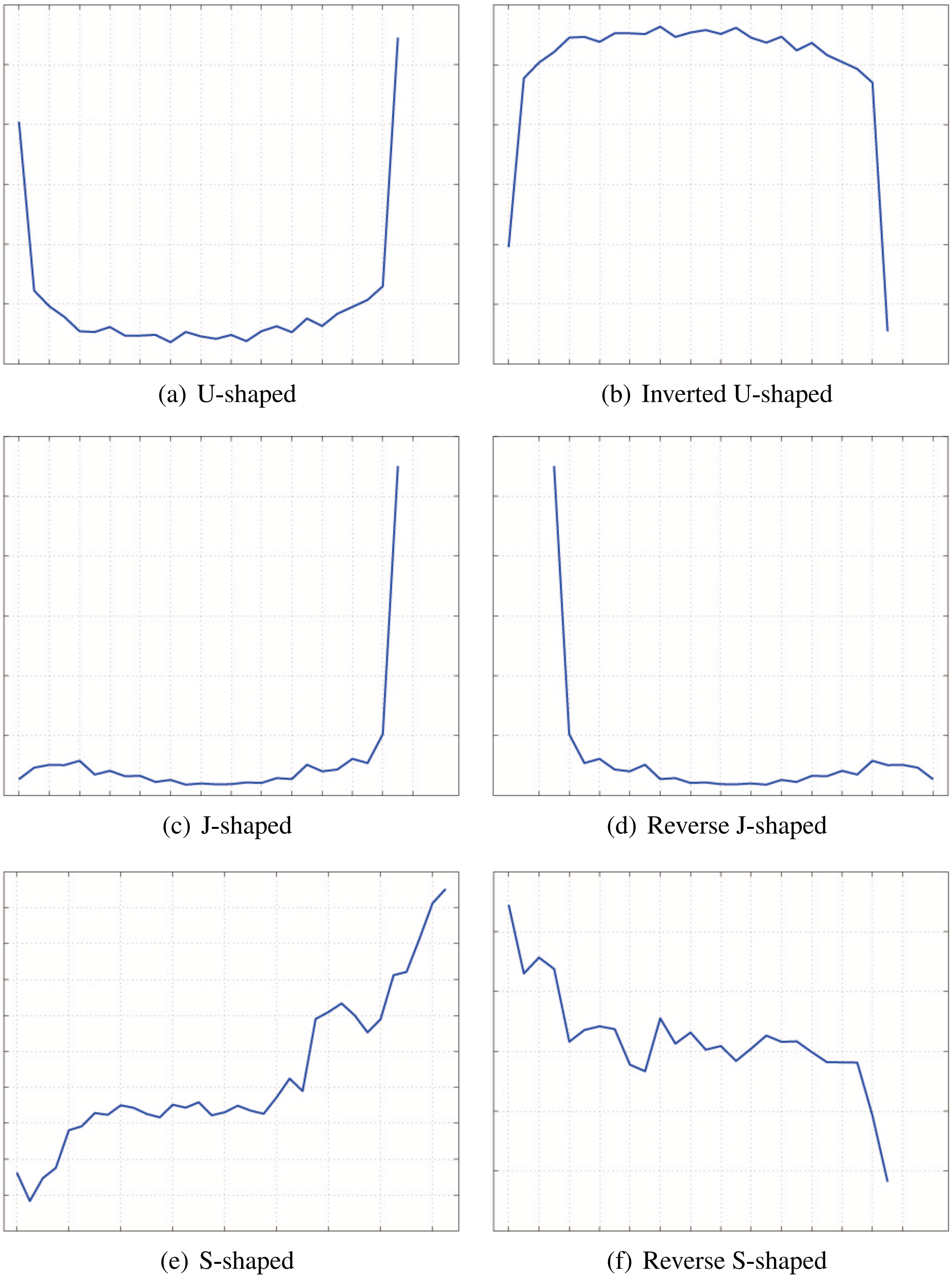
A better understanding of the two questions concerning price impact is helpful for developing more comprehensive price impact models than those reviewed in Section 2.3.1.

2.4 Intraday Patterns

Intraday behaviour of different market variables have been analysed extensively in previous literature which dates back to Wood et al. (1985) and Harris (1986). One of the interesting findings is that many of these variables demonstrate a U-shaped pattern over the trading day: the levels of these variables are relatively high at the beginning and at the end of the trading day, and relatively low in the middle of the day. The empirical evidence on intraday behaviours of these market variables helps us build theoretical models explaining the intraday behaviour of the underlying market.

Different shapes of intraday patterns are defined as below.

- U-shaped pattern: values at market open and close are relatively higher than the values in the middle of the trading day
- Inverted U-shaped pattern: values at market open and close are relatively lower than the values in the middle of the trading day
- J-shaped pattern: values at market close are relatively higher than the values at the rest of the trading day
- Reverse J-shaped pattern: values at market open are relatively higher than the values at the rest of the trading day
- S-shaped pattern: values at market open are relatively lower than the values in the middle of the trading day, and values at market close are relatively higher than the values in the middle of the trading day.
- Reverse S-shaped pattern: values at market open are relatively higher than the values in the middle of the trading day, and values at market close are relatively lower than the values in the middle of the trading day.



(a) U-shaped (b) Inverted U-shaped (c) J-shaped (d) Reverse J-shaped (e) S-shaped (f) Reverse S-shaped

Figure 2.1: Different Shapes of Intraday Patterns

There have been a number of studies which have analysed intraday patterns in financial markets. Admati & Pfleiderer (1988) developed the theoretical foundations of intraday patterns found in financial markets. They argue that both informed and liquidity traders prefer to trade when market is thick, that is when their trades have little impact on price. Thus at periods when market is thick, both informed and liquidity traders are willing to trade. Also the strategic interaction between informed and liquidity traders intensifies the concentration of trading at these periods of the trading day, leading to the observed intraday pattern of trading volume.

Another explanation is provided by Brock & Kleidon (1992). They explain the intraday pattern in a dealer market. They relate the intraday patterns to investors' demand of transactions during the trading period. Due to the accumulation of overnight information or due to the imminent non-trading period, investors have a higher demand for transactions in order to achieve an optimal portfolio mix or transfer overnight risk at market open and close than for the rest of the trading day. This motivates the specialist's interest in exploiting the liquidity inelasticity at these periods, leading to wide spread and high trading volume at open and close.

More recent studies have attributed the intraday pattern to limit order traders. Ahn & Cheung (1999) & Chung et al. (1999) show that limit order traders keep the spread wide in order to offset the high risks of trading against informed traders.

2.4.1 Intraday Price Volatility

Numerous studies examine the intraday variation of bid-ask spread and document a U-shaped intraday pattern (as summarised in Table 2.2). The price volatility is relatively high at market open and close, and relatively low during the middle of the day. The Hong Kong Stock Exchange exhibits a double U-shaped intraday pattern on bid-ask spread as shown by Cheung et al. (1994). This is because that

Table 2.2: Intraday Pattern: Price Volatility

Exchange	**Intraday Pattern**	**Period of Data**	**Authors**
NYSE	U-shaped	Sep. 1971 - Feb. 1972 Jan. - Dec. 1982	Wood et al. (1985)
NYSE	U-shaped	before 1986	Harris (1986)
NYSE	U-shaped	Jan. 1979 - Dec. 1983	Jain & Gun-Ho (1988)
NYSE	U-shaped	1980,1981,1983,1984	McInish & Wood (1990b)
NYSE	U-shaped	Jan. 1964 - Feb. 1989	Lockwood & Linn (1990)
NYSE	U-shaped	Jan. - Jun. 1989	McInish & Wood (1992)
NYSE	U-shaped	Jan. - Dec. 1988	Foster & Viswanathan (1993)
NYSE	U-shaped	Jan. - Dec. 1991	Werner & Kleidon (1996)
NYSE	U-shaped	Jan. - Dec. 1990	Madhavan et al. (1997)
NASDAQ	U-shaped	Oct. - Dec. 1991 Aug. - Sep. 1992	Chan et al. (1995)
LSE	U-shaped	Jan. - Dec. 1991	Werner & Kleidon (1996)
LSE	U-shaped	Jan. - Mar. 1991	Abhyankar et al. (1997)
LSE	U-shaped	Mar. - May 2001	Cai et al. (2004)
HKSE	double U-shaped	before 1994	Cheung et al. (1994)
TkySE	U-shaped	Jan. - Nov. 1991 Feb. 1992 - Apr. 1993	Lehmann & Modest (1994)
FSE	reverse J-shaped	May 2004 - Sep. 2005	Hussain (2011)

Note: HKSE is short for Hong Kong Stock Exchange.
TkySE is short for Tokyo Stock Exchange.
FSE is short for Frankfurt Stock Exchange.

there was a lunchbreak session for two hours in the middle of the trading day at that time period, dividing the trading day into a morning trading session and an afternoon trading session, leading to two U-shaped patterns in the morning and in the afternoon. Hussain (2011), however, observe a reverse J-shaped pattern while examining the intraday price volatility in the Frankfurt Stock Exchange.

Admati & Pfleiderer (1988) provide an explanation for the U-shaped intraday pattern on price volatility. Their model predicts that price volatility is related to informed trading. Prices are more volatile in high-volume periods than in low-volume periods because high-volume periods attract more informed traders than low-volume periods which makes prices are more informative at high-volume periods.

2.4.2 Intraday Bid-ask Spread

A number of studies examine the intraday behaviour of bid-ask spreads in the NYSE and report a U-shaped intraday pattern (Chung et al., 1999; Lee et al., 1993; Madhavan et al., 1997). The spread is wider near the market open and close than during the middle of the day. The U-shaped intraday pattern on bid-ask spread is also observed in the Toronto Stock Exchange (Vo, 2007), the LSE (Abhyankar et al., 1997), the HKSE (Ahn & Cheung, 1999) and the Tokyo Stock Exchange (Lehmann & Modest, 1994).

Some studies have proposed explanations for the U-shaped intraday pattern on bid-ask spread. Brock & Kleidon (1992) attribute the wide spread at market open and close to the actions of the specialists who exploit the liquidity inelasticity at these time periods. Chan et al. (1995) attribute the narrowing of the spread at market close to the improved quotes from dealers who need to balance their inventories. More recent research (Ahn & Cheung, 1999) argues that the U-shaped pattern in spreads is driven by limit order traders who keep spreads wide at the

Table 2.3: Intraday Pattern: Bid-Ask Spread

Exchange	Intraday Pattern	Period of Data	Authors
NYSE	U-shaped	Jan. - Dec. 1988	Lee et al. (1993)
NYSE	reverse J-shaped	Jan. - Jun. 1989	McInish & Wood (1992)
NYSE	U-shaped	Jan. - Dec. 1990	Madhavan et al. (1997)
NYSE	U-shaped	Nov. 1990 - Jan. 1991	Chung et al. (1999)
NYSE	reverse S-shaped	Jan. - Dec. 1999	Henker & Wang (2006)
NASDAQ	reverse S-shaped	Oct. - Dec. 1991 Aug. - Sep. 1992	Chan et al. (1995)
TrtSE	U-shaped	Sep. 1999 - Nov 1999	Vo (2007)
LSE	reverse S-shaped	Jan. - Dec. 1991	Werner & Kleidon (1996)
LSE	U-shaped	Jan. - Mar. 1991	Abhyankar et al. (1997)
LSE	reverse J-shaped	Mar. - May 2001	Cai et al. (2004)
HKSE	U-shaped	Oct. 1996 - Mar. 1997	Ahn & Cheung (1999)
TkySE	U-shaped	Jan. - Nov. 1991 Feb. 1992 - Apr. 1993	Lehmann & Modest (1994)
FSE	reverse J-shaped	May 2004 - Sep. 2005	Hussain (2011)
ISE	reverse J-shaped	May - Jul. 2008	Koksal (2012)

Note: TrtSE is short for Toronto Stock Exchange.
HKSE is short for Hong Kong Stock Exchange.
TkySE is short for Tokyo Stock Exchange.
FSE is short for Frankfurt Stock Exchange.
ISE is short for Istanbul Stock Exchange.

open and close to offset high adverse selection costs. Similar arguments are proposed by Chung et al. (1999) who show that the U-shaped intraday pattern of spread on the NYSE largely reflects the intraday behaviour of the spread established by limit order traders.

In contrast to the U-shaped pattern noted above, reverse J-shaped intraday patterns on bid-ask spread are demonstrated by Cai et al. (2004), Hussain (2011), Koksal (2012) and McInish & Wood (1992). The spread opens wide, narrows down in the first few hours and is relatively stable for the rest of the trading day. As explained by the model of Madhavan (1992), the high spread in the morning is due to greater uncertainty. As information asymmetry is gradually resolved over the trading day, more market participants become informed by observing the market, leading to a decline in the spread during the day.

Besides, some studies (Chan et al., 1995; Henker & Wang, 2006; Werner & Kleidon, 1996) document that the spread declines over the trading day. As pointed out by Chan et al. (1995), the difference of spread pattern between various markets can be attributed to the differences in trading structure. A summary of the studies documenting intraday patterns of the bid-ask spread is provided in Table 2.3.

2.4.3 Intraday Trading Volume

U-shaped intraday patterns on trading volume are documented in many studies (summarised in Table 2.4). Trading volume is high at market open, is relatively low during the day, and rises again as market close approaches. Jain & Gun-Ho (1988) explain that high trading volumes at market open arise because investors trade on information received overnight, and high volumes at market close arise because investors close open positions that they can not monitor overnight. Atkins & Basu (1995) examine the public announcements after trading hours and have the same speculation.

Table 2.4: Intraday Pattern: Trading Volume

Exchange	Intraday Pattern	Period of Data	Authors
NYSE	U-shaped	Sep. 1971 - Feb. 1972 Jan. - Dec. 1982	Wood et al. (1985)
NYSE	U-shaped	Jan. 1979 - Dec. 1983	Jain & Gun-Ho (1988)
NYSE	U-shaped	before 1990	Lockwood & Linn (1990)
NYSE	U-shaped	1980,1981,1983,1984	McInish & Wood (1990b)
NYSE	U-shaped	Jan. - Jun. 1989	McInish & Wood (1992)
NYSE	U-shaped	Jan. - Dec. 1988	Foster & Viswanathan (1993)
NYSE	U-shaped	Jan. - Dec. 1988	Lee et al. (1993)
NASDAQ	U-shaped	Oct. - Dec. 1991 Aug. - Sep. 1992	Chan et al. (1995)
TrtSE	U-shaped	before 1990	McInish & Wood (1990a)
TrtSE	S-shaped	Sep. 1999 - Nov 1999	Vo (2007)
LSE	double-humped	Jan. - Mar. 1991	Abhyankar et al. (1997)
LSE	U-shaped	Jan. - Dec. 1991	Werner & Kleidon (1996)
FSE	reverse J-shaped	May 2004 - Sep. 2005	Hussain (2011)
HKSE	U-shaped	before 1991	Ho & Cheung (1991)
HKSE	U-shaped	Oct 1996 - Mar 1997	Ahn & Cheung (1999)
TWSE	J-shaped	Mar. 1995	Lee et al. (2001)
TkySE	U-shaped	Jan. - Nov. 1991 Feb. 1992 - Apr. 1993	Lehmann & Modest (1994)
ASX	U-shaped	Sep. 1990 - Aug. 1993	Aitken et al. (1995)
SSE	U-shaped	Nov. 1996 - Jan. 1997	Al-Suhaibani & Kryzanowski (2000)
ISE	U-shaped	May - Jul. 2008	Koksal (2012)

Note: TrtSE is short for Toronto Stock Exchange.
FSE is short for Frankfurt Stock Exchange.
HKSE is short for Hong Kong Stock Exchange.
TWSE is short for Taiwan Stock Exchange.
TkySE is short for Tokyo Stock Exchange.
ASE is short for Australia Stock Exchange.
SSE is short for Saudi Stock Exchange.
ISE is short for Istanbul Stock Exchange.

Other shapes of intraday patterns on trading volume are also observed in prior literature. Abhyankar et al. (1997) find a two-humped pattern with highs at 9:00 a.m. and 3:00 p.m. on the LSE. This confirms the prediction in Admati & Pfleiderer (1988), who argue that periods of concentrated trading may occur at some points in the trading day not necessarily at market open or close. Lee et al. (2001) show that trading volume exhibits a J-shaped pattern in the Taiwan Stock Exchange (TSE): the volume is stable during the day, but increases as market close approaches and reaches a peak at market close. Vo (2007) observes that trading volume increases over the trading day in the Toronto Stock Exchange (TrtSe), and attributes the difference of U-shaped pattern in the NYSE and increasing pattern in TrtSE to the structural difference between the two exchanges. Hussain (2011) finds a reverse J-shaped pattern of trading volume in the Frankfurt Stock Exchange (FSE): the volume is highest at market open, decreases in the first ten minutes and is stable for the rest of the day.

2.4.4 Intraday Market Depth

Market depth has received relatively less attention in the literature on intraday market phenomena. Both Lee et al. (1993) and Li et al. (2005) document an inverted U-shaped intraday pattern of market depth on the NYSE. Market depth is lowest at the opening and then rises monotonically until the close, at which point it suddenly drops. These findings are corroborated by Ahn & Cheung (1999) and Vo (2007) who examine the depth on the Hong Kong Stock Exchange and Toronto Stock Exchange respectively. A summary of these studies is given in Table 2.5.

These studies all report a negative association between intraday spreads and depth. The depth displays a reverse pattern of the spread: wide spreads are associated with small depths and narrow spreads are accompanied by large depths. Lee et al. (1993) show that both depth and spread are associated with trading vol-

Table 2.5: Intraday Pattern: Market Depth

Exchange	Intraday Pattern	Period of Data	Authors
NYSE	inverted U-shaped	Jan. - Dec. 1988	Lee et al. (1993)
NYSE	inverted U-shaped	Nov. 1990 - Jan. 1991	Li et al. (2005)
HKSE	inverted U-shaped	Oct. 1996 - Mar. 1997	Ahn & Cheung (1999)
TrtSE	inverted U-shaped	Sep. 1999 - Nov 1999	Vo (2007)

Note: HKSE is short for Hong Kong Stock Exchange.
TrtSE is short for Toronto Stock Exchange.

ume. In response to increased volume, spread widens and depth drops. This is consistent with the prediction of the theoretical model of market microstructure in Easley & O'Hara (1992).

Ahn & Cheung (1999) attribute the reversed U-shaped pattern on depth to the adverse selection problem for limit order traders. Due to high information asymmetry at market open and close, limit order traders suffer high expected adverse selection cost, and thus place passive limit orders in order to protect themselves from informed trades, leading to an inverted U-shaped intraday pattern on market depth, as well as a U-shaped intraday pattern on spread. Li et al. (2005) and Vo (2007) also find limit order traders play an important role in forming the intraday pattern on market depth.

The above review of literature on intraday phenomena shows that intraday patterns have been observed in different markets and in different time periods. However, the shapes of intraday patterns differ across markets and time periods. Thus, it is necessary to provide up-to-date evidences on intraday phenomena as the above literature is dated.

2.5 Trade Execution Strategies

Trade execution strategies are predefined sets of instructions for trade execution, which include how best to break up a large order for execution and the actual mechanisms for placing and managing the orders. When deciding whether or not to submit an order to the market place an algorithm must decide on an order's:

- Submission Time - when should the order be placed?
- Size - what size order should be sent to the market?
- Type - should the order be a market, limit or a hidden order?
- Pricing - at what price should the order be placed?
- Destination - there are many market destinations and types, which one will provide the best conditions of execution for the order?
- Management - if a limit order has been submitted, how should this order be managed post submission?

The literature on trade execution strategies can be classified into two categories: optimal trading strategies and order submission strategies. The first category of studies focuses on how to split a large order in order to minimise price impact cost. The other investigates how to choose order type and order aggressiveness. The following sections review these two streams of research separately.

2.5.1 Optimal Trading Strategy

Bertsimas & Lo (1998) are the first to analyse how to control the implicit execution cost of trading a large order. They argue that large-order execution problem is a dynamic optimisation problem, and form the basic framework to solve for this problem.

Supposing that an investor wants to trade a large order $\overline{S}$ of shares within a fixed finite number of periods T, [1] and denoting S_t as the number of shares traded in period t and P_t as the corresponding execution price, the optimal trade execution problem is how to split the order in order to achieve minimised expected execution cost, namely $\mathbb{E}\left[\sum_{t=1}^{T} P_t S_t\right]$, where $t = 1, ..., T$ and $\sum_{t=1}^{T} S_t = \overline{S}$.. Subsequent studies are largely extensions of their work.

Recent years have seen a burgeoning range of studies on this issue. Some have extended this model to the case of portfolios (e.g., Moazeni et al. (2010), Kharroubi & Pham (2010)). Some have considered the case of FX markets (e.g., Schmidt (2010)). In this thesis only the studies focusing on single asset are reviewed.

The previous work on optimal trading execution differs in the way of modelling price/liquidity dynamics. One stream of research led by Bertsimas & Lo (1998) and Almgren & Chriss (1999) decomposes the price impact into a temporary component and a permanent component. The other stream of research led by Obizhaeva & Wang (2012) and Alfonsi et al. (2010) models the price impact in the limit order book and explicitly models the resilience of price impact. The first approach is denoted by Almgren approach, and the other is denoted by Obizhaeva approach.

Almgren Approach

Within this approach, the design of optimal execution strategies are primarily conducted in the context of markets in which the law of one price holds. Trades have two different types of price impact on the price process of the market. One is

[1] The length of a 'period' can be some fraction of a single day, e.g., 30 minute interval (Bertsimas & Lo, 1998).

permanent price impact which alters the market price permanently. The other is temporary price impact which affect only the execution price of the corresponding trade.

Bertsimas & Lo (1998) assume a linear and permanent price impact function depending on the trade size, with the objective of minimising the expected cost of execution. They derive optimal trade execution strategies using stochastic dynamic programming. In the case when the asset price process follows a zero-drift arithmetic random walk, they show that the optimal execution strategy is to break the trade equally across time. This optimal strategy is independent of the function of permanent price impact and the dynamic of the asset price, because the dynamic of the asset price is only affected by the whole trade size and independent of the divided pieces. Although the optimal execution strategy derived in this work seems counter-intuitive, it provides a basic framework to solve the optimal trade execution problem.

Based on the work of Bertsimas & Lo (1998), Almgren & Chriss (1999, 2000) introduce a temporary price impact function of trade size. The temporary price impact depends on the pace of trades. Fast trading causes larger impact on the dynamic of the asset price than slow trading. In other words, the temporary impact function is used to penalize speedy trades. Thus, the optimal trade execution strategy depends on the dynamic of the asset price, the permanent impact function and the temporary impact function. Their objective function is a mean-variance criterion. With a linear specification of price impact, Almgren & Chriss (1999) show that the optimal execution strategy quickly finishes all trades over the first several periods, and Almgren & Chriss (2000) demonstrate that the optimal strategy becomes more aggressive after the volatility increases, and becomes passive after the volatility decreases.

Although Almgren's model (Almgren & Chriss, 1999, 2000) is tractable, the assumption of a linear price impact function is inconsistent with empirical findings. A number of studies have extended the model of Almgren/Chriss by considering nonlinear price impact functions. With a nonlinear specification of temporary price impact, Almgren (2003) shows that the optimal execution strategy is independent of the scale of trading period and the whole trade size.

Obizhaeva Approach

While Bertsimas & Lo (1998) and Almgren & Chriss (1999, 2000) make computations feasible and lead to relatively nice and robust trading strategies, they do not adequately model the empirically observed transience of price impact.

The price-impact-function based models do not take into account the dynamic property of supply and demand of market liquidity. This property is explicitly obvious in an limit order market, where the order book is able to rebuild itself after being hit by a trade, which is referred to as the *resilience* of the order book.

Obizhaeva & Wang (2012) are the first to model the dynamic supply and demand in a limit-order-book market. They consider a limit order market with an order book which has a constant density, namely the financial securities are uniformly distributed at different levels of the order book. In the case when the price impact of a trade decays exponentially over time, they show that the optimal execution strategy crucially depends on shape of the limit order book and the speed of its resilience.

Alfonsi et al. (2010) extend Obizhaeva & Wang (2012)'s model by allowing for nonuniformly distributed shares within the limit order book. In other words, the density of limit orders is not constant, but varies as a function of price. They demonstrate that trading is discrete and for an optimal purchasing strategy all purchases except the first and last are of the same size. Furthermore, the sizes of the intermediate purchases are chosen so that the price impact of each purchase is exactly offset by the order book resiliency before the next purchase.

A continuous-time generalization of the volume impact version of the model has been introduced by Predoiu et al. (2011). They restrict trading to buy-only or sell-only strategies, so price manipulation is excluded by definition. Predoiu et al. (2011) permit the order book shape to be completely general. All price impact is transient. Assuming exponential decay of price impact, Fruth et al. (2011) analyse the specific form and regularity of optimal order execution strategies when liquidity can be time-dependent or even stochastic. Gatheral (2010) analyse the existence and the behaviour of optimal order execution strategies for certain market microstructure parameters. As shown by Alfonsi & Schied (2010) and further discussed in Gatheral et al. (2011), these results depend strongly on the way in which nonlinearity of price impact is modelled.

2.5.2 Order Submission Strategy

Market and limit orders can be categorised into various aggressivenesses in terms of their quantities and prices (Biais et al., 1995). For example, a buying limit order's level of aggressiveness increases with its limit price and its quantity.

Specifying an order's aggressiveness is a critical decision which a trader has to make when he wants to transact an order. This decision directly affects the quality of trade execution, and indirectly influences the final investment return. An understanding of the dynamic order submission behaviours will contribute to our understanding of the dynamic price formation process.

Market orders and limit orders both have their advantages and disadvantages. Although market orders guarantee order executions, they suffer the cost of paying higher execution prices. While limit orders provide price improvement over market orders, their submissions involve two risks (Handa & Schwartz, 1996). The first risk is *adverse selection risk*, also known as *free trading option risk* (Glosten

& Milgrom, 1985), arising from adverse information events which trigger an undesirable execution. The other risk is *nonexecution risk*, which arises when the market price moves away from the limit price and results in the limit order not executing. The outcome of the second risk is that the limit order traders suffer *opportunity cost* as they have to accept inferior prices.

Empirical studies have shown that different types of market participants have distinguished preferences concerning order aggressiveness. Keim & Madhavan (1995a) find that value investors are more likely to use limit orders instead of market orders. Aitken et al. (2007) find that limit orders placed by proprietary trading desks and hedge funds are more aggressive than mutual funds, index funds and insurance companies. While Glosten (1994) and Seppi (1997) argue that informed traders are more likely to use market orders, Bloomfield et al. (2005), Chakravarty & Holden (1995), Kaniel & Liu (2006), Harris (1998) and Wald & Horriga (2005) suggest that informed traders use limit orders more often than market orders.

The execution performance of aggressive and passive orders have been investigated in the previous studies. Generally, passive orders outperform aggressive orders in terms of investment return (Handa & Schwartz, 1996) and trading cost (Harris & Hasbrouck, 1996). Aggressive orders have larger price impact but smaller opportunity cost than passive orders (Griffiths et al., 2000). Moreover, the execution performance of institutional traders and individual traders is compared by Anand et al. (2005), which finds that aggressive orders placed by institutional traders outperform those placed by individual traders.

Order Aggressiveness and State of The Order Book

The state of the order book is a critical determinant of traders' order aggressiveness. This assertion has been suggested by recent theoretical developments (Foucault, 1999; Foucault et al., 2005; Goettler et al., 2005; Handa et al., 2003; Liu, 2009; Menkhoff et al., 2010; Parlour, 1998; Rosu, 2009), and is supported by many studies in different markets and over different sample periods. The rest of this section discusses how various order-book related variables affect the aggressiveness of incoming orders.

Bid-ask Spread

Bid-ask spread has a positive relationship with incoming order's aggressiveness (Al-Suhaibani & Kryzanowski, 2000; Beber & Caglio, 2005; Bae et al., 2003; Biais et al., 1995; Cao et al., 2008; Duong et al., 2009; Ellul et al., 2007; Griffiths et al., 2000; Hall & Hautsch, 2007; Pascual & Verdas, 2009; Ranaldo, 2004; Verhoeven et al., 2004; Xu, 2009). A decrease in the bid-ask spread makes the use of aggressive orders more attractive for incoming traders, and an increase in the spread makes the use of passive orders more tempting. However, this relationship is influenced by the proportion of patient traders in the trading population (Foucault et al., 2005). If patient (impatient) traders dominate the trading population, more (less) aggressive orders are submitted than passive orders when the spread is wide.

Volatility

Prior empirical research is inconclusive on the effect of transient volatility on order aggressiveness. Most studies (Ahn et al., 2001; Beber & Caglio, 2005; Bae et al., 2003; Chung et al., 1999; Hall & Hautsch, 2007; Ranaldo, 2004; Verhoeven et al., 2004) support an inverse relation between order aggressiveness and volatility: when market volatility increases, passive orders are more attractive than aggressive orders for incoming traders. Ahn et al. (2001) examine the relationship between order placement and the transitory volatility from different sides of the markets, and find that more buy (ask) limit orders are placed than buy (ask) market orders when the volatility arises from the bid (ask) side.

This inverse relation can be explained by the high probability of mispricing an asset in a volatile market which thus raises the cost of market order trading (Foucault, 1999). In other words, higher price volatility means a greater opportunity of execution an order at a better price. Consistently, Lo et al. (2002) show that in periods of high volatility, the execution probability of limit orders increases, which makes passive orders more tempting than aggressive orders.

Other studies provide different results. A positive relationship between order aggressiveness and volatility is reported in Hasbrouck & Saar (2002); Lo & Sapp (2010); Wald & Horriga (2005). This positive relation is only documented for institutional traders in large cap stocks in Duong et al. (2009). Hall & Hautsch (2006) observe an increase of all kinds of order submission during periods of high volatility. Cao et al. (2008) find that volatility has a minimal effect on order aggressiveness. Pascual & Verdas (2009) find that the relationship is affected by market capitalisation: higher historic volatility suggests limit order submission in mid-cap stocks, but the opposite phenomenon is observed in large-cap stocks.

Some theoretical models (Foucault, 1999; Handa & Schwartz, 1996) of limit order market have predicted that price volatility determines the choice between market and limit orders. Foucault (1999) and Handa & Schwartz (1996) predict that market order is less attractive than limit orders and thus more limit orders are submitted to the market than market orders when price is volatile. They attribute this to asymmetric information between investors. When the market is volatile, the probability of trading against informed investors increases, leading to higher expected loss.

Market Depth

Market depth has also been found to matter. The limit order book at the best quotes generates a *crowding out* effect which affects subsequent order submissions (Parlour, 1998). The incoming trader is more likely to submit a market order if the book has a deep depth on the same side, and is more likely to submit a less aggressive order if the book is deep on the opposite side (Al-Suhaibani & Kryzanowski, 2000; Beber & Caglio, 2005; Cao et al., 2009; Duong et al., 2009; Ellul et al., 2007; Foucault, 1999; Griffiths et al., 2000; Hall & Hautsch, 2007; Omura et al., 2000; Parlour, 1998; Peterson & Sirri, 2002; Potters & Bouchaud, 2003; Ranaldo, 2004; Verhoeven et al., 2004; Xu, 2009) . Especially when both sides of the book are thick, 'fleeting' limit orders may happen (Rosu, 2009): some limit orders placed inside the spread are immediately accepted by some traders from the other side of the book. These studies reveal that time-to-execution is a significant concern of limit order traders when execution cost less matters.

In contrast, depth beyond the best quotes has an inverse relation with order aggressiveness (Goettler et al., 2005). Higher depth above the best ask signals that the best ask may be too low compared with the fundamental value of the asset, so that incoming buyers are more likely to use aggressive orders and incoming sellers are more likely to use passive orders, and vice versa for the case of higher depth below the best bid.

Eisler et al. (2012) show that market orders on one side of the book attract compensating limit orders, leading to smaller conditional impact of subsequent market orders on the same side.

Time of the Day

Order submission strategy is also affected by time of the day (Biais et al., 1995; Ellul et al., 2007). The placement of less aggressive orders tend to be concentrated in the morning, whereas more aggressive orders tends to occur late in the trading day. Passive orders are more likely to happen at the end of the day than in an earlier period (Foucault et al., 2005), because spread improvement via the use of limit order is small.

Controlling price impact is important in financial markets. An increasing number of studies have been denoted to this problem. However, these studies, as reviewed in Section 2.5.1, only adopt market orders, although limit orders are equally important when traders make decisions on order choice as shown in Section 2.5.2. It is still unknown whether order choice plays an important role in determining price impact when trading a large order.

2.6 Summary

The aim of this chapter was to provide a brief overview of market microstructure literature and a review of relevant work to this thesis.

Section 2.2 provided a brief overview of the literature on market microstructure. Market microstructure concerns every aspect of the trading process. Theoretical models on market microstructure can be classified into three categories: inventory-based models, information-based models, and limit-order-book models. As more and more markets adopt limit order books, the number of studies on the trading process in limit order markets is significantly increasing in recent years.

Section 2.3 offered a review of the literature on price impact. Most prior studies on price impact have been devoted to the relationship between price impact caused by market orders and their trade sizes. A nonlinear relationship was discovered by these studies. However, the order size does not tell the whole story about price impact. Recently, a number of studies have found that most of the extreme price changes are not caused by large trading size, but by the gaps between different levels of the order book.

Section 2.4 reviewed the literature on intraday phenomena in financial markets. Previous studies concentrate on four variables measuring market liquidity and trading activity, which are price volatility, bid-ask spread, trading volume and market depth. A number intraday patterns with different shapes were observed in prior literature and were summarised in this section.

Section 2.5 reviewed the literature on trade execution strategies. This review focuses on the literature concerning large-order trading strategies and order submission strategies. The studies on large-order trading strategies differ in how to model price impact. One stream of studies decomposes price impact into a temporary component and a permanent component, while the other stream specifically considers price impact in the limit order book and models both the immediate effect of trading on the order book and the resilience of price impact. Most studies on order submission strategies focus on the relation between order aggressiveness and the state of the order book.

The review in this chapter identifies a number of research gaps. These are summarised as follows:

1. **Analysis of the intraday behaviour of price impact**

   It is unknown about up-to-date intraday behaviour of price impact. An analysis of intraday behaviour of price impact is of great interest to both academics and practitioners. It is helpful for developing a more comprehensive price impact model than those reviewed in Section 2.3.1, as well as the design of trade execution strategies as reviewed in Section 2.3.2. Chapter 5 conducts such analysis using the data drawn from the NYSE-Euronext TAQ database and the LSE ROB database.

2. **Examination of the relation between agent intelligence and price impact**

   No study has been conducted to investigate whether agent intelligence plays an important role in determining the magnitude of price impact. This examination is of great importance to understand the driving factors of the concavity of price impact. An examination of this is undertaken in Chapter 6.

3. **Investigation of the effect of order choice on price impact of trading large orders**

It is unknown that whether order choice plays an important role in determining price impact when trading a large order. This gap is addressed in Chapter 7.

The next chapter will present the methodology adopted in two experimental studies of this thesis.

Chapter 3

Agent-based Modelling in Finance

> "... the possibility of building a mathematical theory of a system or of simulating that system does not depend on having an adequate microtheory of the natural laws that govern the system components. Such a microtheory might indeed be simply irrelevant."
>
> Nobel laureate in Economics: Simon (1996) (p. 19)

This thesis adopts an agent-based modelling approach in two of its experimental chapters. This chapter reviews studies on agent-based modelling in financial markets, and describes the core elements of agent-based market simulation as well as the Genetic Algorithm as a learning algorithm in agent-based modelling.

Agent-based modeling (ABM) is a computer modelling approach to construct complex systems consisting of a number of decision makers (agents) and institutions interacting through prescribed rules (Farmer & Foley, 2009). It creates a virtual universe in which agents act in complex and realistic ways.

The potential scope of applications of ABM is wide. Economics and finance are among the most promising areas where ABM can be applied. Financial markets are always considered as complex systems, which are constantly evolving and becoming more complex, and thus difficult to analyse and understand. ABM has provided a powerful tool to analyse market behaviours and the effects of the market mechanism in financial markets. The applications of ABM to financial markets are vast. The review here can not be exhaustive. For general reviews, see for example Chen (2007, 2012); Chen et al. (2012); Chiarella et al. (2009a); Consiglio (2007); Cristelli et al. (2011); Hommes & Wagener (2009); LeBaron (2000, 2001, 2006); Lux (2009); Samanidou et al. (2007); Tesfatsion (2006).

The rest of this chapter is organised as follows. Section 3.1 briefly reviews the literature on agent-based financial markets, followed by a description of the core elements of agent-based market simulation in Section 3.2. An introduction to a learning algorithm in agent-based modelling, the Genetic Algorithm, is given in Section 3.3. A summary of this chapter is provided in Section 3.4.

3.1 Agent-based Financial Markets

Agent-based modelling has been applied to economics and finance since 1971 (Schelling, 1971). As defined by Tesfatsion (2006), agent-based computational economics is "the computational study of economic processes modelled as dynamic systems of interacting agents". It offers an alternative approach to modelling and studying financial markets. It differs from the traditional equilibrium models which can not always provide analytical solutions (Parlour & Seppi, 2008). Agent-based models of financial markets can help us to understand how markets operate and can guide policy formulation.

Agent-based markets have been used to study price bubbles and market crashes in a number of studies, like Feldman & Friedman (2010), Kim & Markowitz (1989), Miller (2008) and Palmer et al. (1994). The Nobel laureate Harry Markowitz (Kim & Markowitz, 1989) is among the earliest researchers using an agent-based market simulation. Their study proposes an agent-based simulation to explain the 1987 crash. The agent-based artificial market with agents that learn their trading behaviours using genetic algorithm in Palmer et al. (1994) exhibit financial bubbles. Feldman & Friedman (2010) investigate how human traders and robot traders behave and interact in a market prone to bubbles and crashes. They find that human traders have little impact on market crashes as they are more experienced. Miller (2008) finds that the bubbles and crashes seen in experiments conducted with robot traders are much less drastic than those with human subjects. This difference is caused by the restriction that robot agents are prohibited from engaging in actions that could result in a loss and are thus not allowed to speculate.

Some studies have used agent-based financial markets to study the statistical regularities, often know as *stylised facts*, observed in real markets. Examples of stylised facts are fat tails of price return and volatility clustering (Cont, 2001). An foreign exchange market with chartist and fundamentalist agents is simulated in Lux (1998) and exhibits fat tails of price return. LiCalzi & Pellizzari (2003) produces a leptokurtic distribution of short-term log-returns in an artificial order-driven market consisting of fundamentalist and Zero Intelligence (ZI) value-based agents under budget constraints. Their results support the conjecture in Bouchaud et al. (2002) and Daniels et al. (2003) that the emergence of some of the statistical properties of order-driven markets is mostly due to their microstructure. Chiarella & Iori (2002) produce volatility clustering in an artificial markets composed of fundamentalist, chartist and noise traders. Chiarella et al. (2009b) show that the chartist strategy is mainly responsible of the fat tails and clustering in their artificial markets. Their results offer evidence that large price impacts are mostly caused by the presence of large gaps in the order book.

In addition, ABM has been applied to other problems in finance, like tests of economic theories (Chen & Yeh, 2002), market design (Darley & Outkin, 2007), option pricing (Suzuki et al., 2009), evaluation of automated trading strategies (Izumi et al., 2009), and analysis of liquidity costs (Huang et al., 2012). Chen & Yeh (2002) use an agent-based stock market with a number of evolving agents to test both the efficient market hypothesis and the rational expectations hypothesis, and show that their agent-based model can replicate some economic behaviours empirically. Darley & Outkin (2007) examine the effect of the tick-size reduction in the NASDAQ market using an agent-based artificial market. Suzuki et al. (2009) analyse implied volatility smile and the skewness premium using an agent-based modelling approach. Izumi et al. (2009) use an artificial market to evaluate the risks and returns of automated trading strategies in various market environments and test the market impact of the trading strategies. Huang et al. (2012) report a higher liquidity costs of market orders in their agent-based order-driven stock market than in the Taiwan Stock Market.

3.2 The Core Elements of Agent-based Market Simulation

The core elements of constructing an agent-based artificial market are

- *agents*,
- *market mechanism*.

3.2.1 Agents

Agents are the essential elements of an agent-based market (LeBaron, 2001). The agents can range from "active data-gathering decision-makers with sophisticated learning capabilities to passive world features with no cognitive functioning", according to Tesfatsion (2006). Chen (2012) undertakes a comprehensive review of various types of agents designs.

Intelligent Agents

At one end of the spectrum are intelligent agents who can learn, adapt and evolve. A variety of techniques from artificial intelligence have been used to model agent learning. The celebrated example is the Sante Fe Institute Artificial Stock Market (SFI ASM) (Arthur et al., 1997). In their model, agents are heterogeneous in the way they form their expectations. Each agent, possessing a number of linear forecasting models at any time, learns and evolves by discovering which forecasting model works best as well as developing new ones from time to time, via an inductive algorithm, namely the *Genetic Algorithm*. They show that both herd effects and systematic speculative profits are possible in an endogenous-expectations market.

Zero-intelligence Agents

The zero-intelligence (ZI) agents are the other extreme of the spectrum of agent types. Zero-intelligence modelling was originally pioneered by the Nobel laureate Becker (1962), who showed that some aspects of supply and demand curves could be understood without any reliance on strategic thinking. Later, it was popularised by Gode & Sunder (1993). As characterised in the work of Gode & Sunder (1993), these agents are *random behaving agents* who randomly generate buying and selling orders subject to budget constraints. The orders submitted by impatient traders are executed against patient orders previously submitted to the market. The patient orders are placed in the order book according to price and time priorities. Despite their simplicity, ZI agents are able to get a remarkable allocation efficiency in the artificial double auction market. A good review of ZI agents based artificial markets is given by Iori et al. (2003); Ladley & Schenk (2009); Zovko & Farmer (2002).

Despite their simplicity, the zero-intelligence artificial markets are able to generate many non-trivial behaviours seen in real markets. Bak et al. (1997) show that their zero-intelligence model is able to reproduce phenomena seen in actual markets, including larger than Gaussian fluctuations at short time scales and some of the power-law forms of price variations. The zero-intelligence model in Maslov (2000) can produce approximate several regularities observed in actual markets, including the autocorrelation of the absolute value of price changes and short-range correlations in the signs of price movements, as well as the fat tails of price changes.

3.2.2 Trading Mechanism

Trading Mechanism is a critical design decision which the builder has to consider when creating an agent-based financial market. It relates the agents' trading demands represented by their orders and the prices of the market.

Most artificial financial markets implement simplified market mechanisms, which omit some institutional details of trading, but serve their research needs sufficiently. One example is the SFI market (Arthur et al., 1997), which implements a simple market-clearing mechanism. At the end of each trading period, the market specialist collects the accumulated buy and sell orders of all the agents, and clears the market at a new market price. If the demand excesses the supply, the specialist increases the market price, otherwise he decreases the price.

A number of studies have implemented a fairly realistic market, which incorporates explicit market microstructure, like a continuous double auction mechanism with limit order books. As pointed out in LeBaron (2001), this type of trading mechanism is "most appealing for modelling high-frequency data" and is only easy to implement when there is no human intervention (like specialists or market makers) in the price formation process. All ZI artificial markets replicating a limit order market are of this type. Some examples are Chan et al. (2001); Raberto & Cincotti (2005); Yang (2002).

The next section presents the Genetic Algorithm, which is often used to simulate agent learning in agent-based modellings.

3.3 Genetic Algorithm

The genetic algorithm (GA) is an optimisation algorithm, founded upon the evolutionary principle, "survival of the fittest". It provides a systematic search process directed by performance feedback. In financial markets, GA methodologies have been successfully used for solving a broad selection of problems, ranging from predicting movements in current values to optimizing equity portfolio composition. Examples of applications of GA in economics and finance are Allen & Karjalainen (1999), Arifovic (1995), Arifovic (1996), Brabazon et al. (2008), Chavez-Demoulin & McGill (2012), Chen & Yeh (1997), Dempster & Jones (2001), Dworman et al. (1996), Kozhan & Salmon (2011), Neely et al. (1997), Neely & Weller (1999), Neely & Weller (2001), Neely & Weller (2003), and Sefiane & Benbouziane (2012).

The evolutionary processes in GA represent an archetype, whose application transcends their biological root. In biological evolution, species are positively or negatively selected depending on their relative success in surviving and reproducing in the environment. Differential survival, and variety generation during reproduction, provide the engine for evolution (Darwin, 1859; Spencer, 1864) (Figure 3.1).

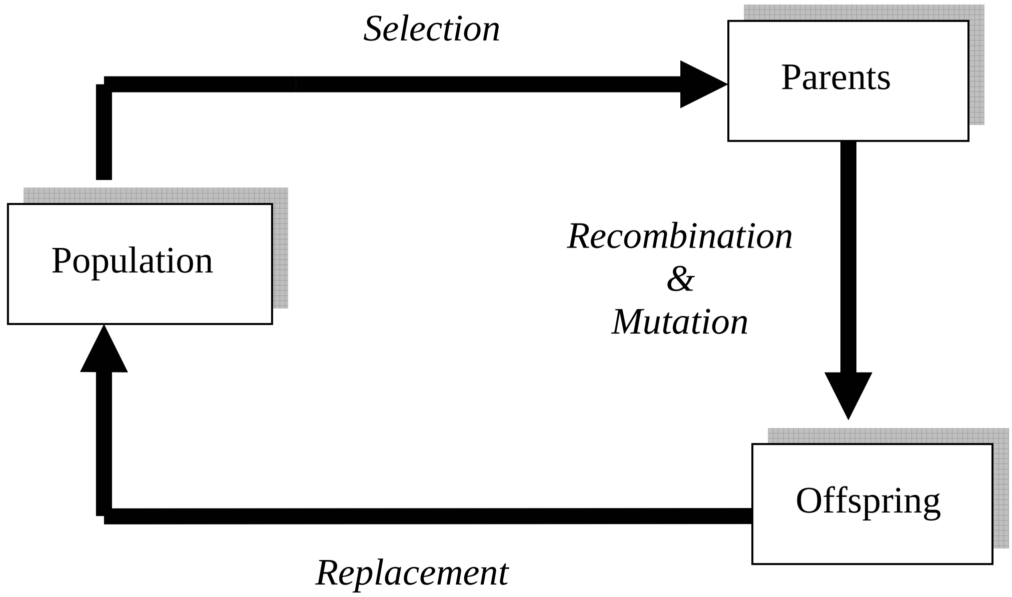


Figure 3.1: Evolutionary Cycle

These concepts have metaphorically inspired the field of evolutionary computation (EC). Algorithm 1 outlines the evolutionary meta-algorithm. There are many ways of operationalising each of the steps in this meta-algorithm, consequently, there are many different, but related, evolutionary algorithms (EA). Just as in biological evolution, the selection step is a pivotal driver of the algorithm's workings. The selection step is biased in order to preferentially select better (or 'more fit') members of the current population. The generation of new individuals creates children which bear some similarity to their parents but are not identical to them. Hence, each individual represents a trial solution in the environment, with better individuals having increased chance of influencing the composition of individuals in future generations. This process can be considered as a 'search' process, where the objective is to continually improve the quality of individuals in the population. In financial markets, EC methodologies including Genetic Algorithm

(GA), Genetic Programming (GP) and Grammatical Evolution (GE), have been used for solving a broad selection of problems, ranging from prediction, asset selection, portfolio optimization, derivatives pricing and credit risk assessment. An overview of some EC applications in finance can be seen in Brabazon & O'Neill (2006) and Brabazon et al. (2008).

```
Initialise the population of candidate solutions;
repeat
    Select individuals (parents) for breeding from the current population;
    Generate new individuals (children) from these parents;
    Replace some / all of the current population with the newly-generated individuals;
until terminating condition;
```

Algorithm 1: Evolutionary Algorithm

3.3.1 Grammar-based Genetic Algorithm

The grammar-based genetic algorithm, which is Grammatical Evolution (Dempsey et al., 2009; O'Neill & Ryan, 2003), is an evolutionary methodology, and can be used to evolve structures, programs or 'rule sets'. These rule sets can be as general as a functional expression which produces a good mapping between a series of known input-output data vectors.

GE is an Evolutionary Automatic Programming (EAP) technique which allows the generation of computer programs (or 'rule sets') in an arbitrary language. GE can conduct an efficient exploration of a search space, and notably permits the incorporation of existing domain knowledge in order to generate 'solutions' with a desired structure. In finance (for example), this allows the users to seed the evolutionary process with their current trading strategies in order to see what improvements the evolutionary process can uncover. Recently GE has been successfully applied to a number of financial problems. These include financial time series modelling, intraday financial asset trading, corporate credit rating, and the uncovering of technical trading rules (Brabazon & O'Neill, 2006).

A particular strength of the methodology is that the form of the model need not be specified *a priori* by the modeler. This is of particular utility in cases, such as in this study, where there is a theoretical or intuitive idea of the nature of the relevant explanatory variables, but a weak understanding of the functional relationship between the explanatory and the dependent variable(s). GE does not require that the model form is linear, nor does the method require that the measure of model error used in model construction is a continuous or differentiable function. Another useful feature of a GE approach is that it produces human-readable rules that have the potential to enhance understanding of the problem domain. Interestingly, GP/GE methods enjoy wide application beyond finance (O'Neill & Brabazon, 2009), and routinely produce human-competitive performance, with some solutions being patentable in their own right (Koza, 2010).

Genotype-phenotype Mapping

A genotype-phenotype mapping is employed such that each individual's variable length binary string, contains in its codons (groups of 8 bits) the information to select production rules from a Backus Naur Form (BNF) grammar. GE is therefore a grammar-based approach to GP (McKay et al., 2010). The user can tailor the grammar to produce solutions that incorporate domain knowledge by biasing the grammar to produce very specific forms of sentences. BNF is a notation that represents a language in the form of production rules. It is comprised of a set of non-terminals that can be mapped to elements of the set of terminals (the primitive symbols that can be used to construct the output program or sentence(s)), according to the production rules. A simple example BNF grammar is given below, where <expr> is the start symbol from which all programs are generated. The grammar states that <expr> can be replaced with either <expr><op><expr>

or <var>. An <op> can become either +, −, or ⋆, and a <var> can become either x, or y.

```
<expr> ::= <expr><op><expr>        (0)
         |  <var>                  (1)
  <op> ::= +                       (0)
         |  -                      (1)
         |  *                      (2)
 <var> ::= x                       (0)
         |  y                      (1)
```

The grammar is used in a developmental process to construct a program by applying production rules, selected by the genome, beginning from the start symbol of the grammar. In order to select a production rule in GE, the next codon value on the genome is read, interpreted, and placed in the following formula:

$$Rule = Codon\ Value\ Mod\ Num.\ Rules$$

where Mod represents the modulus operator. Given the example individual's genome (where each 8-bit codon has been represented as an integer for ease of reading) in Fig.3.2, the first codon integer value is 20, and given that there are 2 rules to select from for <expr> as in the above example, $20\ Mod\ 2\ =\ 0$ is obtained. <expr> will therefore be replaced with <expr><op><expr>.

Beginning from the left hand side of the genome codon integer values are generated and used to select appropriate rules for the left-most non-terminal in the developing program from the BNF grammar, until one of the following situations arise:

- A complete program is generated. This occurs when all the non-terminals in the expression being mapped are transformed into elements from the terminal set of the BNF grammar.

- The end of the genome is reached, in which case the *wrapping* operator is invoked. This results in the return of the genome reading frame to the left hand side of the genome once again. The reading of codons will then continue unless an upper threshold representing the maximum number of wrapping events has occurred during this individual's mapping process.

- In the event that a threshold on the number of wrapping events has occurred and the individual is still incompletely mapped, the mapping process is halted, and the individual assigned the lowest possible fitness value.

Returning to the example individual, the left-most `<expr>` in `<expr><op><expr>` is mapped by reading the next codon integer value 124 and used in $124\ Mod\ 2\ = 0$ to become another `<expr><op><expr>`. The developing program now looks like `<expr><op><expr><op><expr>`. Continuing to read subsequent codons and always mapping the left-most non-terminal the individual finally generates the expression `y*x-x-x+x`, leaving a number of unused codons at the end of the individual, which are deemed to be introns (unused codons) and simply ignored.

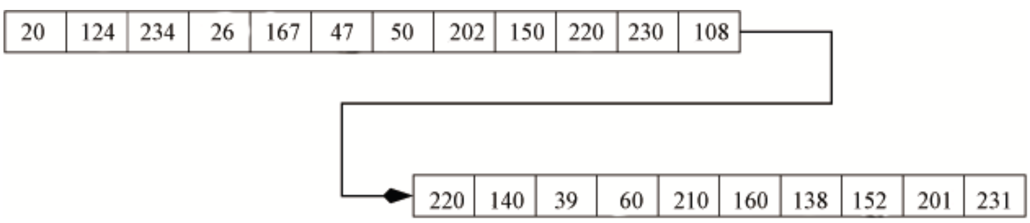


Figure 3.2: An Example GE Individual's Genome Represented as Integers for Ease of Reading.

In the context of this study, GE is used to evolve trade execution strategies.

A population of strategies is maintained and iteratively improved via a simulated evolutionary process. The structure of these rules is governed by a choice of grammar and the utility of evolved strategies is assessed by testing them in an artificial stock market environment. Both the choice of grammar and the design of the stock market environment are discussed in later chapters.

3.4 Summary

This chapter presented the ABM methodology and briefly reviewed its applications to the area of finance. A description of the three essential components for constructing agent-based artificial markets is also given. Moreover, a popular learning algorithms, the Genetic Algorithm, was introduced. Although ABM has been applied to economics and finance for a long time, no previous studies have investigated price impact using ABM.

In Part II of this thesis, the three empirical studies are presented. Chapter 4 initially provides a description of the data used and of the necessary data preprocessing steps.

Part II

Experiments

In Part II, several empirical investigations of price impact are presented. The literature review in Part I has identified a number of research gaps concerning our understanding of intraday behaviour of price impact, whether agent intelligence affects the magnitude of price impacts, and the effects of order choice on the price impact of trading large-orders.

The analysed data is drawn from the NYSE-Euronext TAQ database and the LSE ROB database. In Chapter 4, the data is introduced and a description of the data preprocessing steps is given.

Three studies are presented in Chapters 5-7 to address the three research questions presented in Chapter 1. Chapter 5 studies the intraday behaviour of price impacts and market liquidity using TAQ data and the ROB data. Chapter 6 presents a study investigating whether agent intelligence determines the magnitude of price impact. In Chapter 7, the effect of order choice on the price impacts of trading large-orders is examined.

Chapter 4

Datasets and Data Processing

Parts of this chapter have appeared in:

- Cui, W. and Brabazon, A. (2012). An Analysis of Liquidity in London Stock Exchange. *Proceedings of the 25th Irish Accounting and Finance Association Conference*, pp. 1-45, Galway, Ireland. (Cui & Brabazon, 2012b)
- Cui, W. and Brabazon, A. (2011). Ultra High-frequency Financial Data: Analysing NYSE Euronext TAQ Data. *Proceedings of the 24th Irish Accounting and Finance Association Conference*, pp. 1-38, Cork, Ireland. (Cui & Brabazon, 2011)

More than ever before researchers today face the challenges of working with high-frequency datasets from financial markets. Examples of this data include NYSE's TAQ data, and LSE's ROB data. These datasets are very large and very difficult to process. The data records all visible order activities, i.e. limit order submissions, cancellations and executions, and can be used to reconstruct the historical limit order book up to any required precision. However, this creates a challenge as it is necessary to reconstruct the limit order book using the same rules that were used by the matching algorithm applied by the exchange. Although simple in principle, such algorithms need to take into account market-specific issues and considering the large volume of data, work is extremely difficult.

The previous two chapters conducted a literature review in the field of market microstructure relating to the work in this thesis. This chapter gives an introduction to the two datasets used in this thesis. One is the NYSE Euronext trade & quote dataset (TAQ), and the other is the LSE rebuild order book dataset (ROB). Both of these are intraday high-frequency data, recording information on both trades and quotes throughout each trading day. The ultra high-frequency data provided by NYSE-Euronext is unfiltered and contains errors and noise. Analysing the raw data can produce inaccurate results which lead to imprecise conclusions. Filtering the raw NYSE-Euronext data is important before data analysis. Likewise, the ultra high-frequency data provided by LSE is also raw data, for example, the original LSE data is out of time sequence. Preprossing the raw LSE data is necessary before data analysis.

The remainder of this chapter is organised as follows. Section 4.1 introduces the datasets. Section 4.2 discusses the issues in data processing and gives the steps used to process the data. A summary of this chapter is provided in Section 4.3.

4.1 Datasets

This section presents descriptions of the two databases, namely the NYSE Euronext TAQ database and the LSE ROB database.

4.1.1 NYSE Euronext TAQ Database

Source of the Data

The NYSE Euronext Trade and Quote (TAQ) database contains intra day information on trades and quotes for equity securities listed on most of the major US exchanges. The database commenced in 1993, and the information contained is in turn drawn from the 'Consolidated Tape' which takes an information flow on real-time trades and quotes from all the participating markets and 'consolidates' this to provide traders with a complete picture of trades and quotes across all markets. The current participants in the system include all major US equity trading venues (further background information on these venues is provided in Sect. 4.1.1 below).

Technically, there are two discrete elements of the consolidated tape, the Consolidated Tape Plan, which governs trades and the Consolidated Quotation Plan, which governs quotes. Since the late 1970s, all SEC-registered exchanges and market centers that trade NYSE or AMEX-listed securities send their trades and quotes to a central consolidator where the Consolidated Tape System (CTS) and Consolidated Quote System (CQS) data streams are produced and distributed worldwide.

For investors and traders, the CQS provides *pre-trade transparency* as it displays the best bid and offer price and volume (and all revisions of quotes), posted by specialists (NYSE, AMEX) and by market makers (NASDAQ), for all stocks for a particular security on each exchange on which it trades, allowing investors and traders to decide how best to route their orders. In contrast, the CTS provides *post-trade transparency* allowing investors and traders to compare the execution price of their trades against that of other trades in the market place which took place at the same time.

The way that trade and quote information reaches the CTS and CQS have changed over the years, moving from largely manual systems in traditional stock exchanges (such as use of floor reporters by the NYSE until 2001) with the attendant risk of delay and errors in recording,[1] to largely automated recording of this data in modern markets. As noted above, the TAQ database is extracted from the consolidated tape and consequently includes all transactions data, time-stamped to the nearest second, reported during the Consolidated Tape hours of operation (currently 4am EST until 6.30pm EST). The data does not include information on the identity of the trading parties or whether the trade is buyer or seller initiated.

TAQ data (along with other tick data products) is available for purchase from NYSE Euronext in a number of delivery formats. The academic research version of the TAQ data is supplied via monthly data DVDs. In the following sections, the scale of trade and quote activity for a number of equities will be illustrated, but in advance of that, it is worth noting the large number of equity securities that are covered by the consolidated tape (as at 27 January 2011, there were 3,243 firms listed on the NYSE, 2,875 listed on Nasdaq and 551 listed on AMEX).

Accessing the Data

The academic version of the data is available for purchase approximately four weeks after the end of each month. The data is accessible via a supplied front-end extraction program. Figure 4.1 illustrates the first screen wherein the user is able to select the desired date ranges and trade/quote information. A number of other screens (including Figure 4.2) allow the user to select the relevant ticker codes and exchanges of interest, along with a choice of data output file format. Users

[1] For example, using data drawn from July 1994 to June 1995, Blume and Goldstein Blume & Goldstein (1997) reported a median delay of 16 seconds between the execution and the consequent reporting of NYSE trades, with Peterson and Sirri Peterson & Sirri (2003) reporting a median delay of 2 seconds using 1997 NYSE data from the NYSE System Order database File.

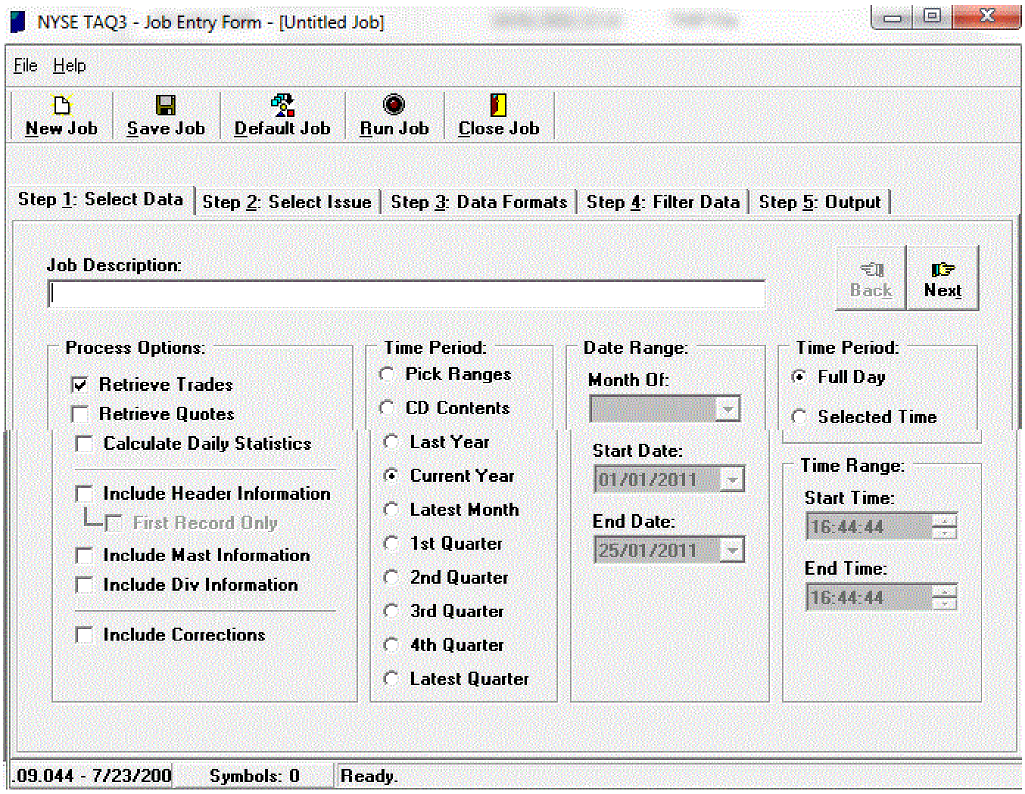


Figure 4.1: Data Selection Screen for TAQ3 Extraction Tool

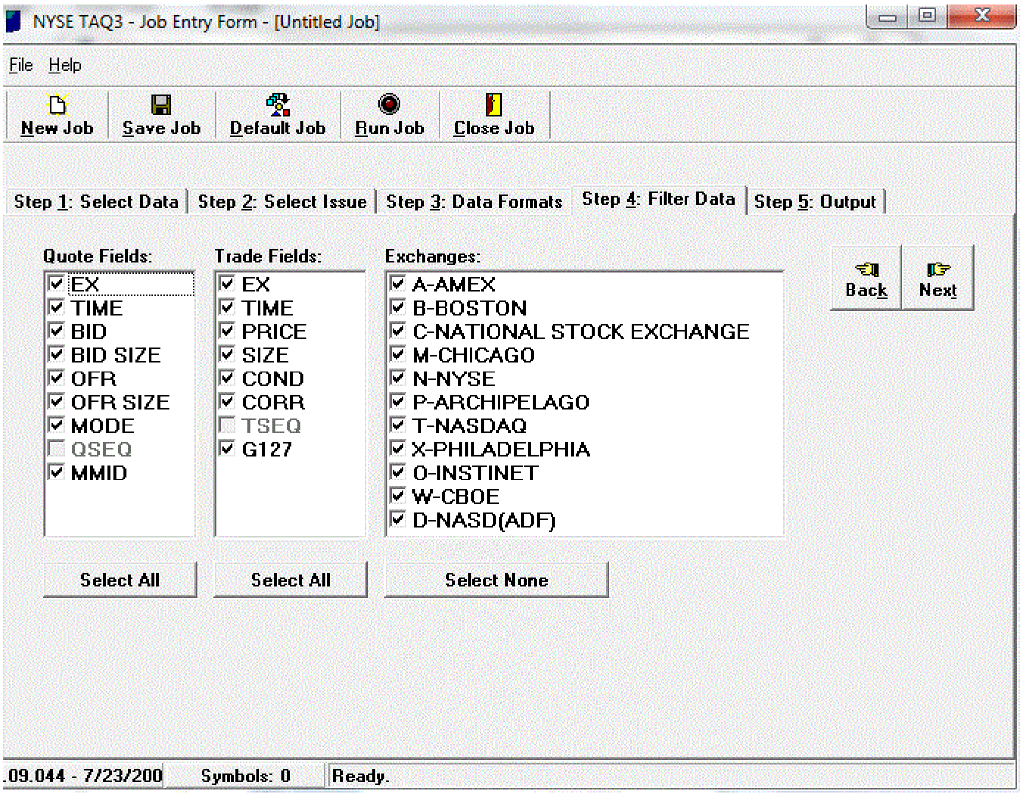


Figure 4.2: Data Filter Screen for TAQ3 Extraction Tool

can also access the raw data files directly via custom written scripts.

Sample Data

A sample of the trade data available is provided in Table 4.1. Starting from the left-most field, there are the stock ticker code (here 'F' for Ford), the exchange on which the trade occurred (here 'N' for NYSE), the date and time of the trade, the trade price and trade size, and a series of data flags which indicate the trade condition, the correction indicator and a flag which indicates whether the trade is a 'G' trade or a rule 127 transaction.[2]

Table 4.1: A Sample of Trade Data for Ford Drawn from 2/2/09

Ticker	Exchange	Date	Time	Price	Size	Cond	Corr	G127
F	N	2/2/09	11:20:48	1.8700	100		0	0
F	N	2/2/09	11:20:51	1.8800	100		0	0
F	D	2/2/09	11:20:54	1.8784	1000		0	0
F	D	2/2/09	11:20:55	1.8800	3000		0	0
F	Z	2/2/09	11:20:55	1.8800	100		0	0
F	D	2/2/09	11:21:04	1.8700	145		0	0

The quote information indicates the best (quoted) bid / offer available on an exchange at a point in time. This changes with each alteration of the best bid/offer or any changes to the number of shares on offer at either of these prices. As would be expected, a very large quantity of quote information is generated each day for the more actively traded stocks. Taking the sample information illustrated in Table 4.2, the leftmost column indicates the ticker code, followed by the exchange generating the quote ('C' = Cincinnati,[3] 'T'=Nasdaq, 'N'=NYSE etc.), date and time of quote, best bid/offer, and the volume (in 100 share lots) at each of these prices.

[2]"G" trade: A member firm trading for its own account must publicly identify that the order is principal. Rule 127: An NYSE trade reported as having been executed as a block position.

[3]Known as the National Stock Exchange since 2003 and headquartered in New Jersey.

Table 4.2: A Sample of Quote Data for Ford Drawn from 2/2/09

Ticker	Exchange	Date	Time	Bid	Size	Offer	Size	Mode	MMID
F	C	2/2/09	11:20:56	1.8700	175	1.8800	75	12	
F	T	2/2/09	11:20:56	1.8700	458	1.8800	597	12	
F	C	2/2/09	11:20:56	1.8700	170	1.8800	75	12	
F	N	2/2/09	11:20:56	1.8700	425	1.8800	162	12	
F	T	2/2/09	11:20:56	1.8700	458	1.8800	517	12	
F	N	2/2/09	11:20:56	1.8700	426	1.8800	162	12	
F	I	2/2/09	11:20:58	1.8700	1005	1.8800	555	12	
F	N	2/2/09	11:20:58	1.8700	426	1.8800	182	12	

TAQ User Guide

The TAQ 3 User's Guide NYSE (2006) provides a description of the operation of the extraction software and the associated meanings for the possible values in each field. However, it should be noted that the guide has not been updated since 30 October 2006 and hence a data user should examine the expanded list of field values for both the quotes and trades files which can be found on http://www.nyxdata.com/Data-Products/Daily-TAQ. For example, the range of trade venues participating in the consolidated tape and the range of related exchange identifier codes in the TAQ database has increased since 2006. The current listing is shown in Table 4.3.

As can be seen from the listing (shown in Table 4.3), the US equity market is spread across a large number of trading venues. Recent years have seen a consolidation and a renaming of some of these trading venues as discussed below.

The main variants of the NYSE are the traditional NYSE floor (now a hybrid of an auction and an automated market) and NYSE Arca (formerly known as the Archipelago Exchange and the Pacific Exchange). NYSE Arca is a fully electronic stock exchange. The American Stock exchange (historically known as the New York Curb Exchange) was acquired by the NYSE in 2008 and is now known as NYSE Amex Equities.

Nasdaq (formerly, the 'National Association of Securities Dealers Automated

Table 4.3: Exchange Codes as at 31 January 2011

Code	**Exchange**
A	American Stock Exchange
B	Boston Stock Exchange (now Nasdaq OMX BX)
C	National (Cincinnati) Stock Exchange
D	National Association of Securities Dealers (ADF)
E	Market Independent (SIP - Generated)
I	ISE (International Securities exchange)
M	Chicago Stock Exchange
N	New York Stock Exchange
P	ARCA
T/Q	NASDAQ Stock Exchange
S	Consolidated Tape System
W	CBOE
Z	BATS
J	DirectEdge A
K	DirectEdge X
X	NASDAQ OMX PSX
Y	BATS Y-Exchange Inc.

Quotations') is an electronic screen-based market. Like the NYSE, the Nasdaq consists of a cluster of trade venues. For example, the Boston stock exchange was formerly a regional stock exchange but was acquired by Nasdaq in 2007. It now trades as Nasdaq OMX BX. Likewise, the Philadelphia stock exchange was also acquired by Nasdaq in 2007, and trades as Nasdaq OMX PHLX. One interesting recent development is the creation of a new Nasdaq market, the Nasdaq OMX 'PSX' platform (launched in October 2010). In most electronic markets, priority for limit orders in the book is decided by price and then time of placement. In other words, priority in the limit order book is given to orders with the best price and then on a time basis (in a price tie, the oldest orders are executed first). In contrast, Nasdaq OMX 'PSX' provides priority to limit orders on a price and *volume* (displayed) basis (i.e. large limit orders are executed first). The object of this is to attract large block trades to the exchange, away from dark pool trading venues. Since its launch, the average trade size on PSX platform is 30% higher than on the main Nasdaq market, and the new venue already accounts for about 1% of US equity trading (FinancialTimes, 2011b)

The National Association of Securities Dealers (now the Financial Industry Regulatory Authority - FINRA) (ADF) is not technically an exchange as it does not offer trade execution facilities. Instead it is an ‘alternative display facility’ which collects and disseminates quotes and trade information, for example from some electronic communication networks (ECNs).[4]

The national stock exchange is an electronic stock exchange based in Chicago. Prior to 2003 it was known as the Cincinnati stock exchange.

Table 4.4: US Daily Market Share Equities for 2 February 2011 (FinancialTimes, 2011a).

Market	**% Daily Turnover**
OTC	27.6
Nasdaq	18.1
NYSE Arca	15.4
NYSE	13.8
BATS	12.3
Direct Edge	9.5
Boston	1.3
Other	2.0

Note in this table, OTC includes trades not undertaken on regulated exchanges and therefore includes dark pool venues and electronic communication networks.

BATS is an electronic stock market, founded in 2005, trading approximately 10% of US equities. In 2010, BATS launched a second US equities exchange (BATS Y-Exchange). Direct Edge is also an electronic stock market and was

[4]ECNs are regulated by the SEC but are not recognised as stock exchanges.

granted SEC approval to become a stock exchange in 2010. Formerly it had been an ECN.The International Securities Exchange (ISE) took a shareholding in Direct Edge in 2008. The ISE formerly traded equities and options. The ISE stock exchange (although not the ISE options exchange) was decommissioned in 2010. A guide to the relative sizes of each exchange is provided in Table 4.4.

4.1.2 LSE ROB Database

Compared to the NYSE-Euronext TAQ database, which contains the best quotes and the corresponding depths, the LSE Rebuild Order Book data has richer information. The LSE is the fourth largest exchange in the world and the largest in Europe in terms of market capitalisation. It operates an electronic order book platform, Stock Exchange Electronic Trading System (SETS). This trading system has three sessions on each trading day:

- opening auction session, from 07:50 to 08:00, where buy and sell orders are matched at 08:00,
- continuous trading session, from 08:00 to 16:30, where buy and sell orders are continuously matched at price/time priority, and
- closing auction session, from 16:30 to 16:35, where buy and sell orders are matched at 16:35.

The ROB data records order activities, which pooled together comprise the quote prices and depths. It can be used to reconstruct the limit order book up to any quote level. The data consists of every order and trade that occurred at the LSE. It contains information on order submissions, order cancelations, order modifications and executions. Every trade is linked to the corresponding orders through a trade code. Therefore, if a large order is executed against many small orders resulting in many small trades, one can trace back to each small order using corresponding trade code. With this information one is able to reconstruct the limit order book for each millisecond of the trading day. The LSE provides three separate files to record trade and order information for each trading day. These files are:

- 'order details' file, which contains details of every new persistent order[5] entering the electronic order book,
- 'order history' file, which contains information on changes to each persistent order, including order partial/full matching, order deletion, order expiration and order modification, and
- 'trade report' file, which contains details of every trade execution.

4.1.3 Sample Data

A sample of the Order Detail file is shown in Table 4.5. As shown in the table, each order has an order code which is unique to the order. The stock code in the second column of the table corresponding to a stock listed in the LSE. In the third column, "B" stands for a buy order and "S" represents a sell order. The next four columns show the price, visible size, submission date and time for each submitted order. The last column is the message sequence number which is unique for each event (e.g. limit order submission, order cancellation, order matching, and order modification) occurring in the market. The message sequence number is usefull to distinguish the time priorities of two events which occur at the same time (accurate to one millisecond). It also appears in the Order History file and the Trade Report file.

[5]Persistent orders are the orders which are stored on the electronic order book after they are submitted to the market.

Table 4.5: A Sample of Order-Detail Data

OrderCode	StockCode	BuySellInd	Price	AggregateSize	Date	Time	MessageSequenceNumber
B101ZLMIUO	3989	B	620.00	1000	30072010	09:01:47.661	568016
B1021HZOGE	8210	S	11.50	5000	20082010	10:02:13.942	1366931
B1021PXYRW	3990	S	0.05	50000	24082010	07:55:43.761	5519
B1022A4QAL	1490	S	110.00	2640	31082010	07:50:53.980	6914
B1022AL8NY	3989	S	721.00	654	31082010	08:43:28.465	445616

A sample of the Order History file is shown in Tables 4.6 & 4.7. As shown in Table 4.6, there is an order action type attached to each entry. It records the action made to the order submitted previously and still listed in the order book. There are six order action types, which are explained in Table 4.9. After actions “D” or “E”, the order will be removed from the order book. Partial match means that part of the order is executed, and the unexecuted part is still left in the order book. Full match means that the order is fully executed and thus its aggregate size is zero as shown in the second column of Table 4.7. If it is either partial matched or fully matched, the counterparty’s code is recorded as shown in the third column of Table 4.6.[6] Action “T” rarely appears in the database.

A sample of the Trade Report file is shown in Table 4.8. As shown in Table 4.6, each order matching entry has a unique trade code.[7] This trade code corresponds to an entry in the Trade Report file. As shown in Table 4.8, each entry records details of each transaction, including the trading price, size, date and time (accurate to one second).

As shown in Tables 4.5, 4.6, 4.7 and 4.8, the information contained in these three files are connected by order codes and trade codes. The order code maps the new persistent orders listed on ‘order details’ file and their order trajectories (full execution, partial execution, deletion, expiration and modification) recorded on ‘order history’ file. Similarly, the trade code is used to track the trading details contained on the ‘trade report’ file for each order matching record on ‘order history’ file.

[6] The counterparty of a buy order is a sell order, and vice versa.

[7] A limit order executed against two or more market orders will have two different trade codes.

Table 4.6: A Sample of Order-History Data (Part I)

OrderCode	OrderActionType	MatchingOrderCode	TradeSize	TradeCode
ZO00UO59SD	P	ZO00UO5CNV	441	ZF001UD7AU
ZO00UL7QO1	M	ZO00UL8QKK	1407	ZF001UCFA1
ZO00ULIX1H	M	ZO00ULIXCB	1000	ZF001UCHP0
ZO00UNKBZE	P	ZO00UNKCD4	1331	ZF001UCZVZ
B700NPHAYC	P	B700NPO2CH	3800	A7001XEFZ1

Table 4.7: A Sample of Order-History Data (Part II)

StockCode	AggregateSize	BuySellInd	MessageSequenceNumber	Date	Time
4154	759	B	2694979	04012010	08:00:20.087
4154	0	S	240963	04012010	13:50:37.040
4154	0	S	466808	04012010	16:19:48.828
4154	1089	B	2130431	04012010	15:01:17.559
9460	3820	S	463254	04012010	14:51:13.631

Table 4.8: A Sample of Trade-Report Data

MessageSequenceNumber	StockCode	TradeCode	TradePrice	TradeSize	TradeDate	TradeTime
5272	9056	A2005L4TG0	10.32	492	01092010	10:36:38
644861	2570	A90022YXOG	146.58	7291	01092010	09:02:02
913100	9454	A700223A2U	43.94	5500	01092010	09:08:15
1281027	1437	ZA0022NJXM	2550.00	2	01092010	09:39:08
99	5657	A90022YQPO	1.67	488	31082010	17:45:33

Table 4.9: Order Action Type (ROB)

Order Action Type	**Description**
D	Deletion
E	Expiry
P	Partial Match
M	Full Match
T	Transaction Limit (150 trades)
Z	Modification (Size decrease)

4.2 Data Processing

The course of data preprocessing is presented in this section.

4.2.1 Preprocessing of TAQ data

This section initially introduces the issues in data preprocessing and then presents the steps used to preprocess the TAQ data.

Issues in Data Preprocessing

Although the data contained in the TAQ database is a valuable resource for research purposes, it is supplied in a raw (unprocessed) format and therefore requires pre-processing in order to prepare it for analysis. Figure 4.3 shows the unfiltered time series for CHDX. It clearly illustrates that there are often extremely abnormal prices in the non-NASDAQ quote time series (green line). Thus it is very necessary to filter the raw time series before using it for analysis. The cleaning of high-frequency financial data have been given special attention in many previous studies, such as Lee & Ready (1991) Bessembinder (2003) Vergote (2005) and Henker & Wang (2006). Issues which can arise in raw high-frequency data include:

Quote data only

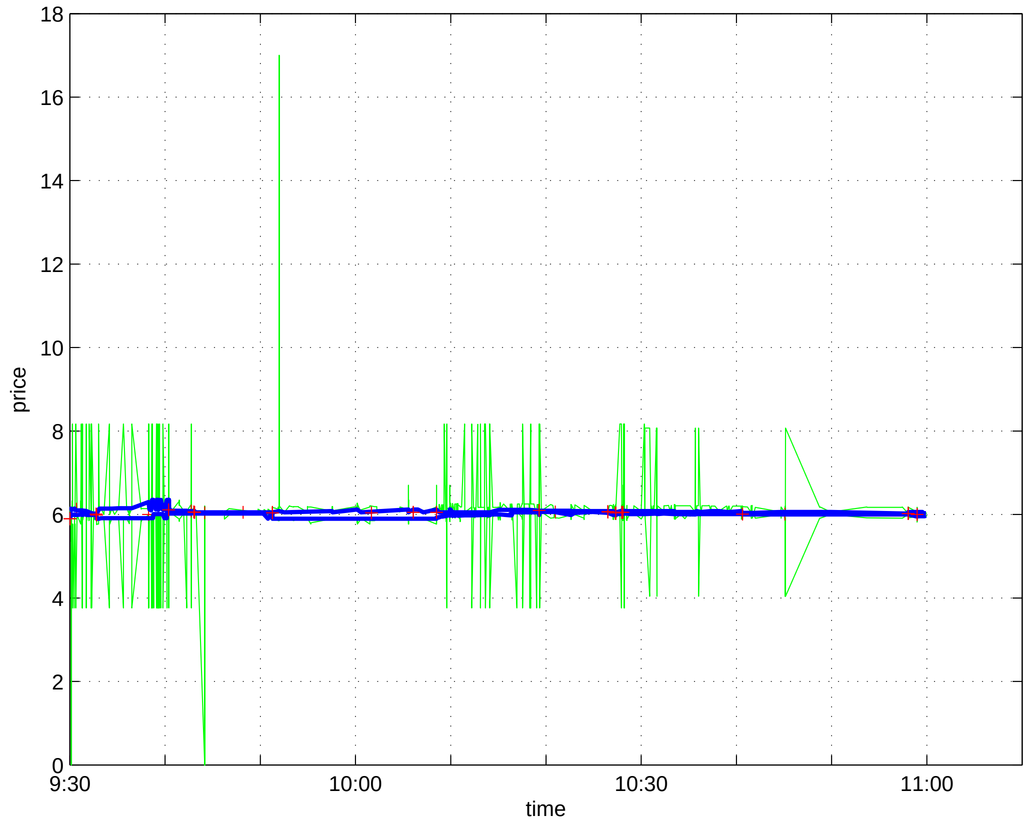


Figure 4.3: Unfiltered Time Series for CHDX at 02/02/09 from 9:30-11:00 am. The blue line depicts quotes occurring at NASDAQ; the green line depicts quotes occurring at all exchanges; the red crosses represent the transaction prices.

Q1. Quotes which are generated by non normal market activity

Trade data only

T1. Trades which are subsequently modified or cancelled

T2. Trades of which the report time is delayed

All data

A1. Ticks occurring outside normal trading period

A2. Ticks with non-positive prices, volumes or bid-ask spreads

A3. Ticks of which the prices appear inconsistent with their neighbouring prices (i.e. 'outliers')

Typically it is relatively easy to identify and correct for items Q1, T1-2 and A1-2 as 'flags' in the database and the relevant 'timestamp' can be used to help filter these items. Items falling into category A3 are trickier to handle as a series of modeller decisions are required in order to classify a data point as an 'error' which therefore should be removed in data preprocessing. In this process, the data analyst is attempting to balance out the risks of under scrubbing vs over scrubbing the data. Under scrubbing the data risks the embedding of noise in the data being analysed whereas over scrubbing the data filters out potentially vital information. The process must consider issues such as the change in price between one trade record and the next simultaneously with the level of trading activity. A trade which is, for example, \$0.10 off the previous recorded trade price would be unusual for a highly-liquid stock which trades many times a second, but may be less suspect for a thinly-traded security which only trades every 10 minutes.

Depending on the research questions of interest, and the associated required analyses (for example, trade classification, calculating effective spreads, and estimating the information context of trades (Henker & Wang, 2006)), it may also be necessary to align the trades and quotes data chronologically into a single time series. This task is not as simple as it appears at first glance as there is no field in the separate trade / quote files which make up the TAQ database which links trades with their corresponding quote update. The problem is made more complex due to the potentially differing levels of latency in data from differing sources in reaching the consolidated tape.

The degree of latency in the recording of both quotes and trades has varied over the years and hence, the most appropriate procedure to adopt in aligning these time series has also varied with obvious implications for research study design, depending on the dates from which the data is drawn. Historically, there has been greater latency in the recording of quote updates than trade updates on the consolidated tape, and consequently, lining up the time series based on time stamp may produce an unintended outcome where a trade is recorded *before* the associated change in the best quote for that exchange (obviously, any trade will impact on the current best quote for an exchange as it will usually alter some element of best bid/ask price or volume at those prices). For example, under the manual system operated in the NYSE in the 1980s, if a specialist assistant was faster in recording a quote revision than the floor reporter was in recording a trade, the corresponding quote update could be recorded before the trade that triggered it. As a result of this problem, Lee & Ready (1991) suggested, based on NYSE data drawn from 1988, that the prevailing quote five seconds before a trade should be considered to be the actual quote at the time of the trade. This rule was used in many academic studies up to and including the early 2000s.

Of course, with the increasing use of technology there have been many structural changes in the way that trade and quote information has been captured over the past 20 years. With the introduction of 'auto-quoting' for all stocks in 2003 on the NYSE, the best bid/offer was automatically updated whenever a limit order was posted to the Display Book at a better price than the previous best bid or offer, and trades taking place at the best bid/offer automatically led to an updating of the best bid/offer (Vergote, 2005). A study after the introduction of the autoquote system, using data drawn from October-December 2003, indicated that a trade - quote change lag still existed, although it had been reduced to about 2 seconds (Vergote, 2005), with another study (Henker & Wang, 2006) (which used earlier data) suggesting that the appropriate lag was one second.[8]

The increasing use of 'unlit' trading venues such as dark pools has resulted in an growing portion of trades taking place 'off exchange'. While these trades are recorded on the CTS, there will not be a corresponding 'quote change' recorded on the CQS. Another issue that should be noted is that the lag from a trade execution to the recording of that trade on the CTS is permitted to range up to 90 seconds and hence, it cannot be assumed that actual trade time and the reporting time of a trade are identical (Hasbrouck et al in a 1993 study (Hasbrouck et al., 1993) reported a median delay of 14 seconds for NYSE trades between the SuperDot execution report time and the CTS print time).

Actual Data Preprocessing

Only the quotes and trades ticks which happen at their main listed exchange are included. So the ticks happened outside NYSE for DAR, F, MDS and XOM, and the ticks happened outside NASDAQ for CHDX and GOOG are excluded. For NYSE traded stocks, it is easy to obtain their records by choosing 'N-NYSE' when using the TAQ3 extraction tool as shown in Figure 4.2. On obtaining NASDAQ traded ticks, this is a little more complicated. As shown in Table 4.3, there are two ex-

[8] While increased use of electronic trading would be expected to enhance the quality and timeliness of trading data capture, this is not always so. A recent case in point was the introduction of a new matching engine on the London Stock Exchange (14 February 2011) where due to computer glitches, price data vendors were displaying incorrect prices, blank prices and incorrect trading volumes (King, 2011).

change codes corresponding to NASDAQ, which are 'T' and 'Q'. But as shown in Figure 4.2, only 'T-NASDAQ' can be chosen when using the TAQ3 extraction tool, where 'Q-NASDAQ' is not listed. In order to obtain all ticks which happen at NASDAQ, analyst needs to select all exchanges when using the TAQ3 extraction tool. By doing this, all ticks from NASDAQ can be included in the extracted file labeled with 'Q' or 'T' in the 'Exchange' column.

In preprocessing the raw data this study follows the lead of Brownlees & Gallo (2006). For both quote and trade data, initially all ticks recorded outside the normal trading day are removed. Following Brownlees & Gallo (2006), an expanded trading period spanning between 9:30 AM and 4:05 PM is adopted, in order to insure that closing prices possibly recorded with a delay are accounted for. For trade data, all incorrect and delayed trades are initially discarded, which are indicated by the CORR field differing from 0 and the COND field equaling 'Z' respectively in the trade file. Typically, this removes only a small amount of data from the initial dataset (usually less than $< 0.5\%$). On quote data, all records generated by non normal activities are deleted, as indicated by the MODE field values 1,2,3,6 or 18 in the quote file. Then all quote and trade ticks with non-positive prices or volumes and all quotes with non-positive bid-ask spreads are deleted.

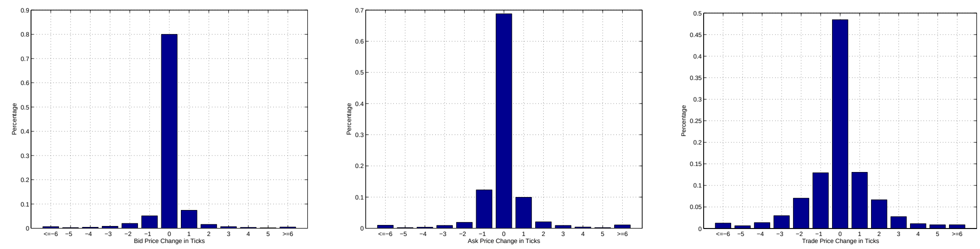
Removing Outliers

To remove outliers, observations are omitted when the absolute difference between the current price and the average neighbour price is outside three standard deviations plus a parameter that controls for the minimum price variation. To illustrate, let p_i be a value of the price time series $P_{i=1}^{N}$. The proposed method in Brownlees & Gallo (2006) to remove outliers is:

$$|p_i - \overline{p}_i(k,\delta)| < 3s_i(k,\delta) + \lambda \tag{1}$$

Here $\overline{p}_i(k)$ and $s_i(k)$ denote respectively the δ-trimmed sample mean and sample standard deviation of a neighborhood of k observations around i and λ is a granularity parameter[9]. The k neighbor observations are chosen based on the following rules: the neighborhood of an observation in the middle of the day are the first preceding $k/2$ observations and the following $k/2$ ones; the neighborhood of one of the first $k/2$ observations for the trading day are the first k observations of the same trading day; the neighborhood of one of the last $k/2$ observations for the trading day are the last k observations of the same trading day. For any given observation price p_i, if the condition above is true, this observation is kept, otherwise discarded.

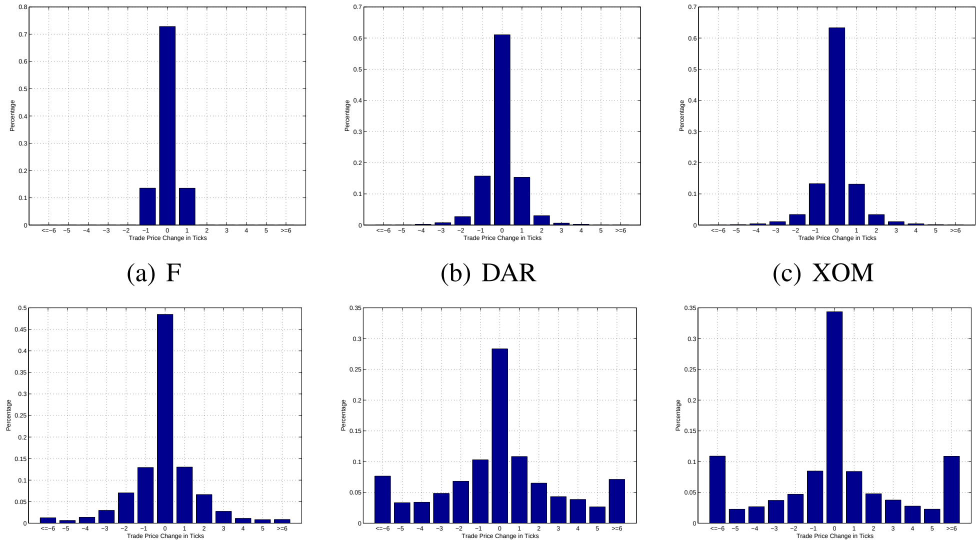
A very heuristic procedure is introduced in Brownlees & Gallo (2006) for choosing these three parameters. The parameter k is chosen based on the level of trading intensity, the more active the trading, the larger the k. The parameter δ is chosen on the basis of the frequency of outliers, the higher the frequency, the higher the δ. The choice of the granularity parameter λ is guided by the frequencies of the price changes which is always multiples of the minimum price variance for that stock.



(a) (b) (c)

Figure 4.4: Frequencies of Price Changes for CHDX. (a) frequency of bid price change (b) frequency of ask price change (c) frequency of transaction price change.

[9]The parameter λ is used to avoid zero variances from the time series of k equal prices around the observation price p_i.



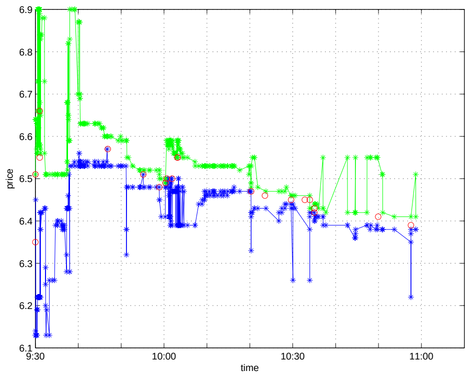
(a) F (b) DAR (c) XOM

(d) CHDX (e) MDS (f) GOOG

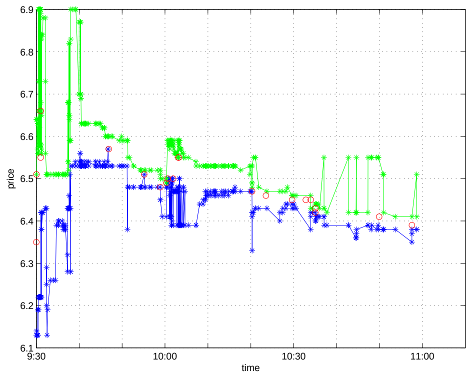
Figure 4.5: Frequencies of Transaction Price Changes for Six Stocks

Figure 4.4 illustrates the frequencies of bid, ask and transaction price changes for CHDX, which guides the choice of parameter λ. Figure 4.5 depicts the frequencies of transaction price changes for the six stock. As shown in Figure 4.5(f), there is a significant fraction of price changes equal or larger than six minimum ticks.[10] In contrast, there is almost no price change larger than two ticks for F as shown in Figure 4.5(a). The distributions of price changes for the other four stocks are between these two extremes. It is possible that more outliers are identified in the price time series with higher dispersion level of price change. The parameter δ is fixed at 10%, the same as Brownlees & Gallo (2006). Several sets of parameters (k, λ) are tested. The quality of the cleaning is judged by a visual inspection of the clean tick-by-tick price series graph. Figure 4.6 depicts the time series on 03 Feb 2009 from 9:30 to 11:00 for CHDX. The plot in Figure 4.6(a) shows the time series of raw data. The time series in Figures 4.6(b), 4.6(c) & 4.6(d) demonstrate the clean data with three different sets of parameters (k, λ). 4.6(d) provide

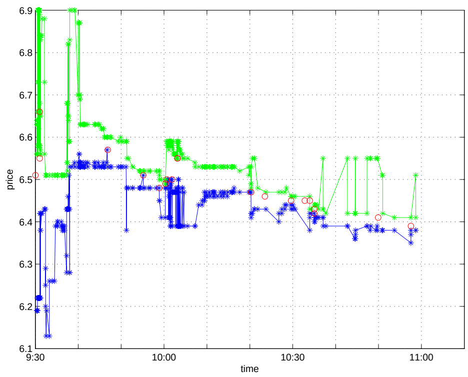
[10]One minimum tick equals 0.01.



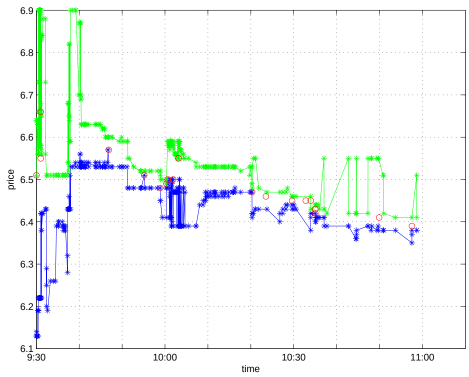
(a) raw data



(b) clean data with $k_q = 40$ $\lambda_q = 0.06$ for quote data, and $k_t = 10$ $\lambda_t = 0.08$ for trade data



(c) clean data with $k_q = 80$ $\lambda_q = 0.02$ for quote data, and $k_t = 10$ $\lambda_t = 0.04$ for trade data



(d) clean data with $k_q = 40$ $\lambda_q = 0.02$ for quote data, and $k_t = 10$ $\lambda_t = 0.04$ for trade data

Figure 4.6: Tick-by-tick Price Series on 03 Feb 2009 (9:30-11:00) for CHDX. The blue line (below) represents the bid price time series; the green line (above) represents the ask price time series; the red circles represent the transaction prices.

the most satisfactory result. The sets of most satisfactory parameters (k, λ) are reported in Table 4.10.

Table 4.10: Parameters for Removing Outliers

	Quotes Data			Trades Data		
	δ	k	λ	δ	k	λ
MDS	0.1	30	0.02	0.1	10	0.06
CHDX	0.1	40	0.02	0.1	10	0.04
DAR	0.1	50	0.02	0.1	20	0.03
F	0.1	60	0.01	0.1	40	0.02
GOOG	0.1	120	0.01	0.1	60	0.02
XOM	0.1	120	0.01	0.1	60	0.02

There are three steps to clean the raw quote data: the first step is to remove all quotes which are generated by non normal market activities; the second step is to remove all quotes with non-positive prices, volumes or bid-ask spreads; the third step is to remove the outliers. Similarly, there are three steps to clean the raw trade data: the first step is to remove all trade ticks which are modified, cancelled or delayed; the second step is to remove all trades with non-positive prices or volumes; the third step is to remove the outliers.

Tables 4.11 & 4.12 summarize the results of cleaning raw quote and trade data respectively. The values in bold represent of the percentage of the final filtered data in the raw data. Most of these percentages are larger than 99%, indicating that the filtering process generally does not alter the raw data. From the tables, one can also find that non-positive values did not appear in either raw quote data or trade data, in contrast that some zero values were found in Barndorff-Nielsen et al. (2009). This indicates that the consolidated tape system has been improved and zero values are avoided.[11] Among the three steps, the last step removes a

[11]In Barndorff-Nielsen et al. (2009), the period of analysed data is from January 3 to June 29, 2007. The time of the data analysed here is February 2009.

Table 4.11: Cleaning Quote Data.

	MDS		**CHDX**		**DAR**		**F**		**GOOG**		**XOM**	
	No.	Perc.	No.	Perc.	No.	Perc.	No.	Perc.	No.	Perc.	No.	Perc.
Raw Data	34891	1.0000	47365	1.0000	160759	1.0000	508852	1.0000	1786150	1.0000	1951197	1.0000
Non-normal Quotes	25	0.0007	0	0.0000	25	0.0002	45	0.0001	0	0.0000	42	0.0000
Non-positive Values	0	0.0000	0	0.0000	0	0.0000	0	0.0000	0	0.0000	0	0.0000
Outliers	711	0.0204	452	0.0095	59	0.0004	99	0.0002	42254	0.0237	13832	0.0071
Final Sample	34155	**0.9789**	46913	**0.9905**	160675	**0.9995**	508708	**0.9997**	1743896	**0.9763**	1937323	**0.9929**

Note: No. means the number of observations.
Perc. means the percentage of the number of observations in the raw data.

Table 4.12: Cleaning Trade Data.

	MDS		**CHDX**		**DAR**		**F**		**GOOG**		**XOM**	
	No.	Perc.	No.	Perc.	No.	Perc.	No.	Perc.	No.	Perc.	No.	Perc.
Raw Data	1339	1.0000	3128	1.0000	14588	1.0000	54874	1.0000	443140	1.0000	511462	1.0000
Incorrect Trades	0	0.0000	0	0.0000	0	0.0000	0	0.0000	0	0.0000	0	0.0000
Non-positive Values	0	0.0000	0	0.0000	0	0.0000	0	0.0000	0	0.0000	0	0.0000
Outliers	5	0.0037	8	0.0026	16	0.0011	1	0.0000	4063	0.0092	796	0.0016
Final Sample	1334	**0.9963**	3120	**0.9974**	14572	**0.9989**	54873	**1.0000**	439077	**0.9908**	510666	**0.9984**

Note: No. means the number of observations.
Perc. means the percentage of the number of observations in the raw data.

larger fraction of the observations than the other two steps. Among the six stocks, F removes a smaller fraction of the observations while GOOG removes a larger fraction of the observations. It can be found that there is a positive relationship between the number of removed observations and the dispersion level of price changes. The more dispersive the price change, the more the observations identified as outliers. Figures 4.7 & 4.8 illustrate the raw data and the filtered data for CHDX on 10 Feb. 2009 respectively.

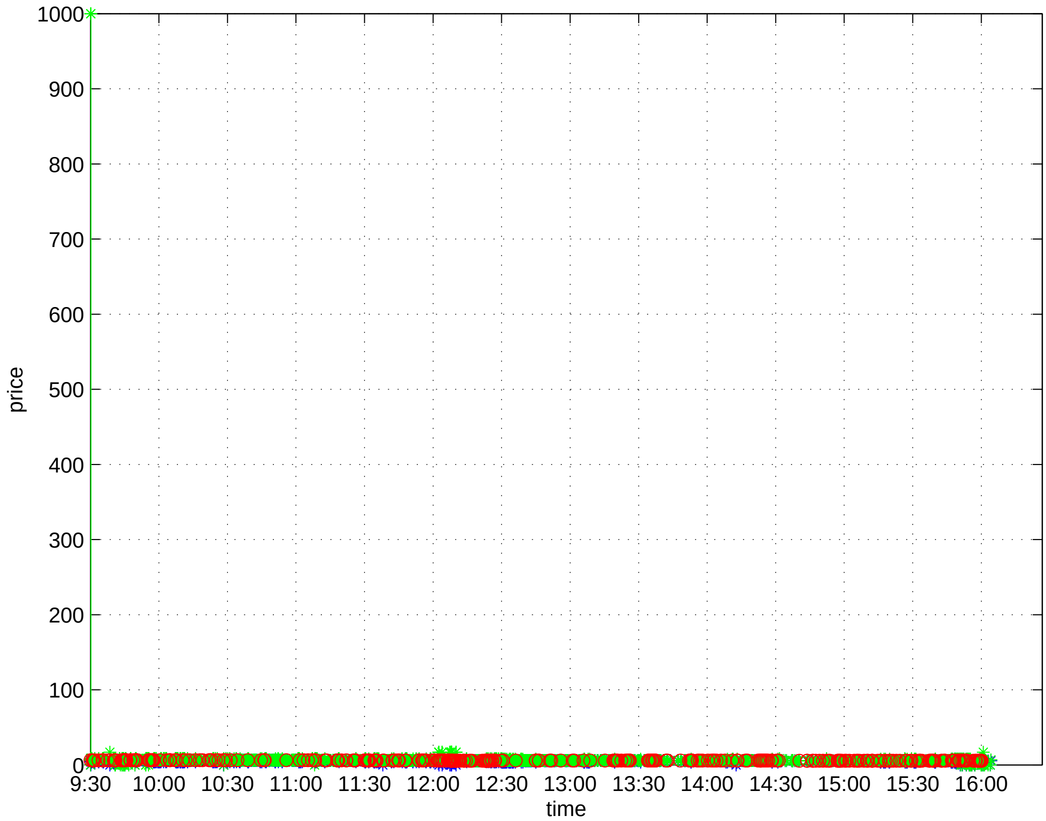


Figure 4.7: Raw Data for CHDX on Feb 10, 2009

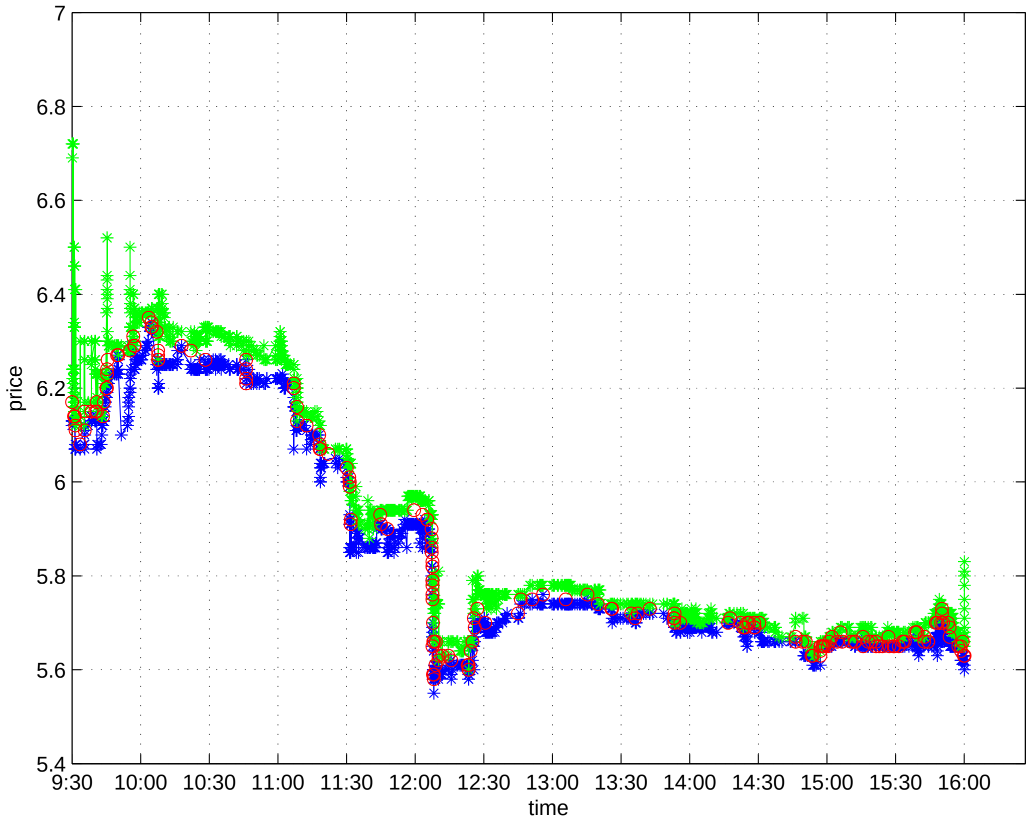


Figure 4.8: Clean Data for CHDX on Feb 10, 2009

Matching Trade and Quote

To match trade ticks with quote ticks, this study follows the one-second rule in Henker & Wang (2006). Any quote less than one second prior to the trade is ignored and the first one at least one second prior to the trade is retained as the prevailing quote. If the transaction occurs above the prevailing mid-quote, it is regarded as a buyer-initiated trade; otherwise if the transaction occurs below the prevailing mid-quote, it is a seller-initiated trade. If the transaction occurs exactly at the mid-quote, it is signed using the previous transaction price, which is called 'tick test' in Lee & Ready (1991). If the transaction price is higher (lower) than the price of the previous trade, it is classified as a buyer(seller)-initiated trade. If the transaction price is the same as the price of the previous trade, it is signed using the previous transaction price change. If the previous price change is upward (downward), the trade is a buy (sell). Sometimes, the trade does not fall into any cases mentioned above, for example the trade is the first trade of the trading day. An alternative test called 'reverse tick test' in Lee & Ready (1991) can be applied which compares the current trade price with the trade prices immediately after the current trade. If the following trade price is higher (lower) than the current trade price, the trade is classified as a sell (buy).

4.2.2 Preprocessing of ROB data

The data contained in the ROB database is supplied in a raw (unprocessed) format and therefore requires preprocessing in order to prepare it for analysis. Unlike the raw data in the TAQ database, the ROB data rarely contains non-positive values, but the ROB data is out of time sequence and more complicated than the TAQ data. The information needed to analyse the order book is not direct to obtain and the some preprocessing steps are needed. For example, market order information can not be directly found in the data files provided and needs to be inferred using the available data information.

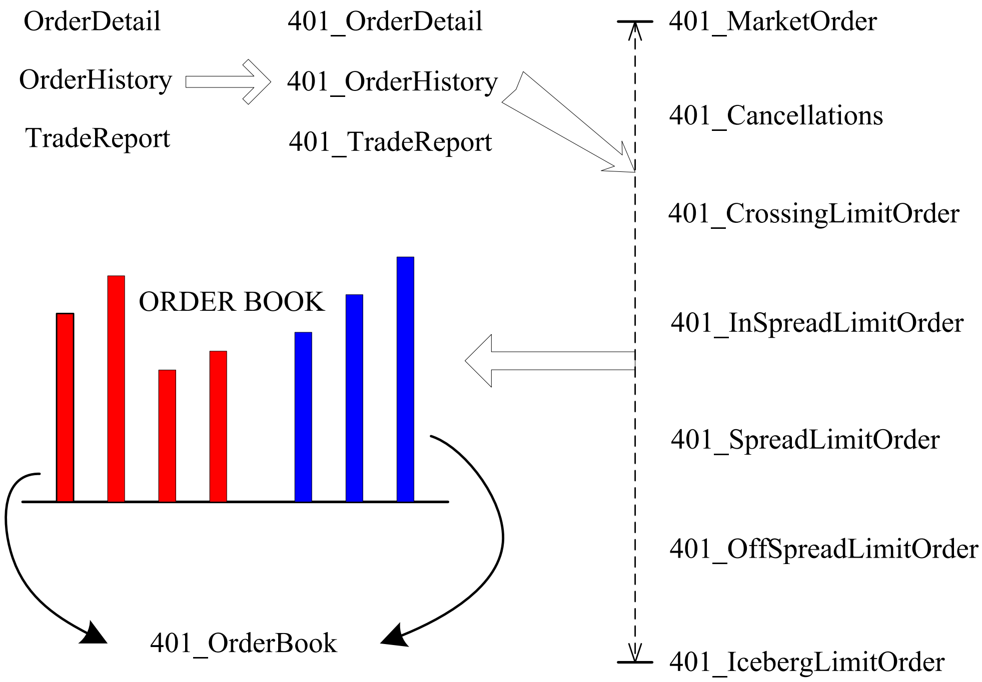


Figure 4.9: ROB Data Preprocessing Chart

Figure 4.9 illustrates how to preprocess the ROB data in order to extract the order book information for the 1st of Aril in 2010 which is denoted by "401". There are three steps explained as follows.

1. The first step is to extract the data for each trading day from the three original files (Order-Detail, Order-History and Trade-Report), because each original file contains information for numerous (about 11) continuous days. As shown in Figure 4.9, the information for 1 Oct. 2010 is then written to 401_Order-Detail, 401_Order-History and 401_Trade-Report.

2. The second step is to infer the information on undisclosed orders, and extracting the information on limit orders with various levels of aggressiveness

(which will be used in the study in Chapter 6). Detail of this step is explained in the next section.

3. The third step is to reconstruct the order book using the information extracted in the last step, and record the information of the order book up to any quote level. As shown in Figure 4.9, the order book information is recorded in 401_OrderBook. Detail of this step is described below.

Inferring Undisclosed Orders

In order to get the full information needed for the experiments in Chapters 5 & 6, non-persistent market orders and three missing events from the given data files have to be inferred. Market orders can be inferred from the 'order history' file as it records information on the matching side of each transaction. In this study, the reconstructing process is implemented in JAVA.

The first missing event in ROB data is the execution of iceberg limit order. LSE allows traders to place iceberg limit orders, part of which are hidden in the order book and not recorded in the ROB data. When the visible part of the iceberg limit order is matched by a market order with larger size, the hidden part will be executed against the rest of the market order. The traded hidden part of the iceberg limit order can be inferred from the records of the limit order whose transaction size is larger than its original size.

The second set of missing events are crossing limit orders which cause immediate trade after submission. The 'order details' file only records the unexecuted part of the crossing limit orders. The traded part can be found from the 'order history' file in that each crossing limit order is matched by one or more persistent limit orders previously submitted to the market.

The third missing event is the limit order which was submitted to the market a few days ago but is executed today. The information on details of these orders needs to be recovered from older data files.

After this step, seven data files are obtained as shown in Figure 4.9. Four types of limit orders with various aggressiveness are distinguished for the use in Chapter 6, which are crossing limit order, in-spread limit order, spread limit order, and off-spread limit order. Details of these order types are explained in Chapter 6.

Reconstructing The Limit Order Book

In order to get information on the limit order book up to any quote level, the order book on each trading day needs to be reconstructed using all the available data, which is already acquired in the second step.

At first, the opening time of the normal trading session needs to be inferred. This time is a random point between 8:00:00 a.m. and 8:00:30 a.m. At this point, all the orders submitted during the opening sessions are executed at the price which can maxmise the trading volume. Therefore, the opening time is the time point at which the number of trades is largest in the period from 8:00:00 a.m. to 8:00:30 a.m. Then all the orders occurring during the same trading day are sorted in time sequence. For those events occurring at the same time, they are distinguished using their sequence numbers. All the rules as in the real market are applied for reconstructing the limit order book, like the price & time priorities. In this study, the reconstructing process is implemented in Matlab.

4.3 Summary

The objective of this chapter was to provide an introduction to two popular high-frequency databases, and using this data, explore some of the issues faced by researchers seeking to use ultra-high frequency financial databases. With the pre-processed data, three studies on price impact are conducted in the following chapters.

Chapter 5

An Analysis of Intraday Behaviours of Price Impact and Market Liquidity

Much of the content of this chapter has been in:

- Cui, W. and Brabazon, A. (2012). An Analysis of Liquidity in London Stock Exchange. *Proceedings of the 25th Irish Accounting and Finance Association Conference*, pp. 1-45, Galway, Ireland. (Cui & Brabazon, 2012b)
- Cui, W. and Brabazon, A. (2011). Ultra High-frequency Financial Data: Analysing NYSE Euronext TAQ Data. *Proceedings of the 24th Irish Accounting and Finance Association Conference*, pp. 1-38, Cork, Ireland. (Cui & Brabazon, 2011)

Intraday phenomena in financial markets have been extensively studied in the literature on market microstructure as shown in Section 2.4 of Chapter 2. However, many of these studies are dated. For example, the most recent studies of intraday phenomena in the US markets and the UK markets are analysed with the 1999 data (Henker & Wang, 2006) and 2001 data (Cai et al., 2004) respectively. Up-to-date evidence is useful for constructing market microstructure theories and is crucial for developing intraday trading strategies. It can also help regulators to design an appropriate measure of liquidity for an efficient and transparent trading system.

There are two objectives of this chapter. The first objective is to examine the intraday behaviour of market liquidity. In this chapter, intraday behaviours of price volatility, bid-ask spread, trading volume, trading frequency, trade size, market depth (cover the best level of the order book) and near market depth (covers the ten best levels of the order book) are investigated.

The second objective of this chapter is to investigate the intraday behaviour of price impact. Several studies have examined the intraday behaviour of price impact. Dufour & Engle (2000); Engle & Patton (2004); Hasbrouck (1999) find that price impact measured as changes of mid-quote, bid and ask prices caused by trades does not have an intraday pattern on the NYSE. In contrast, Chan (2000) shows that price impact displays a U-shaped pattern over the trading day on the HKSE. However, no previous paper examines the intraday behaviour of price impact on the LSE or the NASDAQ. Therefore, it is interesting to know whether the price impact exhibits an intraday pattern in these markets.

In this chapter, the analysed data is drawn from two databases, the NYSE-Euronext TAQ (trade and quote) database and the LSE ROB (rebuild order book) database, and covers three markets, the NYSE, the NASDAQ and the SETS of LSE. This chapter contributes to the literature in the following ways. It documents up-to-date evidences on the intraday behaviours of a number of interesting variables including the price volatility, trading size, trading volume, bid-ask spread, market depths and price impact, in the UK market as well as the US markets, which complements the literature on intraday phenomena.

The chapter proceeds as follows. Section 5.1 introduces various measures of market liquidity which are examined in this chapter. Section 5.2 presents empirical results and discussions. A summary of this chapter is given in Section 5.3.

5.1 Measures of Market Liquidity

There is no single all-encompassing measure of market liquidity as liquidity is a multi-dimensional variable. In this chapter, a number of interesting metrics are used to investigate the intraday pattern of market liquidity, including bid-ask spread, market depth, and price impact ratio. Each trading day is partitioned into successive 15-minutes intervals, and the measure these metrics for each stock during each of the 15-minute intervals on each trading day, which are then averaged over the whole sample period (19 trading days for US stocks and 127 trading days for UK stocks).[1] The methods used to estimate these metrics are described as follows.

5.1.1 Price-based Measures

Following Abhyankar et al. (1997) and Cai et al. (2004), absolute mid-quote returns[2] are used as the return volatility measure. The mid-quote return is calculated as

$$R_t = [log(P_t) - log(P_{t-1})]$$

[1] This thesis chooses a reasonable time length of the data (19 trading days for US stocks and 127 trading days for UK stocks), which is enough to generate intraday phenomena. During these time periods, there were no disturbances in the markets.

[2] The use of mid-quote prices rather than transaction prices should reduce spurious volatility due to the bid-ask bounce.

where P is the mid-quote price at the end of each minute. At first the price returns for each 1-minute interval is calculated and then the average absolute return is used as the return volatility at each 15-minute interval.

The bid-ask spread is the difference between the prevailing best ask price (ASK) and the prevailing best bid (BID). In this chapter the proportional bid-ask spread ($ProSpr$)[3] is used to measure the spread. It is defined as the ratio of the prevailing bid-ask spread and the prevailing mid-quote price ($0.5 * (ASK + BID)$):

$$ProSpr = \frac{ASK - BID}{0.5 * (ASK + BID)}.$$

The 15-minute proportional quoted spread is calculated as the average quoted spread during the period.

5.1.2 Volume-based Measures

Previous work typically only measures the depth at the best available prices (best bid & ask price). However, the quotes and volume beyond the spread also have effects on the market liquidity. Cao et al. (2008) found that the liquidity beyond the spread is informative about future short-term returns. Hopman (2007) has shown that order imbalance inside the order book is an important determinant of simultaneous price changes. Farmer et al. (2004) found that large price impacts caused by market order executions are not always associated with large trading sizes, but are due to gaps at different levels of the order book. In this chapter the near market depth is also investigated which accounts for the 10 highest levels of the order book. The market depth (Dep) is measured as the average value of all shares at the best bid price and all shares at the best ask price:

[3] For the purpose of convenient comparison of different stocks.

$$DepShare = \frac{Shares_{BestBid} + Shares_{BestAsk}}{2}.$$

The near market depth ($NDep$) is measured as the average value of all shares at the best 10 bid prices and all shares at the best 10 ask prices:

$$NDepShare = \frac{\sum_{i=1}^{10} Shares_{bid,i} + \sum_{i=1}^{10} Shares_{ask,i}}{2}$$

where i is the level of the order book. The depth in value is also measured. The depth in value is measured as

$$DepValue = \frac{BID * Shares_{BestBid} + ASK * Shares_{BestAsk}}{2}$$

The near depth in value is measured as

$$NDepValue = \frac{\sum_{i=1}^{10} BID_i * Shares_{bid,i} + \sum_{i=1}^{10} ASK_i * Shares_{ask,i}}{2}.$$

The depth is calculated for each time recorded in the data. It is averaged over each 15-minute interval and over the whole sample period.

The 15-minute trading size is the average trading size at this period and it is then averaged over the whole sample period. The trading size in value which is trading size timed by its corresponding trading price is also measured.

5.1.3 Price and Volume Mixed Measures

More recently, with the increased availability of ultra high-frequency data from stock exchanges and other alternative trading platforms, several papers have looked at measures of liquidity other than bid-ask spread, including Sandas (2001); Ranaldo (2004); Hall & Hautsch (2007); Large (2007).

In this study, *price impact ratio* is proposed to measure liquidity cost. Price impact is the mid-quote price change caused by a trading execution. The price impact ratio ($PiRatio$) is measured as the ratio of price impact and its corresponding trading size:

$$PiRatio = \frac{Price_{after} - Price_{before}}{TradeSize}$$

where $Price$ is the mid-quote price, $Price_{after}$ is the market price immediately after this trading execution and $Price_{before}$ is the market price just before the trading execution. Price impact ratio measures the liquidity cost of each traded share.

5.2 Empirical Results and Discussions

The aim of this experiment is to investigate whether price impact exhibits an intraday pattern and whether this intraday pattern is consistent across different stocks in US and UK markets. Six firms from the TAQ database are selected for analysis (Table 5.1). The firms selected are of varying market capitalisation and industry sector. Table 5.2 reports a number of data descriptives concerning these companies drawn from 2nd February 2009. All of the companies are listed on US markets.

For the ROB database, the six most actively traded stocks, whose market capitalisations exceed 20 billion pounds are selected for analysis. The companies are Barclays, British American Tobacco, BP, Glaxo Smith Kline, HSBC Holdings and Vodafone Group. They are from five different industries: consumer goods, basic materials, healthcare, financial and technology sector (shown in Table 5.3). Also all of them are British multinational companies headquartered in London, United Kingdom and primarily listed in the London Stock Exchange.

Table 5.1: Companies Listed in the US markets. Market capitalisation and share prices as at 01/04/2011

Name	Ticker	Primary Mkt.	Mkt. Cap.	Share Price
Exxon	XOM	NYSE	$419.55 Bn	$84.79
Google	GOOG	Nasdaq	$191.22 Bn	$589.01
Ford	F	NYSE	$57.02 Bn	$15.03
Darling Int'l	DAR	NYSE	$1.85 Bn	$15.48
Chindex Int'l	CHDX	Nasdaq	$0.25 Bn	$16.05
Midas Inc.	MDS	NYSE	$0.11 Bn	$7.76

Table 5.2: Trading Activity of US Stocks (02/02/2009).

Name	Ticker	Consol. Vol	Consol. $ Vol	Consol. Trades	Av. Trade Size
Exon	XOM	36,128,159	$2,761.699m	183,030	197.39
Google	GOOG	5,271,804	$1,789.620m	38,653	136.38
Ford	F	32,404,587	$60.672m	26,246	1,234.65
Darling Int'l	DAR	1,091,619	$4.817m	5,514	197.97
Chindex Int'l	CHDX	39,739	$244,389	264	150.53
Midas Inc.	MDS	39,200	$376,421	284	138.02

Note: Consol. Vol is short for Consolidated Volume.
Consol. $ Vol is short for Consolidated Dollar Volume.
Consol. Trades is short for Consolidated Trades.
Av. Trade Size is short for Average Trade Size.

Table 5.3: Information of Companies Listed in the LSE (28/09/2012)

Ticker	Full Name	Sector	Market Capitalisation
BARC	Barclays PLC	Financial	26.42b
BATS	British American Tobacco PLC	Consumer Goods	62.29b
BP	BP PLC	Basic Materials	83.47b
GSK	Glaxo Smith Kline PLC	Healthcare	70.19b
HSBC	HSBC Holdings PLC	Financial	93.16b
VOD	Vodafone Group PLC	Technology	87.23b

The next section provides summaries of descriptive statistics for both datasets. Then the empirical results on the intraday patterns are presented graphically in Figures 5.2-5.25, and are discussed in Sections 5.3.2-5.3.5.

5.2.1 Summary Statistics

Summary statistics for the companies are provided. Analysed variables include:

- Mid-quote price, which is the average price over the sample time period;
- Number of trades per day, which is the average number of trades over the sample time period;
- Trade size, which is the average number of shares per trade for all the trades in the sample data;
- Buyer-initiated trade size, which is the average number of shares per trade for all the buyer-initiated trades in the sample data;
- Seller-initiated trade size, which is the average number of shares per trade for all the seller-initiated trades in the sample data;
- Bid depth, which is the average number of shares available at the best bid price for the time period in the sample data;
- Ask depth, which is the average number of shares available at the best ask price for the time period in the sample data;
- Buy/sell ratio, which is the average daily ratio of daily number of buyer-initiated trades ($DailyNumber_{buys}$) to daily number of seller-initiated trades ($DailyNumber_{sells}$)

$$daily\ buy/sell\ ratio = \frac{DailyNumber_{buys}}{DailyNumber_{sells}};$$

- Probability of reversal, which is the average daily probability of sign reversal in consecutive trades, for instance a buy trade is followed by a sell trade, and is calculated as

$$daily\ probability\ of\ reversal = \frac{DailyNum_{BS} + DailyNum_{SB}}{DailyNum_{trades} - 1}$$

  where $DailyNum_{trades}$ is the daily number of trades, $DailyNum_{BS}$ is the daily number of reversal from a buy to a sell, and $DailyNum_{SB}$ is the daily number of reversal from a sell to a buy.

Summary Statistics for US Stocks

A summary of descriptive statistics of the six US stocks is reported in Table 5.4. As shown in this table, GOOG has the highest midquote price among the six stocks, because Google has never split its stock. The midquote price of stock F is the smallest which is 1.84.

On the proportional bid-ask spread, the bigger companies (GOOG, XOM) have narrower spreads than smaller companies (MDS, CHDX, DAR, F), because bigger stocks are more liquid to trade than smaller stocks.

The average number of quotes per day of the six stocks is 38,960, which implies that there is more than one quote per second.[4] Especially, there are more than four transactions occurring per second for GOOG and XOM, and more than one for F.

[4]One trading day, from 9:30 to 16:05 equals to 23700 seconds or 395 minutes.

Table 5.4: Descriptive Statistics for US stocks.

	MDS	**CHDX**	**DAR**	**F**	**GOOG**	**XOM**	mean
Midquote price	8.64	5.66	4.83	1.84	349.00	74.21	**74.03**
Proportional bid-ask spread	0.0137	0.0101	0.0051	0.0062	0.0007	0.0005	**0.0060**
Number of trades per day	81	168	767	2888	23109	26877	**8981**
Trade size (no. of shares)	177	207	245	2016	127	331	**517**
Buyer-initiated trade size	157	189	248	1843	129	335	**483**
Seller-initiated trade size	200	223	241	2237	124	329	**559**
Bid depth (no. of shares)	205	436	1025	57839	227	1584	**10219**
Ask depth (no. of shares)	202	247	918	46268	223	1254	**8185**
Buy/Sell ratio	1.1726	0.9295	1.0904	1.2761	0.9873	0.7943	**1.0417**
Probability of reversal	0.4255	0.4500	0.6135	0.9430	0.1816	0.4297	**0.5072**

Note: The numbers reported in this table are the average values over 19 trading days (Feb. 2009).

The average number of trades per day for the six stocks is 8,981, which implies that more than 22 transactions are occurring per minute. About one transaction happens per second for bigger stocks (GOOG and XOM). In contrast, about one transaction happens every five minutes for thinly traded stock MDS. The stocks with higher market capitalisation are traded more frequently than those with lower market capitalisation.

The average trade size of the six stocks is 517, about half of the trade size ten years ago.[5] There are five stocks with an average trade size less than 400. The average buyer-initiated trade size of the six stocks is 483, smaller than the average seller-initiated trade size (559).

The ask(bid) depth is larger than buyer(seller)-initiated trade size for each stock. The buy/sell ratio is the ratio of the number of buyer-initiated trades to the number of seller-initiated trades.

The average buy/sell ratio of the six stocks is 1.0417. Three of the stocks, MDS, DAR and F, have buy/sell ratios larger than one which reflect that there are more buyer-initiated trades than seller-initiated trades, and the other three (CHDX, GOOD and XOM) are smaller than one which suggest that there are more seller-initiated trades than buyer-initiated trades. It is interesting to see how the stock price moves in the sample period. Figure 5.1 plots the price movement during the sample period (Feb. 2009) for stock XOM. Since the buy/sell ratio of XOM is 0.7943, it is not surprising to see that the price falls over the sample period in Figure 5.1: the price rises from 76.69 and reaches a peak (80.34) at the end of the first week in the sample period, and then drops for the rest of the sample period (ending at 67.90).

There are two stocks with a trade reversal probability larger than 0.5, and four

[5]Henker Henker & Wang (2006) investigated the average trade size for 401 stocks and more than 200 trading days in 1999. They found that the average buyer-initiated trade size is 1166 and the average seller-initiated trade size is 1096.

stocks with a trade reversal probability less than 0.5. While F has a very high reversal probability 0.9430, GOOG has a very low reversal probability 0.1816. The average probability of trade reversal is very close to 0.5, suggesting that in general the probability that a trade is followed by another trade with the same sign is similar to the probability that it is followed by a trade with reverse sign for these six stocks.

Figure 5.1: Price of XOM during Feb. 2009 (from Yahoo Finance)

Summary Statistics for UK Stocks

A summary of descriptive statistics of the six UK stocks is reported in Table 5.5. As shown in the table, the proportional bid-ask spreads of these six stocks are almost the same, two are 0.0007, three are 0.0006 and one is 0.0005.

The rest of the variables reported in Table 5.5 reflect trading activity during the period from April 2010 to September 2010 for the six UK stocks. The third row is the values of number of trades per day. Higher values indicate that the stock is more actively traded. As seen from the table, BARC, BP and HSBC are more actively traded during the period than BATS, GSK and VOD.

The fourth row is the values of trade size in shares. While the average trade size of BATS is less than 1,000, the average trade size of VOD is more than 10,000.

Table 5.5: Descriptive Statistics for UK stocks.

	BARC	**BATS**	**BP**	**GSK**	**HSBC**	**VOD**	mean
Midquote price	313.49	2187.29	417.60	1193.99	649.67	146.44	**818.08**
Proportional bid-ask spread	0.0007	0.0006	0.0006	0.0007	0.0005	0.0006	**0.0006**
Number of trades per day	9388	3211	12161	3498	8449	5783	**7082**
Trade size (no. of shares)	3852	657	3339	1573	2434	11176	**3839**
Buyer-initiated trade size	3853	660	3382	1598	2463	11173	**3855**
Seller-initiated trade size	3856	656	3320	1569	2403	11426	**3872**
Bid depth (no. of shares)	8220	2274	8593	9378	6622	60499	**15931**
Ask depth (no. of shares)	8394	2388	8428	10096	6783	65373	**16910**
Buy/Sell ratio	1.0010	1.0036	1.0048	1.0402	1.0059	1.1405	**1.0327**
Probability of reversal	0.4825	0.4857	0.4759	0.4742	0.4762	0.4672	**0.4769**

Note: The numbers reported in this table are the average values over 126 trading days (April-September 2010).

The fifth and the sixth rows are the values for buyer-initiated trade size and sell-initiated trade size respectively. These values indicate that buyer-initiated trade size almost equals seller-initiated trade size. This suggests that it is not necessary to assume that buyer-initiated trade size differs from seller-initiated trade size in theoretical models.

The seventh and eighth rows are the values of bid depth and ask depth separately. Observing these values, the differences between bid depth and ask depth are very small for the six stocks. This suggests that the limit order book is symmetric at least at the best levels in these six stocks.

The ninth row is the values of buy/sell ratio. The buy/sell ratio nearly equals to one in all the six stocks.

The last row in the table reports the values of probability of reversal. All of the six stocks have a trade reversal probability very close to 0.5. This suggests that in general the probability that a trade is followed by another trade with the same sign is similar to the probability that it is followed by a trade with reverse sign for these six stocks. This result is consistent with the finding above for the US stocks.

5.2.2 Intraday Price Volatility

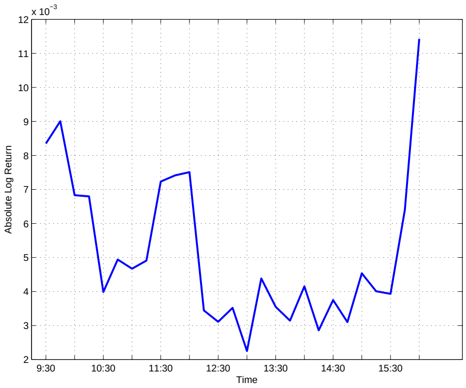
Price volatility is a direct measure of risk and an indirect measure of the level of information (French & Roll, 1986).

Intraday Price Volatility for the US stocks

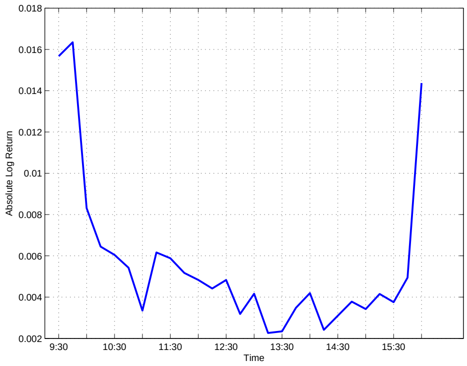
Figure 5.2 shows all the six stocks have U-shaped intraday pattern on price volatility, which is consistent with the findings in prior literature (shown in Table 2.2 in Chapter 2). The results support the predictions of the Brock & Kleidon (1992) model which predicts higher volatility at market open and close.

Intraday Price Volatility for the UK stocks

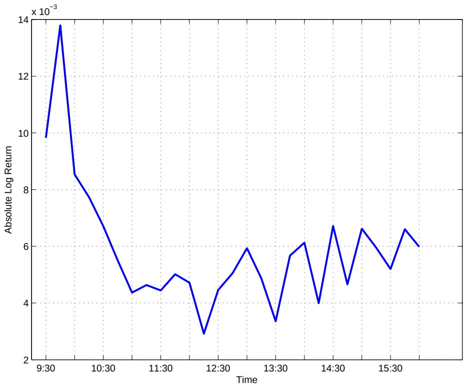
As shown in Figure 5.3, the volatility is relatively higher at market open and close, and relatively lower in the middle of the day for all the six UK stocks. This finding is in line with the U-shaped intraday pattern on price volatility found in previous literature (shown in Table 2.2 in Chapter 2).



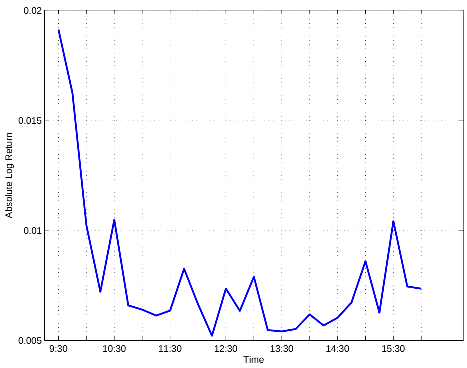
(a) MDS



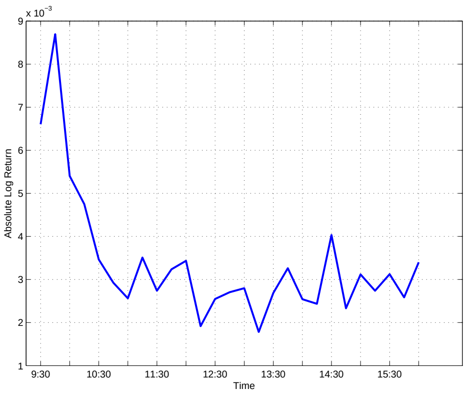
(b) CHDX



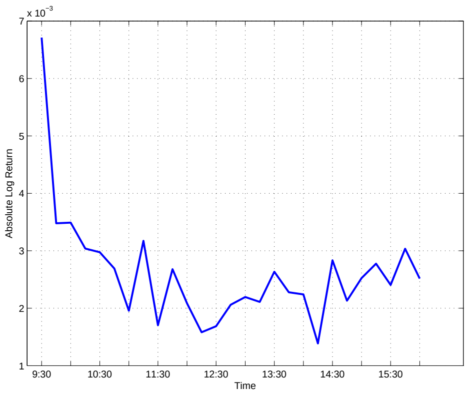
(c) DAR



(d) F

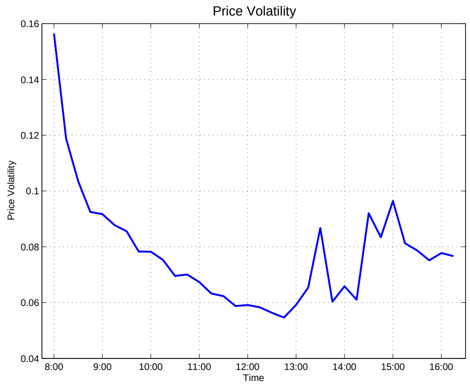


(e) GOOG

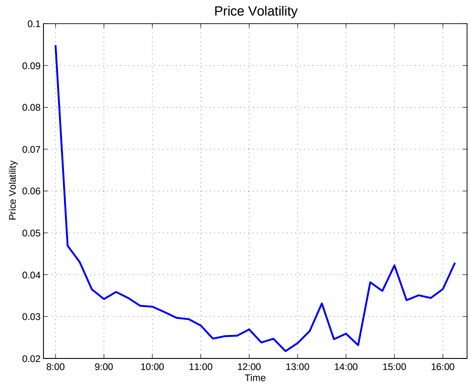


(f) XOM

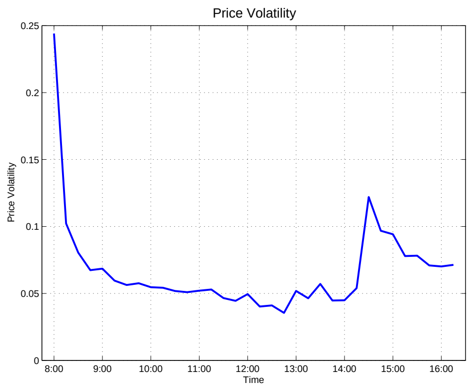
Figure 5.2: Intraday Price Volatility (US stocks)



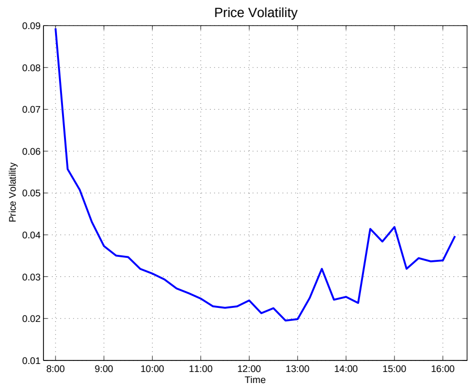
(a) BARC



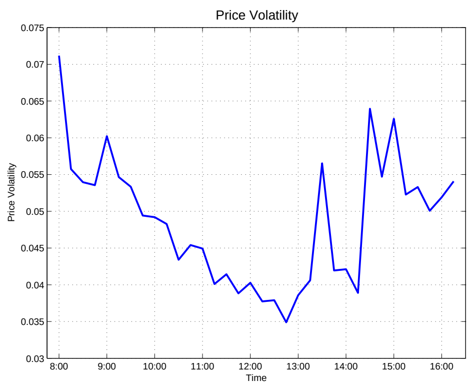
(b) BATS



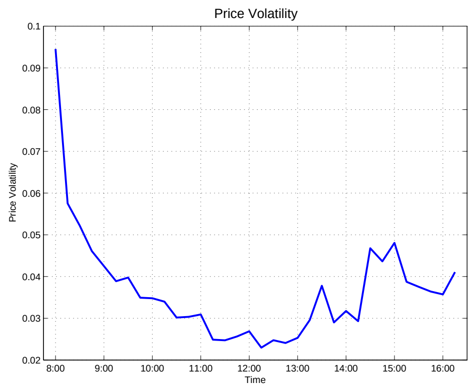
(c) BP



(d) GSK



(e) HSBC



(f) VOD

Figure 5.3: Intraday Price Volatility (UK stocks)

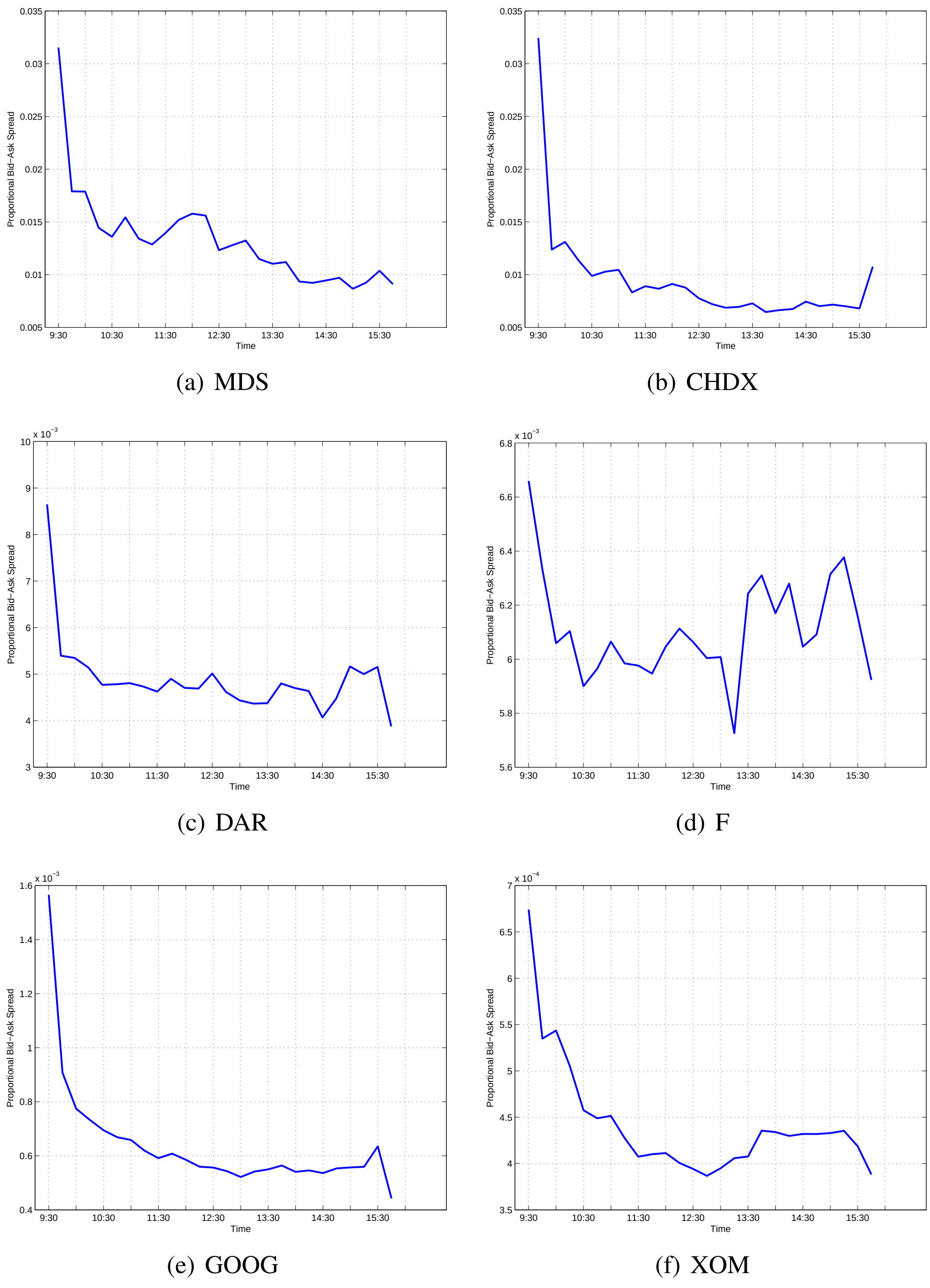
5.2.3 Intraday Bid-ask Spread

Intraday Bid-ask Spread for the US stocks

As shown in Figure 5.4, the proportional bid-ask spreads in five stocks (MDS, DAR, F, GOOG and XOM) do not exhibit the U-shaped intraday pattern as observed by many empirical studies (Chung et al., 1999; Lee et al., 1993; Madhavan et al., 1997) which examine the intraday pattern for the NYSE market. However, the finding is consistent with Henker & Wang (2006), Chan et al. (1995) and Werner & Kleidon (1996), who examine intraday patterns on the NYSE, NASDAQ and the LSE markets respectively, and all demonstrate a reverse S-shaped intraday pattern on bid-ask spread: it is relatively high at market open, is stable in the middle of the trading day, and drops at market close. As shown in Figure 5.4, only CHDX has a U-shaped intraday pattern on bid-ask spread consistent with the findings in Abhyankar et al. (1997), Ahn & Cheung (1999) Chung et al. (1999), Lee et al. (1993), Lehmann & Modest (1994), Madhavan et al. (1997) and Vo (2007).

Intraday Bid-ask Spread for the UK stocks

The intraday proportional bid-ask spreads and intraday time-weighted proportional spreads for the UK stocks are shown in Figure 5.5 and in Figure 5.6 respectively. They both demonstrate that the bid-ask spreads of these six UK stocks have reverse S-shaped intraday pattern: bid-ask spread is highest at market open, stable at the middle of the trading day, and lowest at market close. The result confirms the finding in Werner & Kleidon (1996) who examine British stocks, and is also consistent with the reversed S-shaped pattern on the NYSE market and the NASDAQ market demonstrated by Henker & Wang (2006) and Chan et al. (1995) separately.

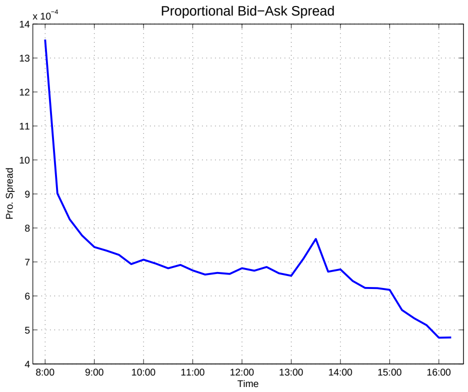


(a) MDS (b) CHDX

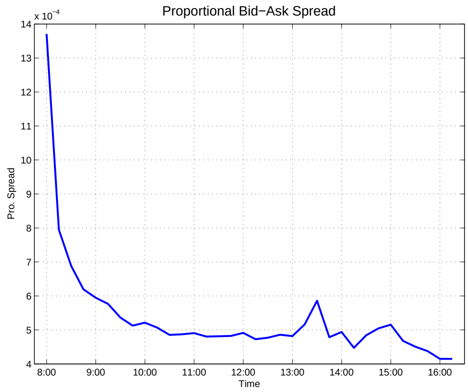
(c) DAR (d) F

(e) GOOG (f) XOM

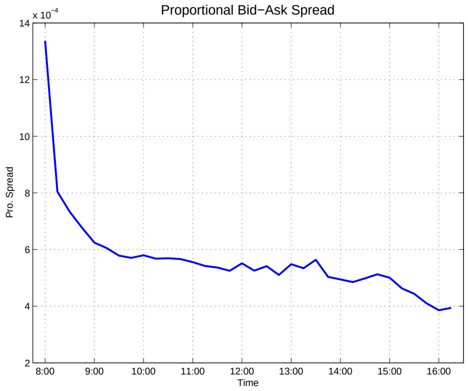
Figure 5.4: Intraday Proportional Bid-Ask Spread (US stocks)



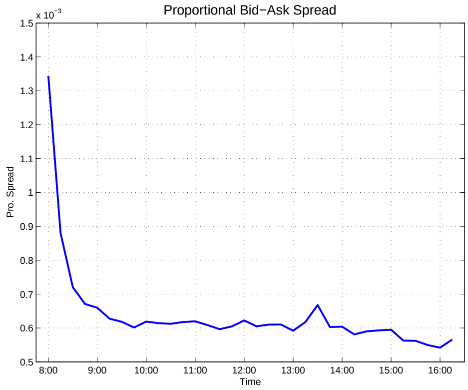
(a) BARC



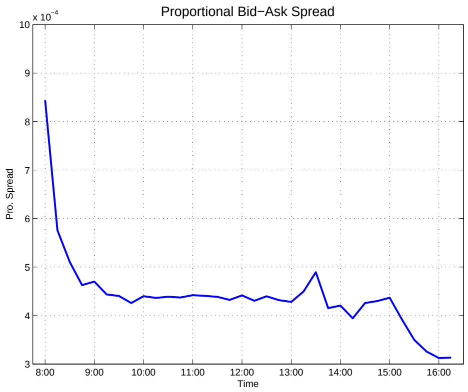
(b) BATS



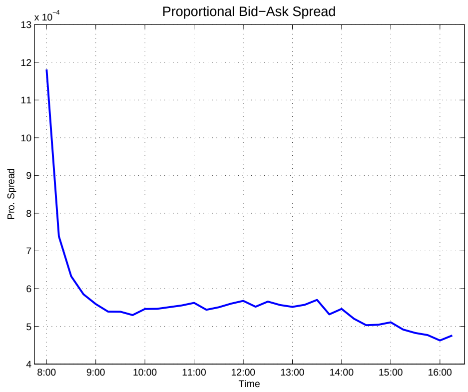
(c) BP



(d) GSK

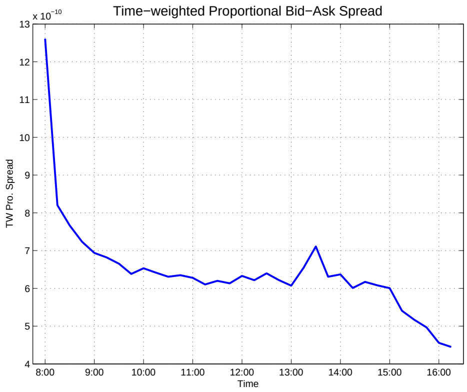


(e) HSBC

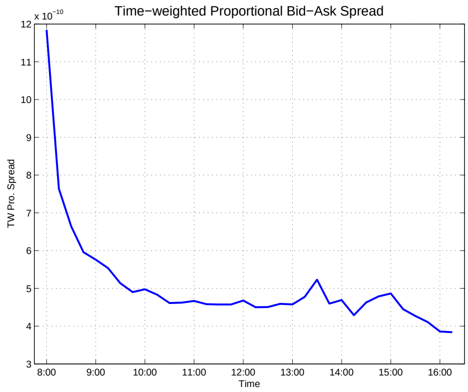


(f) VOD

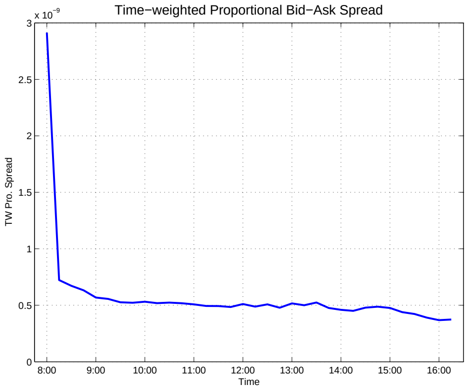
Figure 5.5: Intraday Proportional Spread (UK stocks)



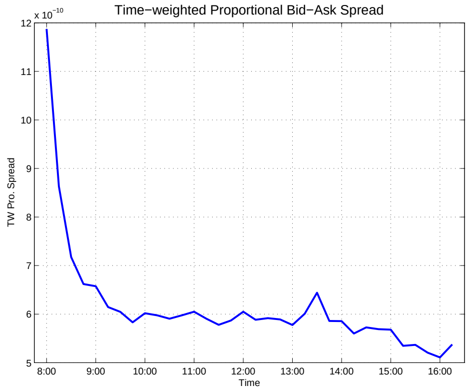
(a) BARC



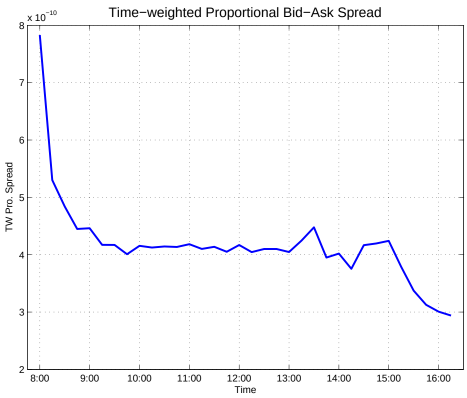
(b) BATS



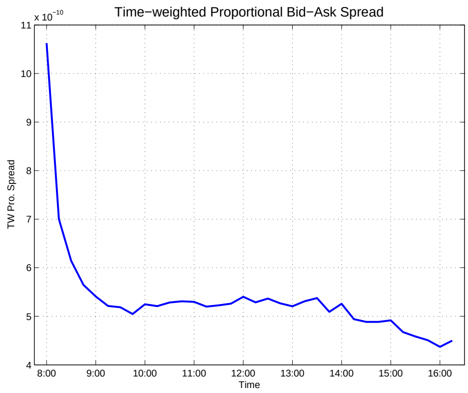
(c) BP



(d) GSK



(e) HSBC



(f) VOD

Figure 5.6: Intraday Time-weighted Proportional Spread (UK stocks)

5.2.4 Intraday Market Depth

Intraday Market Depth for the US stocks

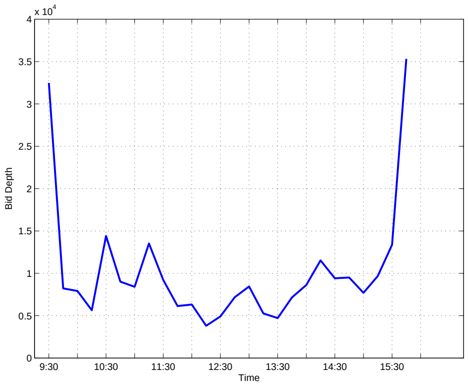
Figures 5.7 & 5.8 demonstrate the intraday behaviours of market depths for the US stocks. As shown in these figures, bid depth and ask depth exhibit an almost identical shape of intraday pattern. There are four stocks (MDS, CHDX, GOOG and XOM) exhibiting a U-shaped intraday pattern on both bid and ask depths. DAR has a J-shaped intraday pattern on market depth, while F has an S-shaped intraday pattern on market depth. The results are inconsistent with the inverted U-shaped pattern on market depth observed by Ahn & Cheung (1999); Lee et al. (1993); Li et al. (2005); Vo (2007). These studies also report a negative association between intraday spread and market depth, which is not observed either in this thesis.

Intraday Market Depth for the UK stocks

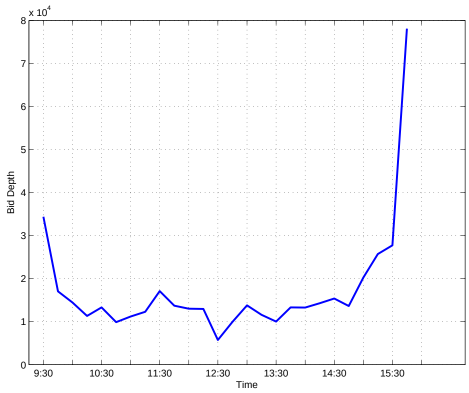
Intraday patterns on market depths for the UK stocks are demonstrated in Figures 5.9-5.11. There are three stocks (BATS, GSK and VOD) exhibiting an S-shaped intraday pattern on market depths. HSBC has a U-shaped intraday pattern on market depths. For BARC and BP, it is clear to see that market depth is very high at market close. As shown in the Figures 5.10-5.11, bid depth has an almost identical intraday pattern with ask depth.

Intraday patterns on near market depths for the UK stocks are demonstrated in Figures 5.12-5.14. All of the six stocks exhibit an S-shaped intraday pattern on near market depth: it is relatively lower at market open, and relatively higher at market close. As seen from the Figures 5.13-5.14, near bid depth has a nearly identical intraday pattern with near ask depth.

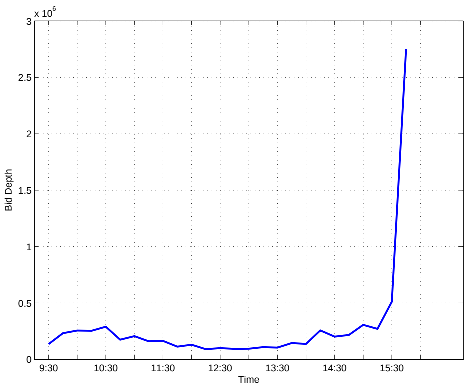
The results on both measures of market depths are inconsistent with the inverted U-shaped intraday pattern observed by Ahn & Cheung (1999); Lee et al. (1993); Li et al. (2005); Vo (2007). However, the near depth (Figure 5.12) shows a negative association with the bid-ask spread (Figures 5.5-5.6), confirming the findings on the relationship between spread and depth in Ahn & Cheung (1999); Lee et al. (1993); Li et al. (2005); Vo (2007). While the near depth is relatively higher during the trading day, that is at market close for the six UK stocks, the spread is relatively lower; while the near depth is relatively lower during the trading day, that is at market open, the spread is relatively higher.



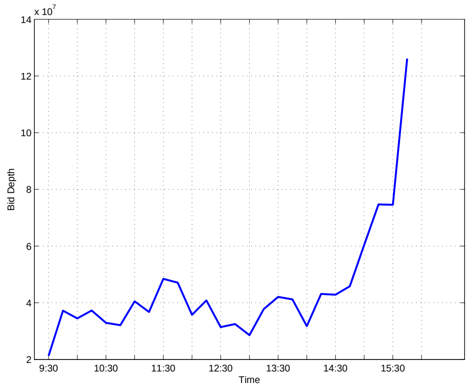
(a) MDS



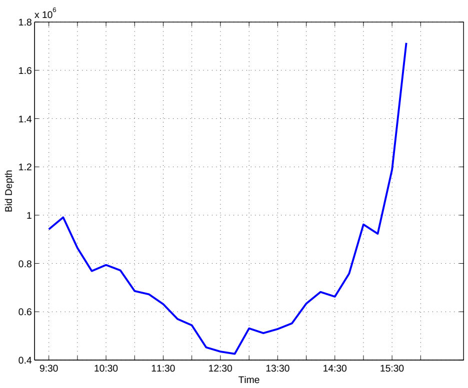
(b) CHDX



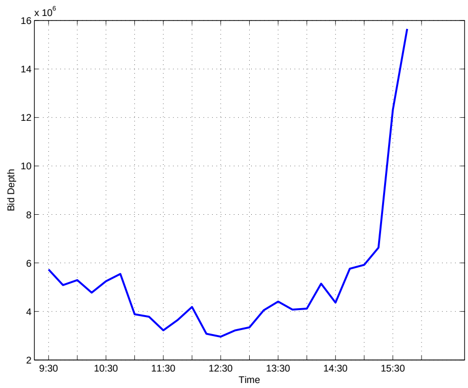
(c) DAR



(d) F

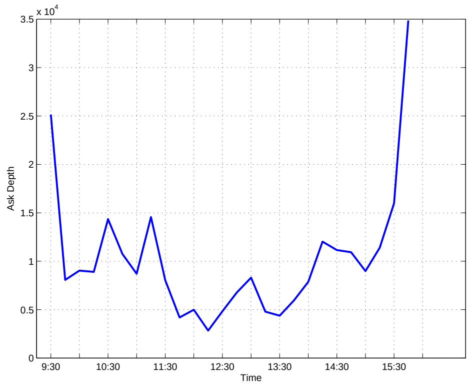


(e) GOOG

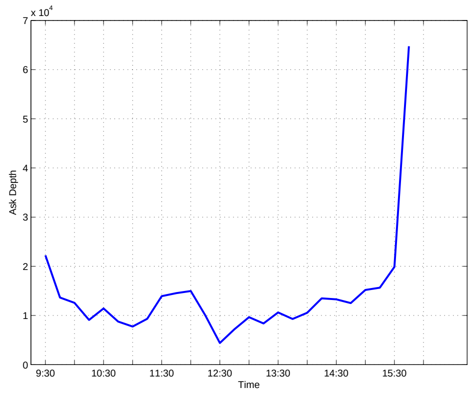


(f) XOM

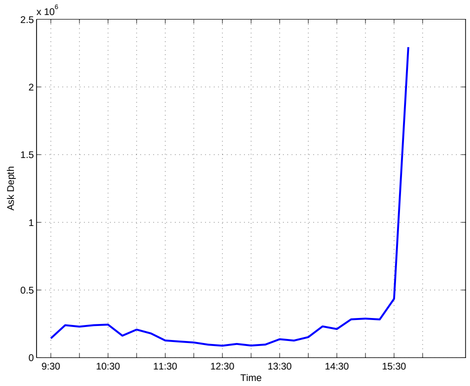
Figure 5.7: Intraday Bid Depth in Share (US stocks)



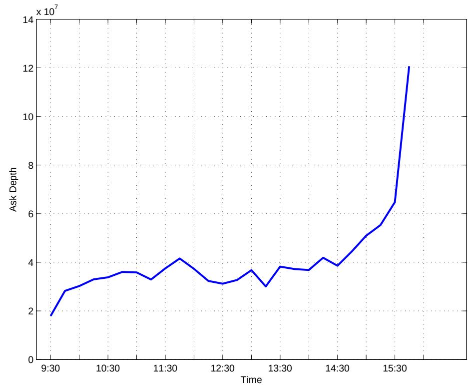
(a) MDS



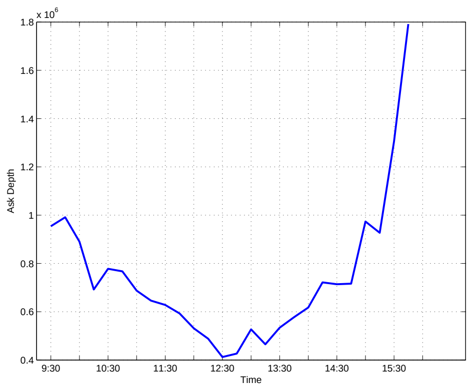
(b) CHDX



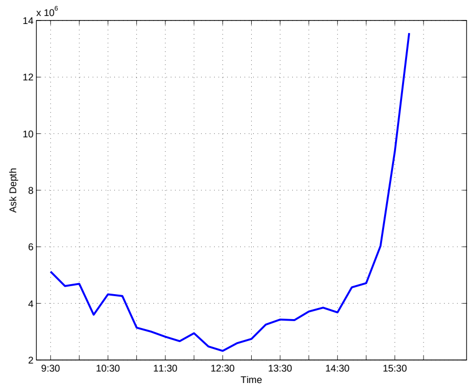
(c) DAR



(d) F

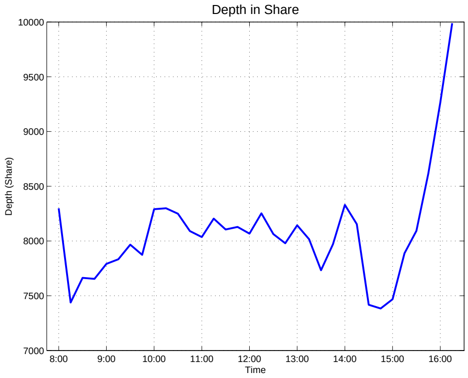


(e) GOOG

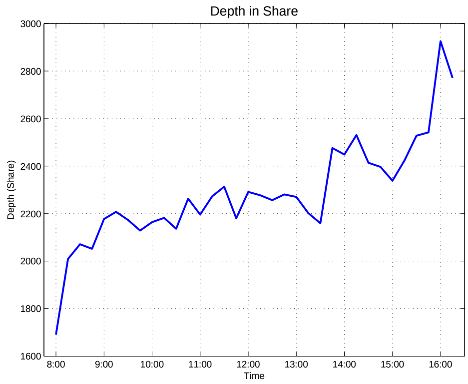


(f) XOM

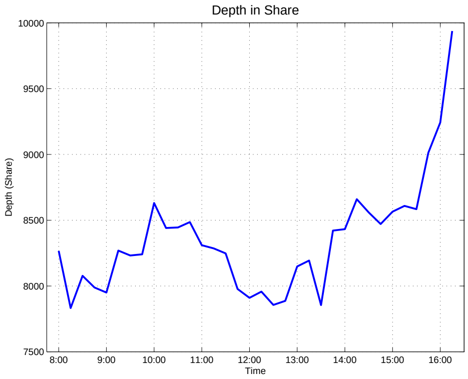
Figure 5.8: Intraday Ask Depth in Share (US stocks)



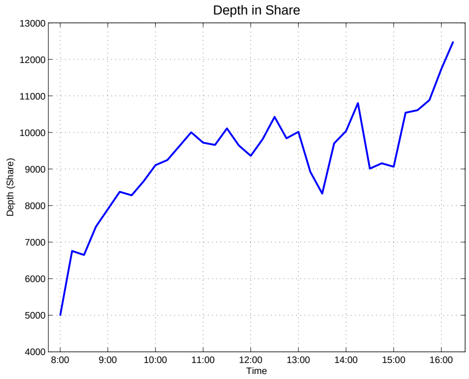
(a) BARC



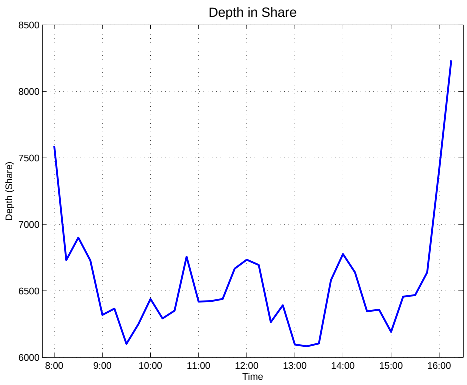
(b) BATS



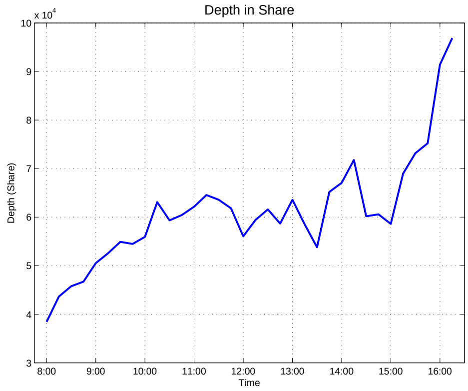
(c) BP



(d) GSK

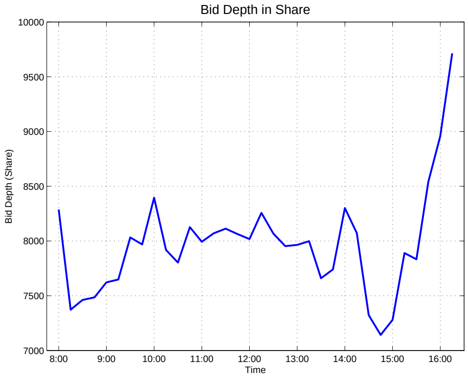


(e) HSBC

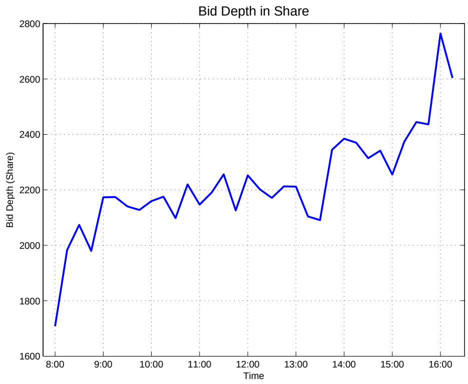


(f) VOD

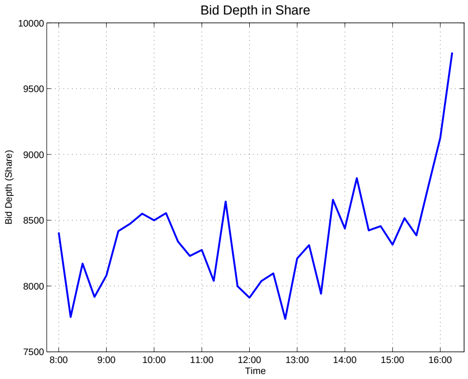
Figure 5.9: Intraday Depth in Share (UK stocks)



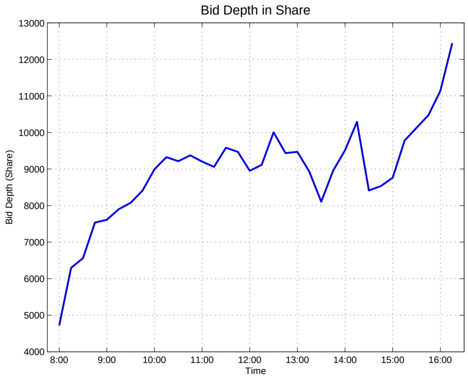
(a) BARC



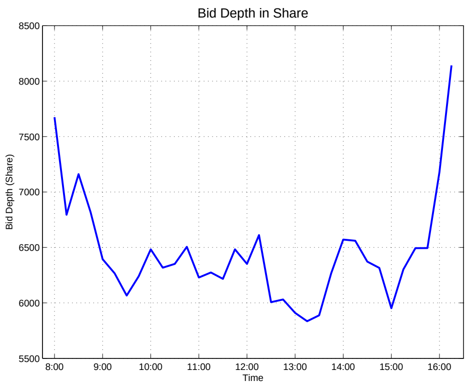
(b) BATS



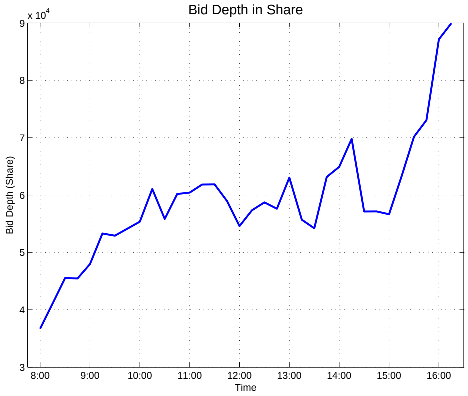
(c) BP



(d) GSK



(e) HSBC



(f) VOD

Figure 5.10: Intraday Bid Depth in Share (UK stocks)

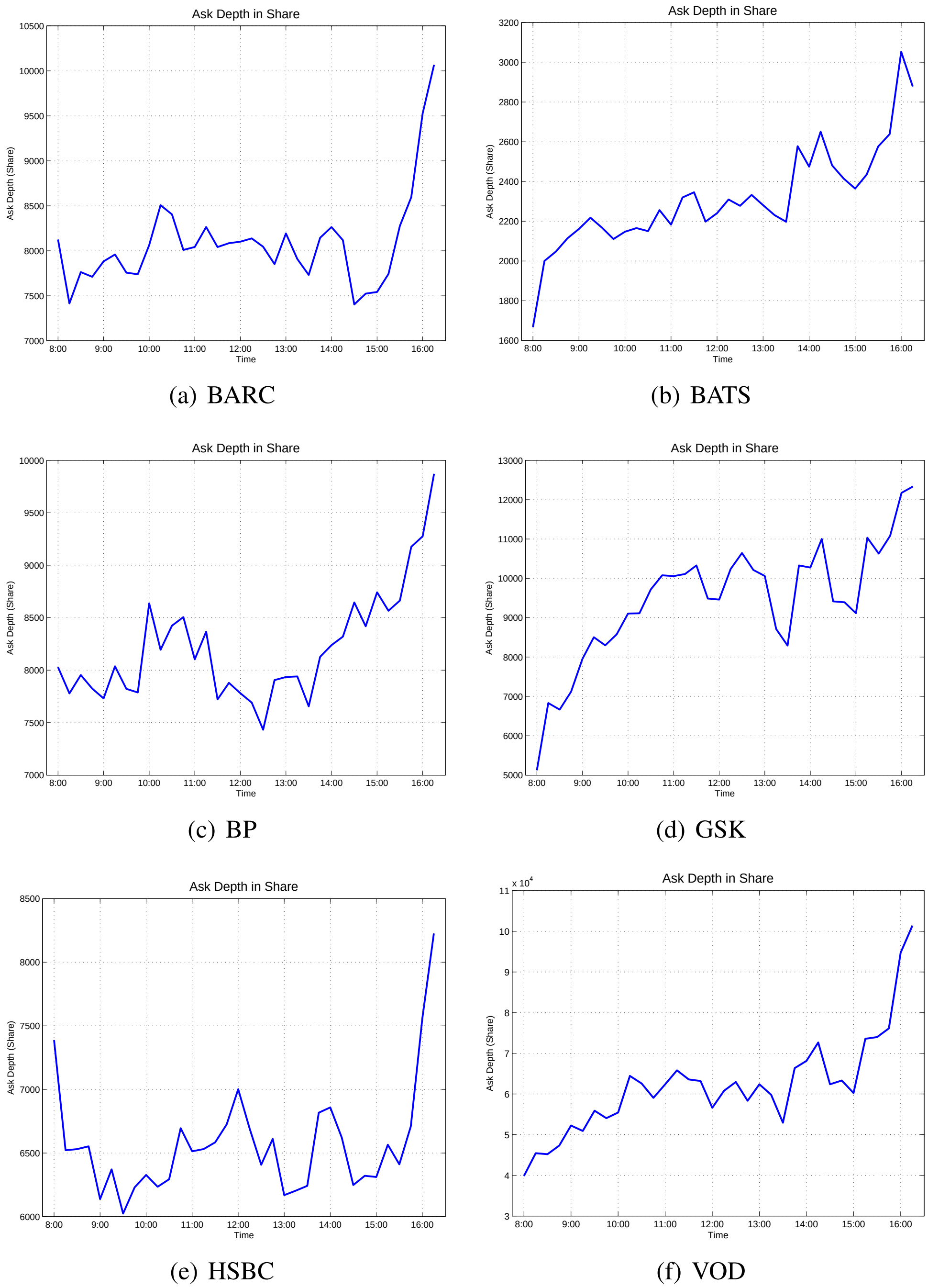
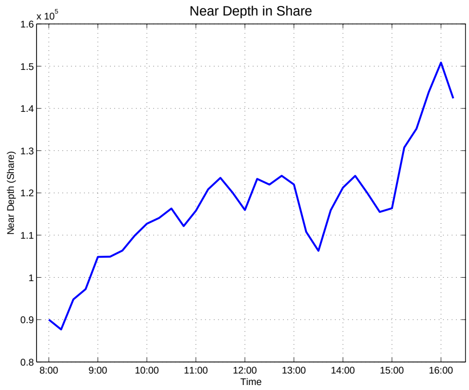
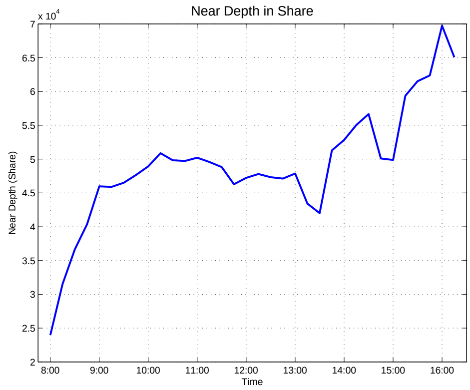


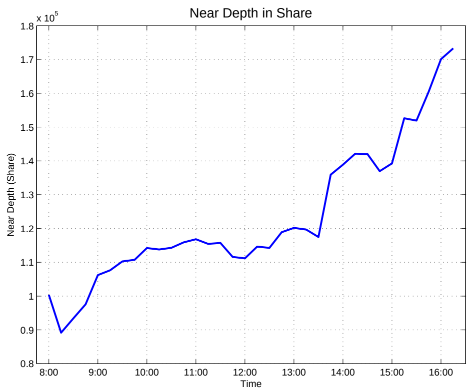
Figure 5.11: Intraday Ask Depth in Share (UK stocks)



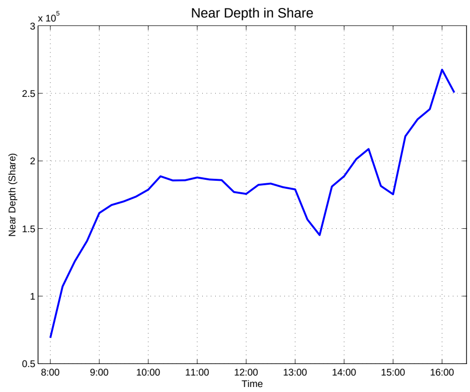
(a) BARC



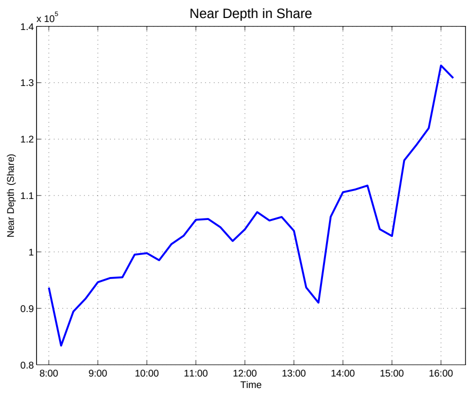
(b) BATS



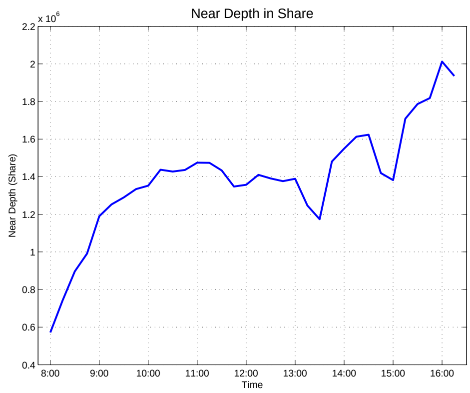
(c) BP



(d) GSK

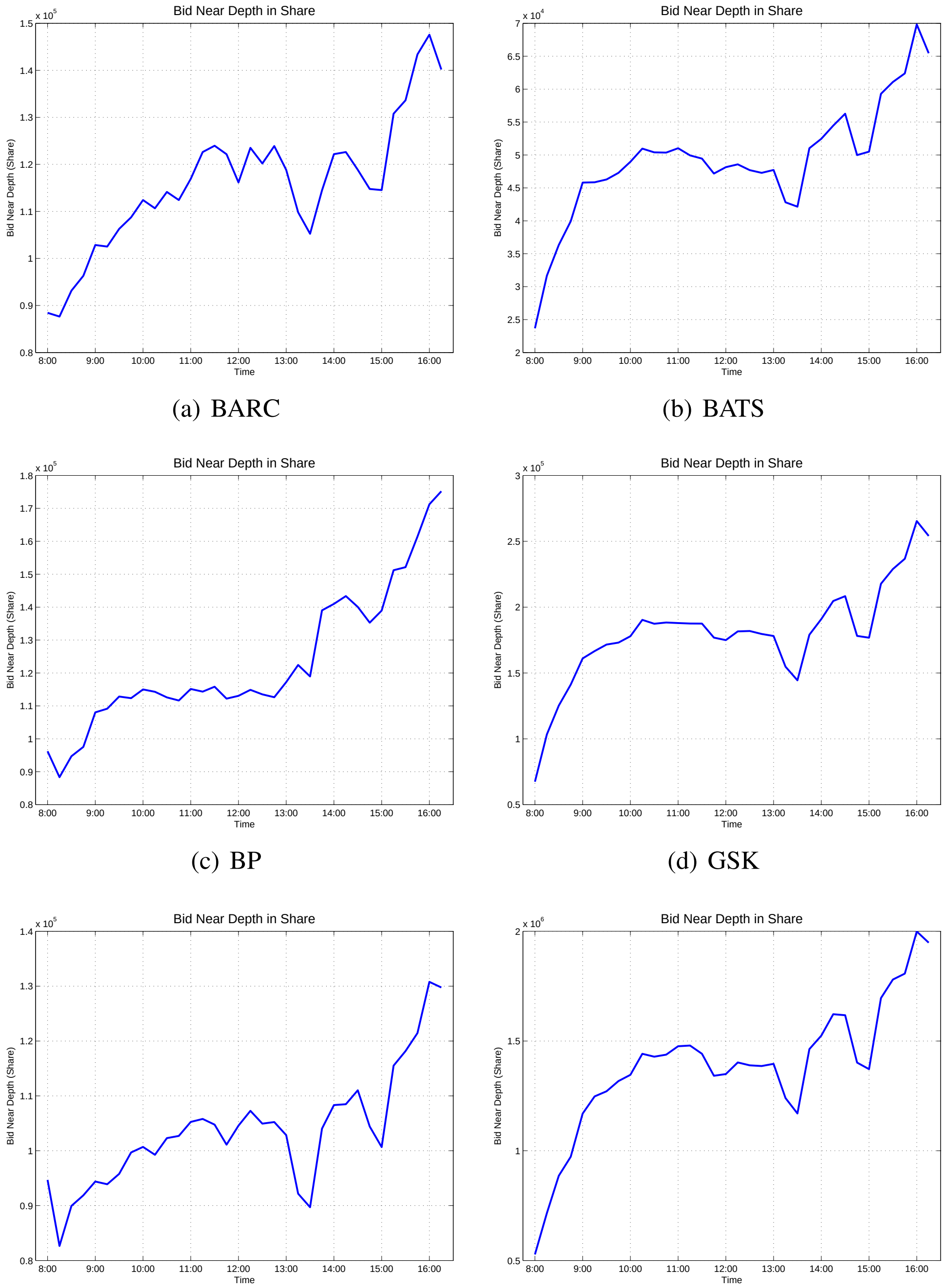


(e) HSBC



(f) VOD

Figure 5.12: Intraday Near Depth in Share (UK stocks)

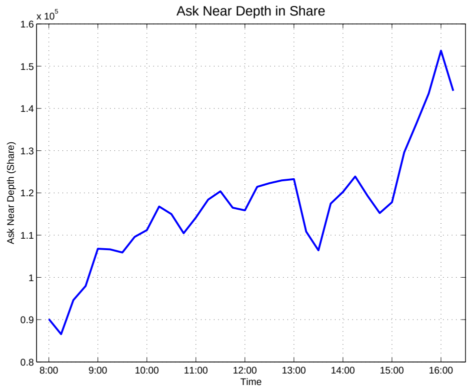


(a) BARC (b) BATS

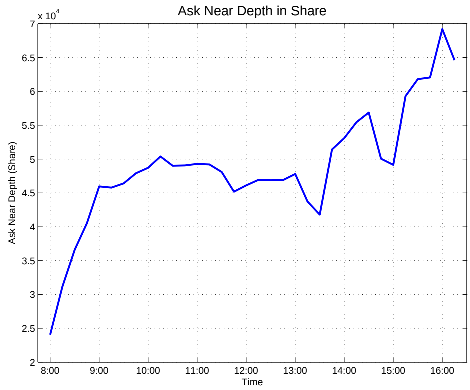
(c) BP (d) GSK

(e) HSBC (f) VOD

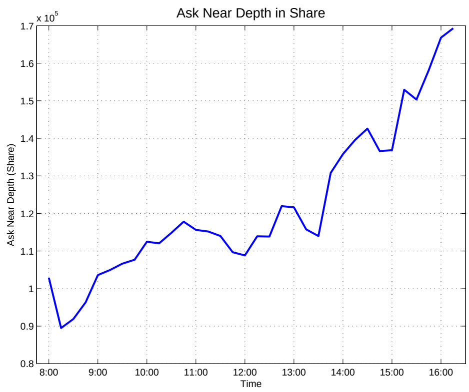
Figure 5.13: Intraday Near Bid Depth in Share (UK stocks)



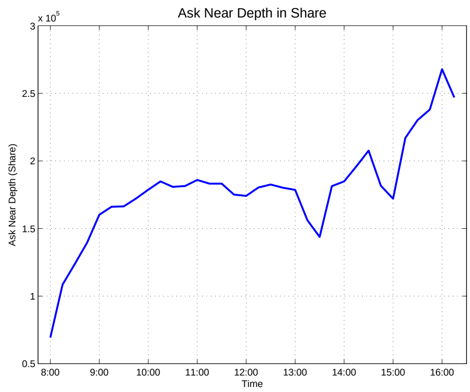
(a) BARC



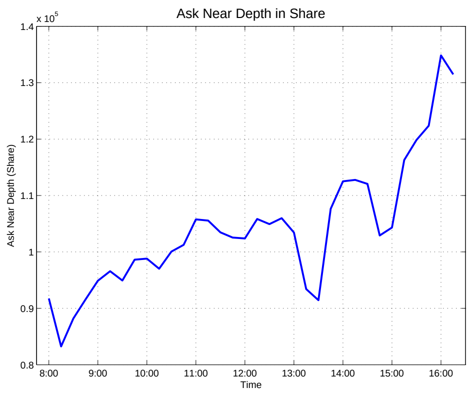
(b) BATS



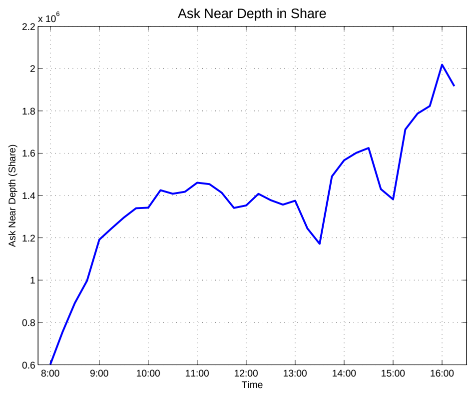
(c) BP



(d) GSK



(e) HSBC



(f) VOD

Figure 5.14: Intraday Near Ask Depth in Share (UK stocks)

5.2.5 Intraday Trade Size and Volume

Volume of trading is a direct measure of trading activity.

Intraday Trading Frequency, Size and Volume for the US stocks

The intraday patterns on trading volume for the six US stocks are shown in Figures 5.15 & 5.16. All of the six stocks have a U-shaped intraday pattern on trading volume. The result confirms the finding in Foster & Viswanathan (1993); Jain & Gun-Ho (1988); Lee et al. (1993); Lockwood & Linn (1990); McInish & Wood (1990b, 1992); Wood et al. (1985) who examine the intraday patterns on the NYSE and the finding in Chan et al. (1995) who examine the intraday patterns in the NASDAQ.

It is interesting to find that trading frequency and trade size also both exhibit a U-shaped intraday pattern as shown in Figures 5.19 & 5.21.

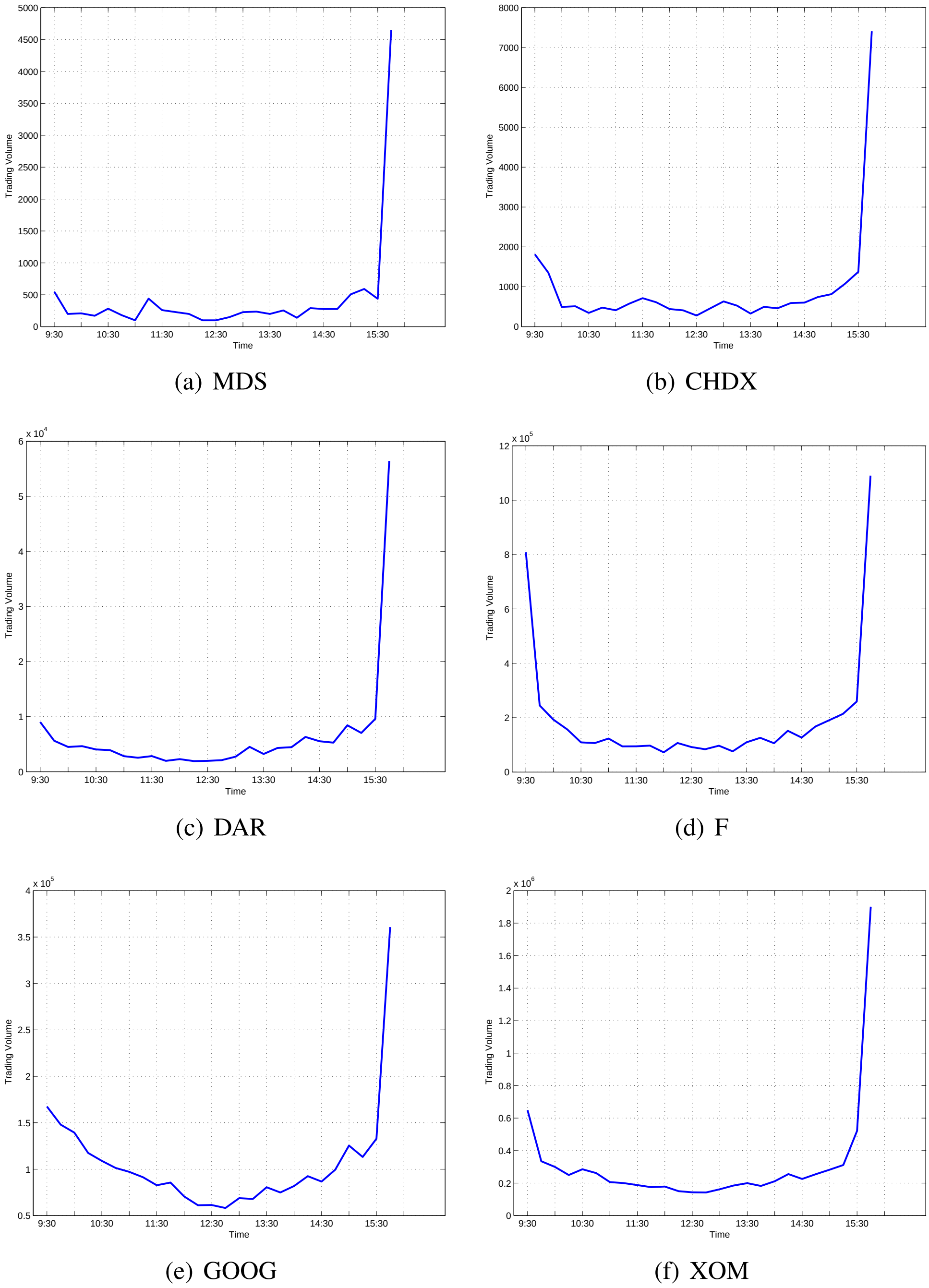
Intraday Trading Frequency, Size and Volume for the UK stocks

Intraday patterns on trading volume for the six UK stocks are shown in Figures 5.17 & 5.18. The six stocks all have a U-shaped intraday pattern on trading volume. Trading frequency shown in Figure 5.20 also exhibits a U-shaped intraday pattern. The results are consistent with the finding of Werner & Kleidon (1996) who examine the intraday pattern on the LSE and the findings for the NYSE market in Foster & Viswanathan (1993); Jain & Gun-Ho (1988); Lee et al. (1993); Lockwood & Linn (1990); McInish & Wood (1990b, 1992); Wood et al. (1985).

It is interesting to observe that there is a sharp increase of trading volume and trading frequency between 2:30 p.m. and 3:00 p.m. for all the six stocks in Figures 5.17-5.20. This sharp increase in trading volume in the afternoon has been also documented by Cai et al. (2004) who investigate the intraday patterns on the LSE. It is due to the open of the US markets, as these six stocks are listed at both the

LSE and the NYSE. US markets opening at 2:30 p.m. GMT positively impact on the trading activities of these stocks in terms of trading volume and trading frequency.

Intraday patterns on trade size for the six UK stocks are shown in Figures 5.22 & 5.23. There are four stocks (BARC, BATS, BP and HSBC) exhibiting a U-shaped intraday pattern on trade size. GSK and VOD have different shapes of intraday pattern with the other stocks: while trade size is very large at market close, it is smallest before market close.

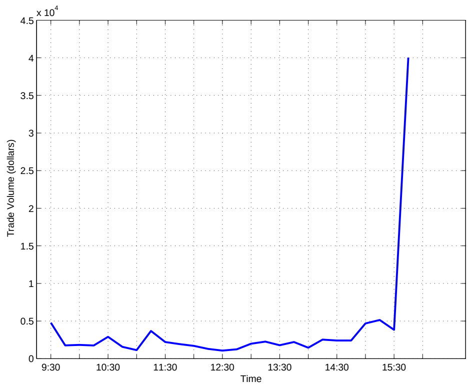


(a) MDS (b) CHDX

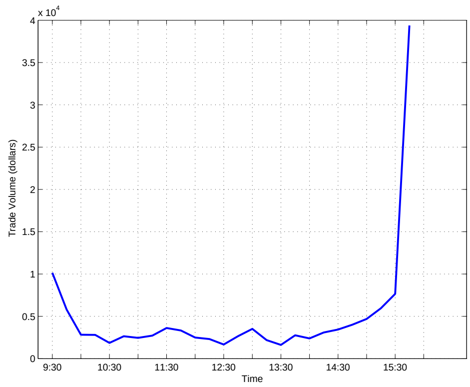
(c) DAR (d) F

(e) GOOG (f) XOM

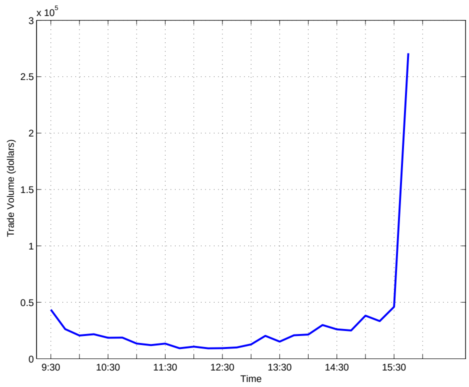
Figure 5.15: Intraday Trading Volume in Share (US stocks)



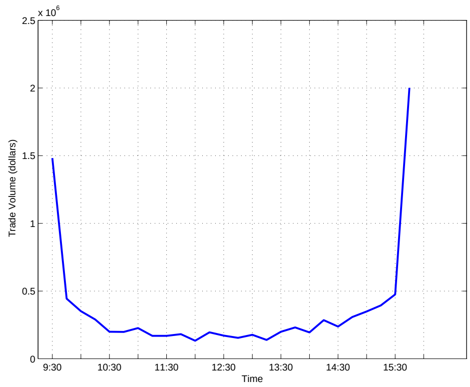
(a) MDS



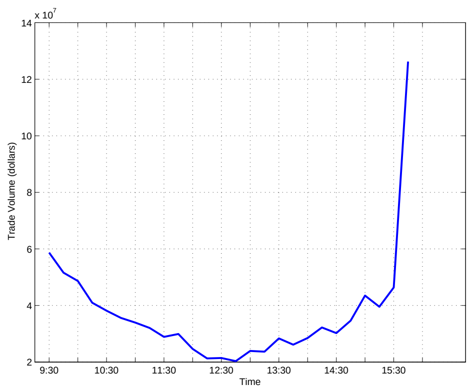
(b) CHDX



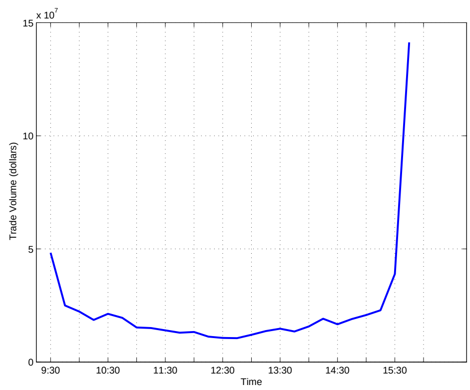
(c) DAR



(d) F

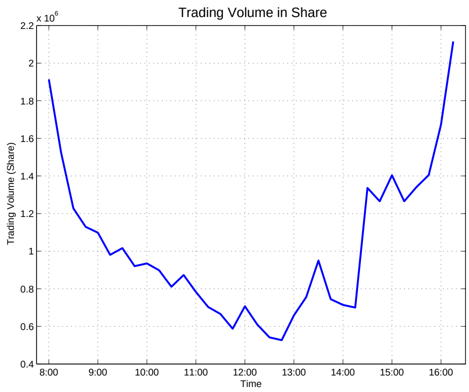


(e) GOOG

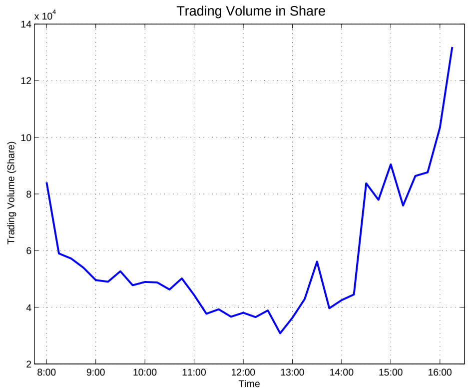


(f) XOM

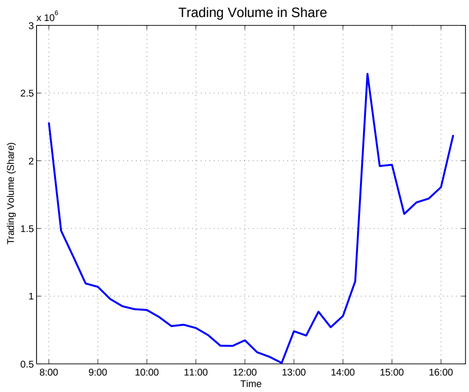
Figure 5.16: Intraday Trading Volume in Value (US stocks)



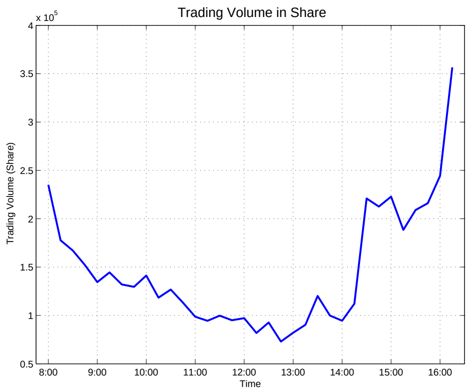
(a) BARC



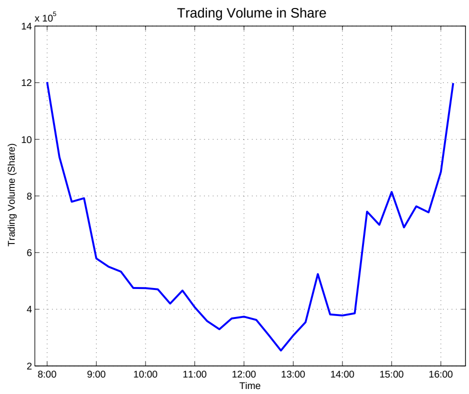
(b) BATS



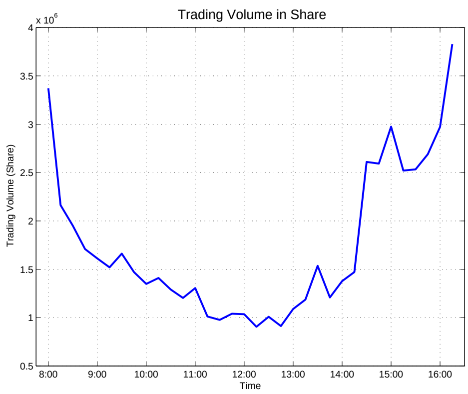
(c) BP



(d) GSK

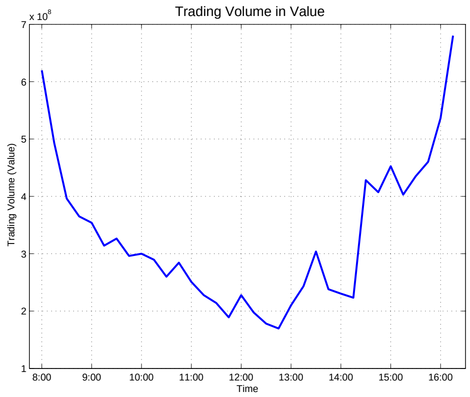


(e) HSBC

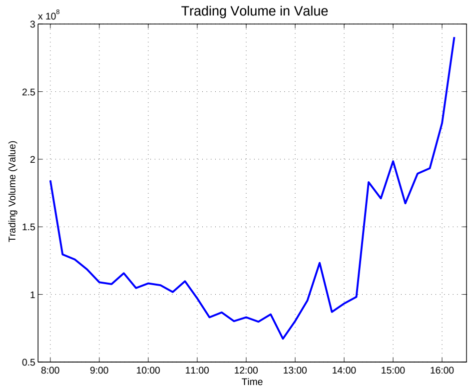


(f) VOD

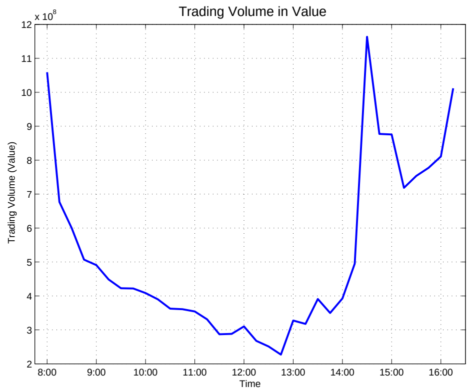
Figure 5.17: Intraday Trading Volume in Share (UK stocks)



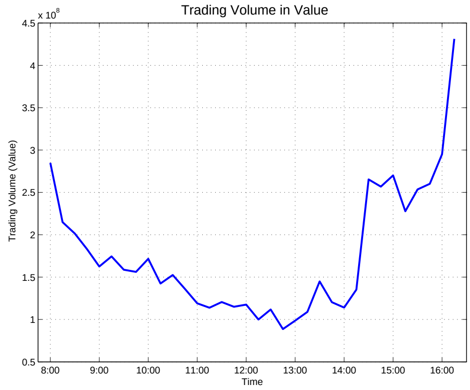
(a) BARC



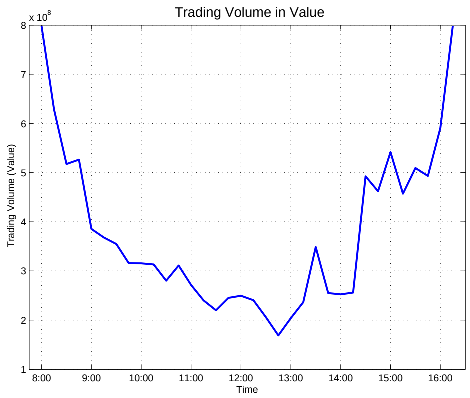
(b) BATS



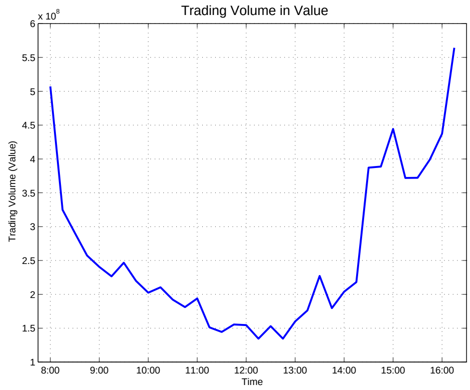
(c) BP



(d) GSK

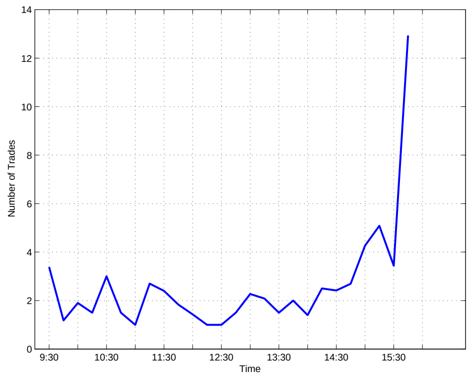


(e) HSBC

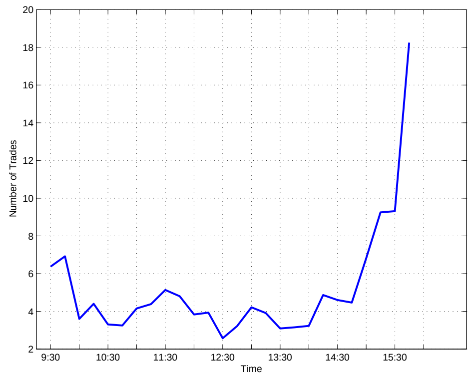


(f) VOD

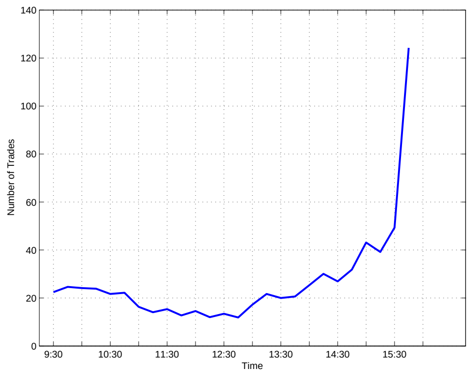
Figure 5.18: Intraday Trading Volume in Value (UK stocks)



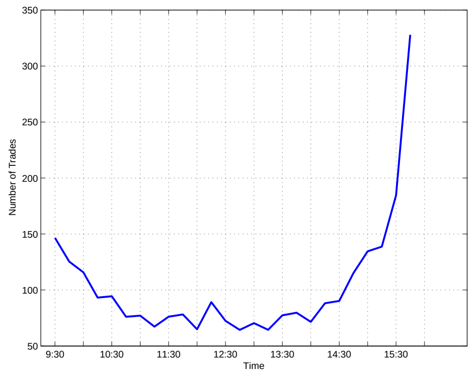
(a) MDS



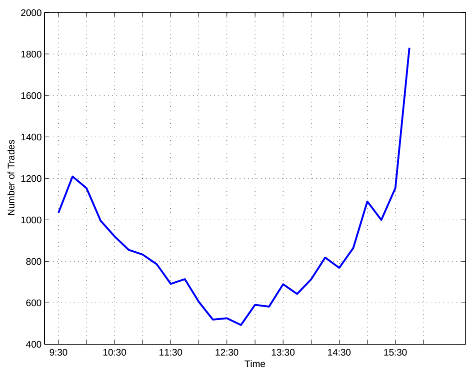
(b) CHDX



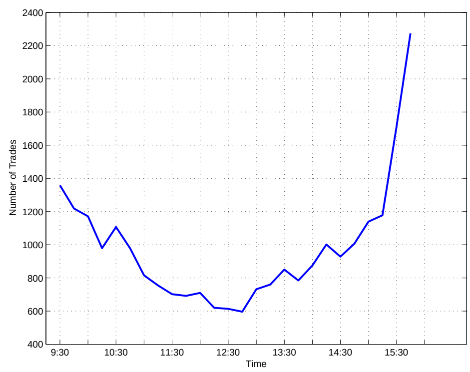
(c) DAR



(d) F

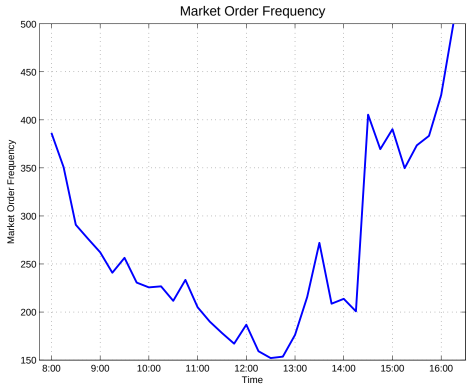


(e) GOOG

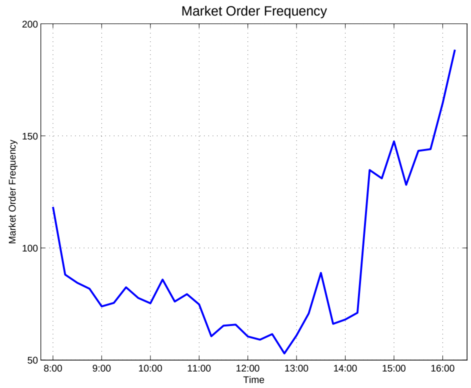


(f) XOM

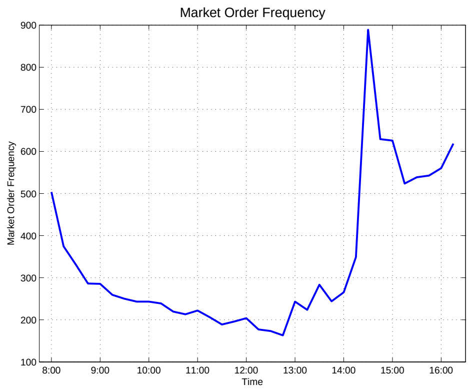
Figure 5.19: Intraday Trade Frequency (US stocks)



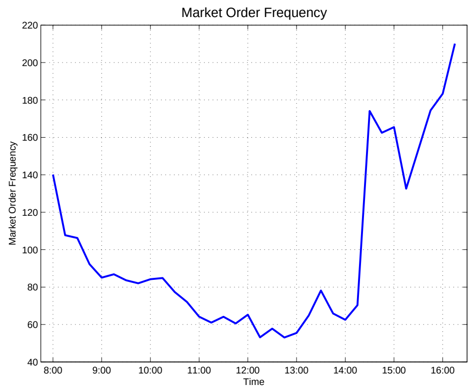
(a) BARC



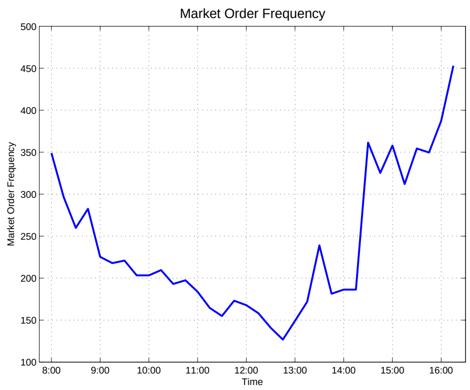
(b) BATS



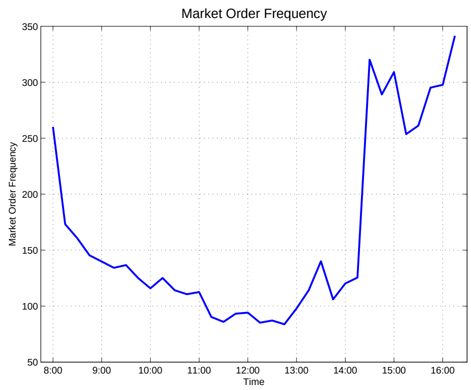
(c) BP



(d) GSK

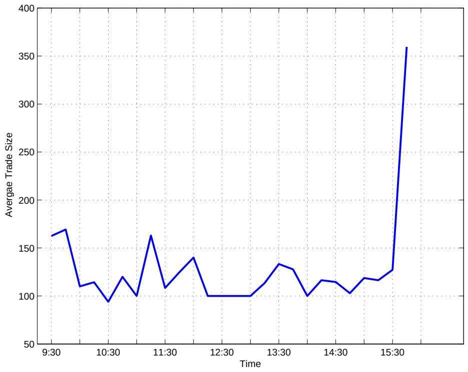


(e) HSBC

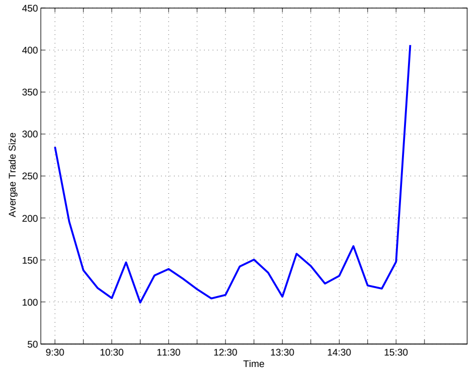


(f) VOD

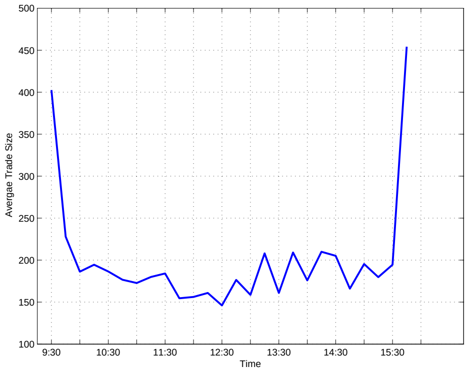
Figure 5.20: Intraday Trade Frequency (UK stocks)



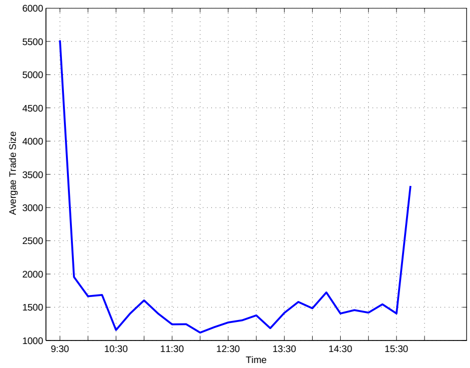
(a) MDS



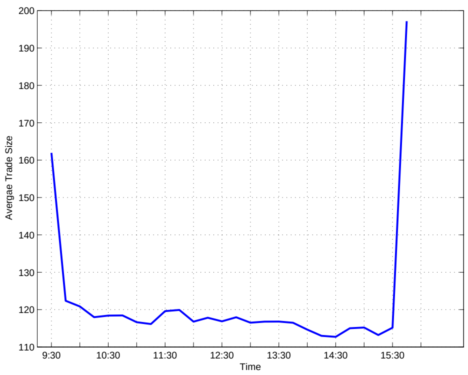
(b) CHDX



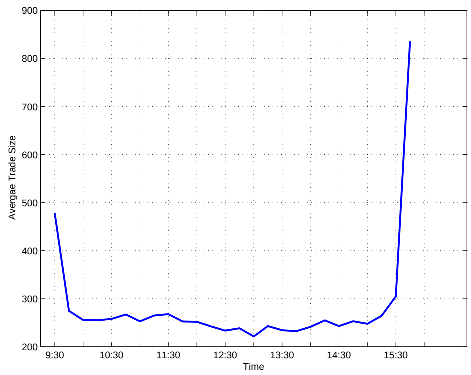
(c) DAR



(d) F

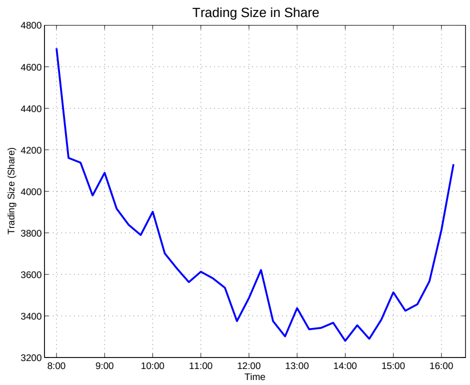


(e) GOOG

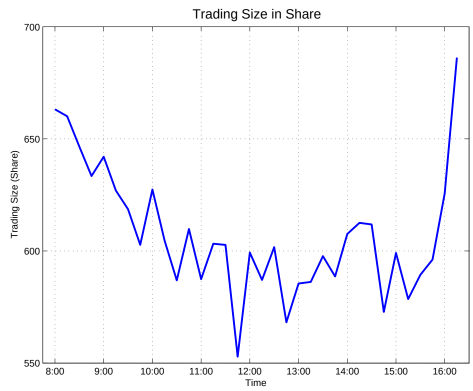


(f) XOM

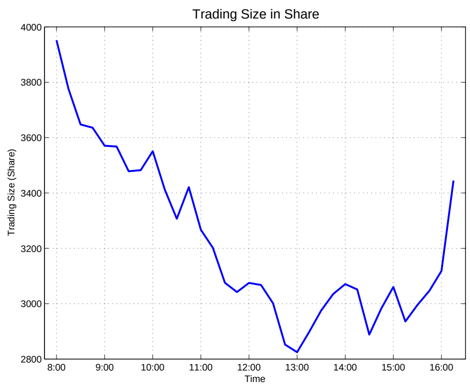
Figure 5.21: Intraday Trade Size (US stocks)



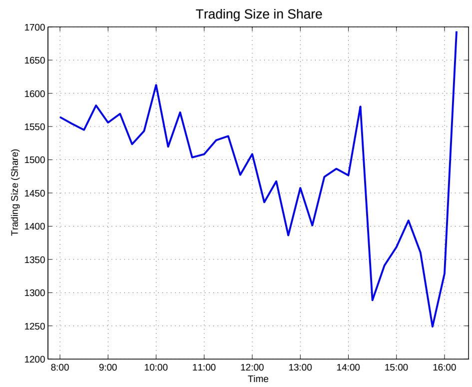
(a) BARC



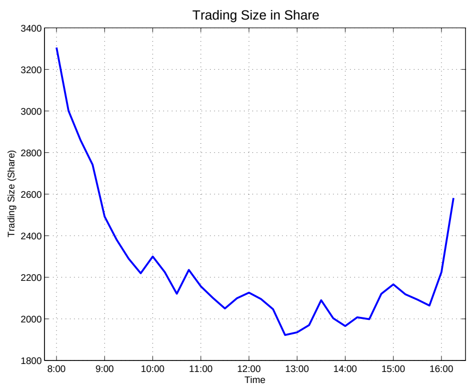
(b) BATS



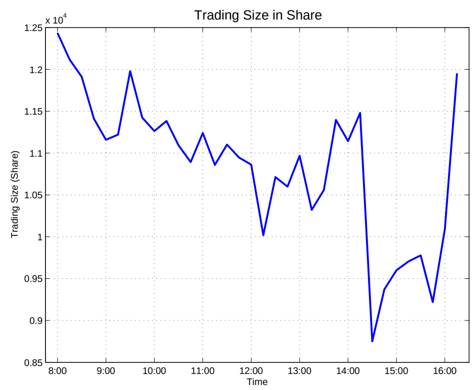
(c) BP



(d) GSK

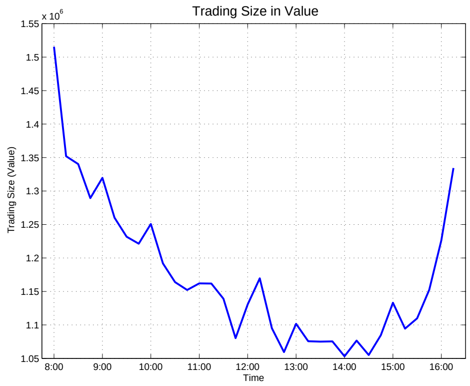


(e) HSBC

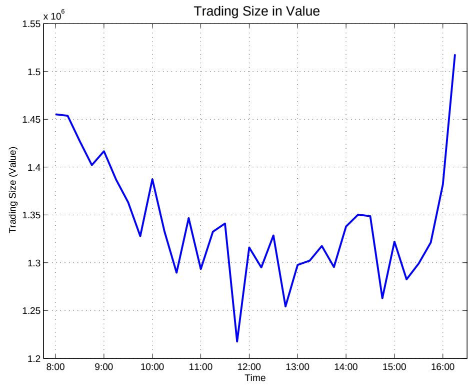


(f) VOD

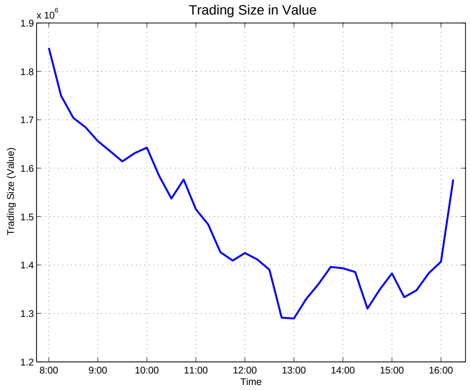
Figure 5.22: Intraday Trade Size in Share (UK stocks)



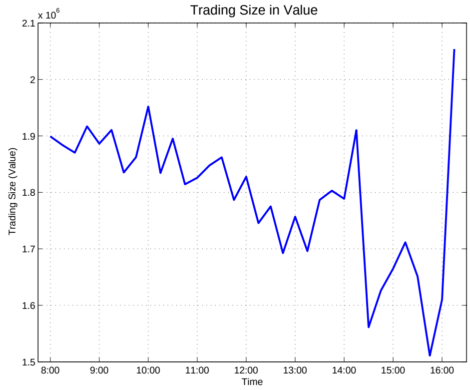
(a) BARC



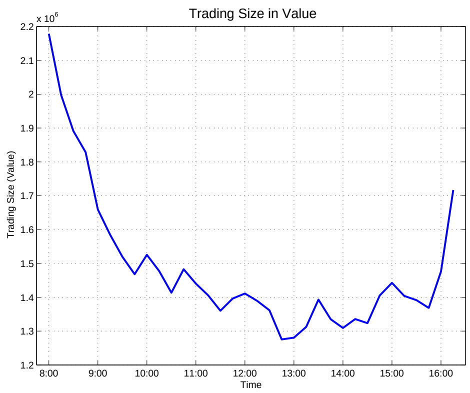
(b) BATS



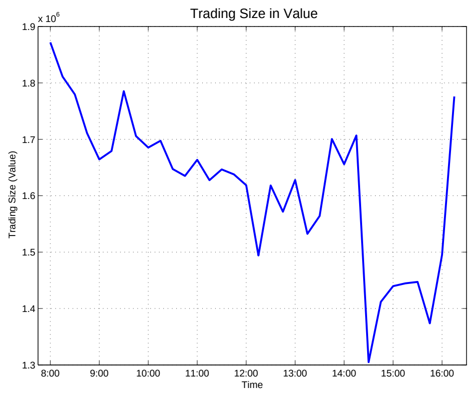
(c) BP



(d) GSK



(e) HSBC



(f) VOD

Figure 5.23: Intraday Trade Size in Value (UK stocks)

5.2.6 Intraday Price Impact

Intraday Price Impact for the US stocks

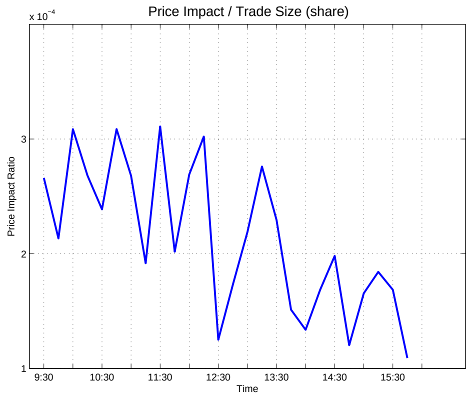
Intraday patterns on price impact for the six US stocks are demonstrated in Figure 5.24. There are four stocks (DAR, F, GOOG and XOM) exhibiting a reverse S-shaped intraday pattern on price impact: it is highest at market open and lowest at market close. For MDS and CHDX, the price impact in the morning is generally higher than the price impact in the afternoon.

Intraday Price Impact for the UK stocks

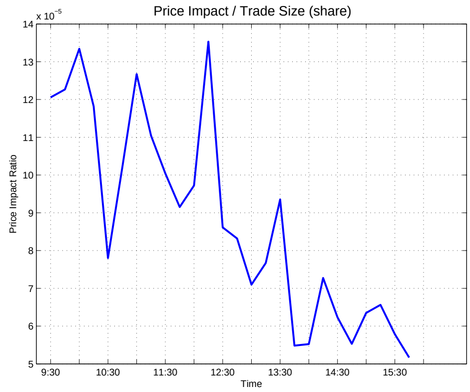
Intraday patterns on price impact for the UK stocks are shown in Figure 5.25. All of the six stocks have a reverse S-shaped intraday pattern on price impact: it is highest at market open, gradually decreases over the trading day, and is lowest at market close. This finding is inconsistent with the U-shaped intraday pattern on price impact observed in the HKSE by Chan (2000). The result indicates that the market liquidity of these six stocks is increasing over the trading day, being lowest at market open and highest at market close. It suggests that time-of-the-day is an important consideration when constructing price impact functions and designing trade execution strategies, which however is ignored in the literature on optimal trading strategy (see the review in Chapter Two).

In contrast to the observation on the HKSE in Chan (2000), price impact does not rebound before the market close as shown in Figures 5.24 & 5.25. This observation cannot be explained by private information as in Chan (2000) who asserts that larger trades reveal more private information leading to larger price impact. As shown in Figures 5.21, 5.22 and 5.23, the average trade size is very large at the market close, negatively associated with low price impact in Figures 5.24 & 5.25.

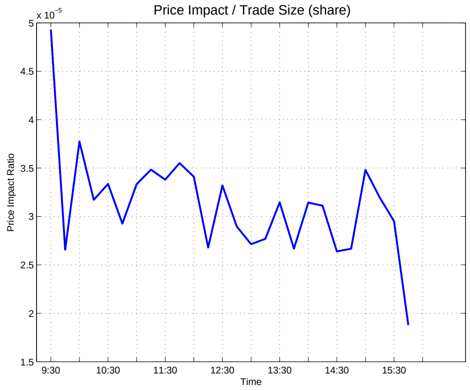
It is interesting to find that price impact (Figure 5.25) has a negative association with near depth (Figures 5.12 & B.7). While near depth is relatively lower during the trading day, price impact is relatively higher; while near depth is relatively higher during the trading day, price impact is relatively lower. The finding confirms the conclusions in Hopman (2007) that price impact is mainly determined by the the order book.



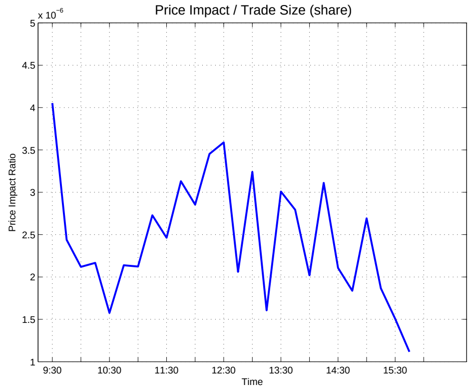
(a) MDS



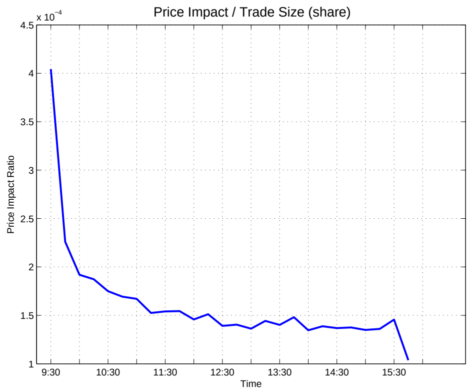
(b) CHDX



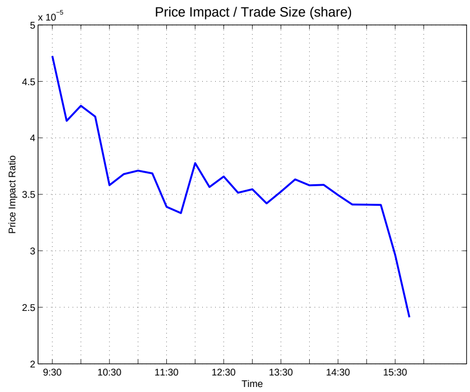
(c) DAR



(d) F

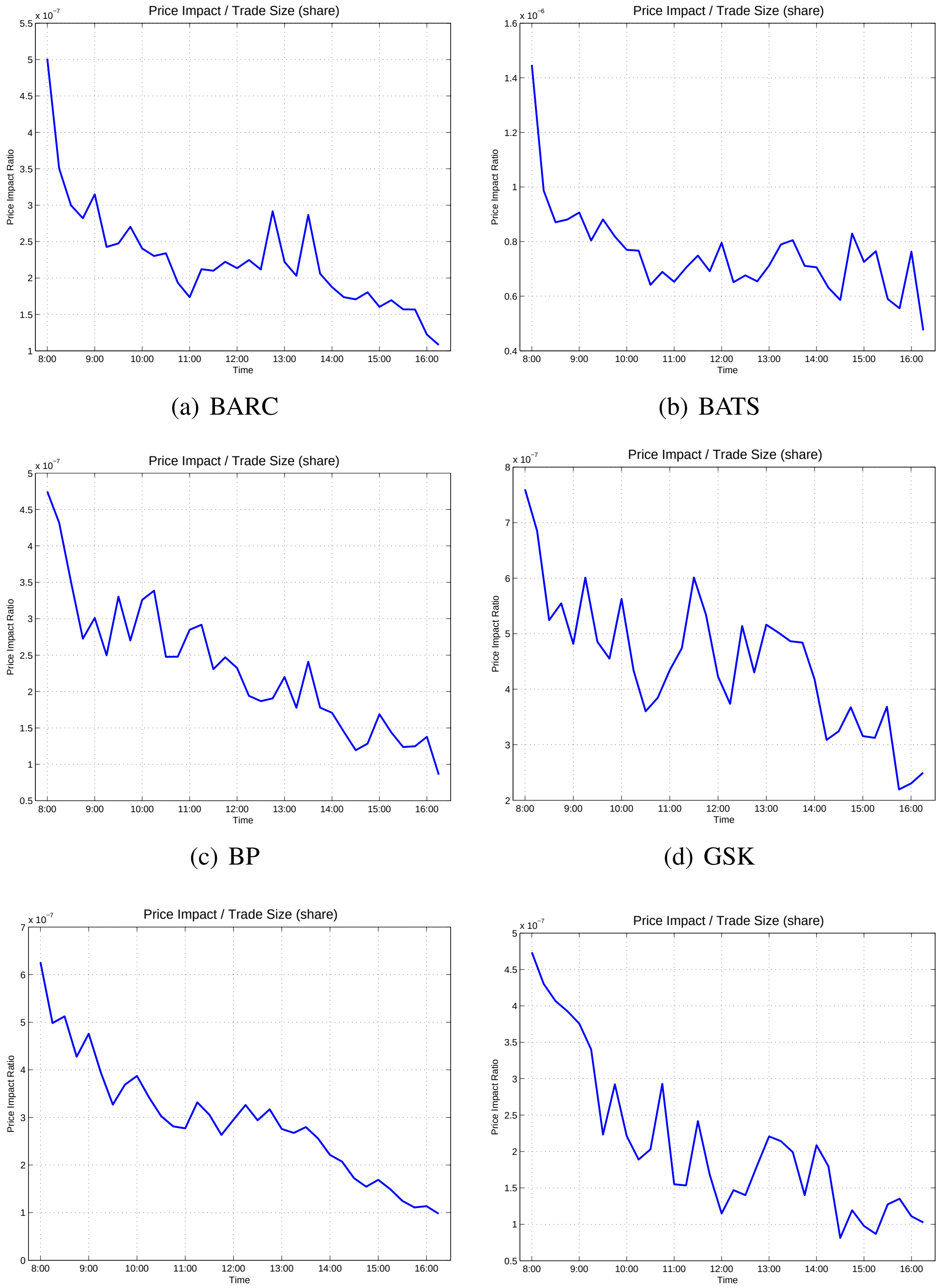


(e) GOOG



(f) XOM

Figure 5.24: Intraday Price Impact Ratio (US stocks)



(a) BARC

(b) BATS

(c) BP

(d) GSK

(e) HSBC

(f) VOD

Figure 5.25: Intraday Price Impact Ratio (UK stocks)

5.3 Summary

This chapter presented an empirical investigation of intraday behaviours of a number of interesting variables using the data drawn from the NYSE-Euronext TAQ database and the LSE ROB database. Six stocks from the US markets and six stocks from the UK markets were selected to perform this task.

This study contributes to the literature by documenting the up-to-date intraday behaviours of a number of liquidity metrics. The findings in this study are of interest to academics and trading desks. A better understanding of intraday behaviours of market liquidity could help theorist in developing comprehensive models. It also emphasises the importance of microstructure tick by tick data, which could be applied to aggregation techniques, with implications beyond the microstructure concerns. The results of this study are summarised as follows.

- Consistent with prior literature, this study documents U-shaped patterns on price volatility for both the US and the UK stocks, consistent with the findings in prior literature (shown in Table 2.2).

- Eleven out of twelve stocks exhibit a reverse S-shaped intraday pattern on bid-ask spread, consistent with results in Chan et al. (1995), Henker & Wang (2006) and Werner & Kleidon (1996).

- The observed intraday patterns on market depths in this study are inconsistent with the inverted U-shaped pattern demonstrated by Ahn & Cheung (1999), Lee et al. (1993), Li et al. (2005) and Vo (2007).

- It is interesting to find that the near market depths of the UK stocks have shown an S-shaped intraday pattern. Moreover, it is found that there is a negative association between the near depth and bid-ask spread, confirming the relationship identified by Ahn & Cheung (1999), Lee et al. (1993), Li et al. (2005) and Vo (2007).

- Trading volume of the 12 stocks examined in this study demonstrates a U-shaped intraday pattern, confirming the findings of most of the studies shown in Table 2.4. Trading frequency also has a U-shaped intraday pattern for all of the 12 stocks. Ten out of twelve stocks exhibit a U-shaped intraday pattern on trade size.

- Ten out of twelve stocks demonstrate a reverse S-shaped intraday pattern on price impact. It is interesting to find that price impact has a negative association with near depth. This indicates that market is relatively more liquid when near depth is relatively higher and price impact is relatively lower, and vice versa. The results provide support for the conclusion in Hopman (2007) that price impact is mainly determined by the limit order book.

A natural question to ask which is worth further research is, what subpopulation of the market participants is constituting the limit order book and what proportion of the market participants is contributing to the trading volume and market depth. To answer this question, further information on the identities of market participants is needed.

Chapter 6

An Investigation of Price Impact and Agent Intelligence

Parts of this chapter have appeared in:

Cui, W. and Brabazon, A. (2012). An Agent-based Modelling Approach to Study Price Impact. *Proceedings of the 2012 IEEE Conference on Computational Intelligence for Financial Engineering & Economics*, pp. 1-8, New York City, USA. (Cui & Brabazon, 2012a)

It is well known that a buying transaction moves the stock price up and a selling transaction drives it down. Price impact is the price change caused by the transaction. In most limit order markets, the order book is transparent to all market participants. A few studies have argued that the degree of price impact is affected by the agent intelligence due to the transparency of limit order books. As suggested by Farmer et al. (2004), the concavity of price impact is due to the agent's selective liquidity taking, namely agents condition the size of their trades on liquidity, executing large orders when liquidity is high and small orders when liquidity is low. Similar arguments are presented in Hopman (2007) and Weber &

Rosenow (2006).

This chapter examines whether the price impact of market orders is affected by the intelligence of the agents who place market orders. The objective of this chapter is to investigate whether the available market information is exploited by "intelligent" agents when placing their market orders, thus affecting the magnitude of the influences of their trades on the asset's price. Novelly, an agent-based modelling is adopted to accomplish this task.

This chapter is organised as follows: a review of related work on agent-based modelling of limit order markets is conducted in Section 6.1; a zero-intelligence based model of limit order market is presented in Section 6.2; a description of experimental setup is given in Section 6.3; results and discussions are provided in Section 6.4, followed by a summary of chapter in Section 6.5.

6.1 Related Work

The development of zero-intelligence models of the continuous double auction markets has a long history. One of the earliest models is built by Bollerslev & Domowitz (1993), who extend the work of Gode & Sunder (1993) by taking into account an order book in order to investigate the effects of order book on market performance. Challet & Stinchcombe (2001) simulate a market which models the Island ECN's order books, while Bouchaud et al. (2002) model the Paris Bourse (now known as Euronext Paris) market. In a series of studies, Zovko & Farmer (2002), Iori et al. (2003), Smith et al. (2003), Farmer et al. (2005b) and Mike & Farmer (2008) simulate a zero-intelligence model of double auction markets based on random order placements to study order flows and market dynamics, and their results demonstrate that the trading mechanism is more important than agent behaviours in determining the statistical properties of the market. In their model, the arrivals of the order flows composed of market orders, limit orders and order cancelations are modeled as Poisson processes. Order sizes are produced from a half-normal distribution.

Another study which simulates a limit order market in the spirit of the ZI idea is conducted by Daniel (2006). In this model, there are 100 ZI traders endowed with the same cash and shares. The arrival of each trade is modeled as a Poisson process. At each arrival time, the trader chooses to buy or sell with probability $1/2$, and decides to submit a market order, or a limit order, or cancels a previous submitted order with probabilities π_m, π_l and $1-\pi_m-\pi_l$ respectively. This model distinguishes two types of limit orders according to order aggressiveness. Traders place a limit order uniformly distributed inside the bid-ask spread with probability π_{in}, or power-law distributed away from the spread with probability $1 - \pi_{in}$. The volumes of incoming limit orders are produced from a log-normal distribution, and the sizes of market orders are the same as those of the best counterpart orders. Their model reproduces the principal stylized facts exhibited by real markets.

The most recent relevant work is the study by Huang et al. (2012). They build an agent-based stock market model consisting of ZI agents to mimic the Taiwan Stock Market (TWSE). Unlike the continuous matching in most limit order markets, order matching on the TWSE is organized every 25 seconds. In their model at each simulated time corresponding 0.01 second of real time, there are five possible events happening in the artificial market: limit order submission, market order submission, order cancelation, order matching and no activity. The probabilities of these events are estimated using real data from TWSE, which are then used for simulating the artificial market. Like Smith et al. (2003), they use a stochastic order size which is generated randomly from a half-normal distribution. In the study, they compare the liquidity costs in the real market to the simulated market, which is measured as the difference of the virtual payment at the disclosed price when the order is entering the market and the actual transaction payment. The result shows that the liquidity cost generated by the simulation data are higher than those for the TWSE data which is possibly caused by their overestimated market order size.

Many zero-intelligence models have been devoted to investigating market behaviours and market structure. A couple of studies have investigated the stylised facts with zero-intelligence markets. These studies, including Ladley & Schenk (2009), Bartolozzi (2010) and Palit et al. (2012), show that some stylised facts can be obtained from the simulated markets with interacting zero-intelligence agents, indicating that many stylised facts of limit order markets do not rely on individual strategic behaviours, but are derived from the interaction of the market mechanism and non-strategic agents. In addition, Duffy (2006) analyses bubbles and crashes with a zero-intelligence market and show that very little trading strategy is is required to reproduce them. LiCalzi & Pellizzari (2007) investigate the performance of four different types of market protocols. Aloud & Tsang (2011) show that the intraday and interweek seasonality of FX market trading activity emerge with a simple agent-based models of zero-intelligence traders.

6.2 A Zero-intelligence based Model of Limit Order Market

The model simulates a limit order market. The trading mechanism follows the price/time priority described in Section 1.2.1 of Chapter 1. Traders are allowed to submit orders at any time, and cancel their limit orders which are not executed. The model extends the model of Smith et al. (2003) by distinguishing different limit orders according to order aggressiveness, considering log-normal distributed order size and taking into account power-law distributed limit order placements in order to come as close as possible to a realistic order flow.

6.2.1 Specification of Model

In the model, there are two groups of agents: buy agents and sell agents. All the buy (sell) limit/market orders are placed by the buy (sell) agent. It is assumed that each hypothetical time in the artificial market corresponds to one millisecond in real market. At each time in the artificial market, a buy agent or a sell agent is chosen with probability $1/2$. The chosen agent at each time can perform one of the four actions as below to fulfill his investment object:

- do nothing,
- submit a market order,
- submit a limit order, or
- cancel an outstanding limit order.[1]

With probability λ_o, the agent will do nothing at all; with probability λ_m, the agent will submit a market order to the market; with probability λ_l, the agent will submit a limit order to the market; with probability λ_c, the agent will cancel a limit order which he previously placed and is not executed. The sum of these probabilities has to be one ($\lambda_o + \lambda_m + \lambda_l + \lambda_c = 1$). Without loss of generality, it is assumed that the agent always cancels the oldest limit order which he previously placed.

If the agent wants to submit a limit order, he can choose from four types of limit order according to the order's aggressiveness. These four types of limit orders are:

- a crossing limit order, which causes an immediate execution,

[1]An outstanding limit order means that the limit order is still listed on the order book

- an inside-spread limit order, which is placed inside the bid-ask spread,
- a spread limit order, which is placed at the best bid (ask) price if it is a buy (sell), and
- an off-spread limit order, which is placed inside the order book with a less attractive price than the best quote.

With probability λ_{crs}, the agent chooses a crossing limit order; with probability λ_{inspr}, the agent uses an inside-spread limit order; with probability λ_{spr}, the agent adopts a spread limit order; with probability λ_{offspr}, the agent places an off-spread limit order. The sum of these probabilities is one ($\lambda_{crs} + \lambda_{inspr} + \lambda_{spr} + \lambda_{offspr} = 1$). It is assumed that crossing buy (sell) limit orders are always placed at the best ask (bid) price.[2] If the size of the crossing buy (sell) limit order exceeds the depth at the best ask (bid) quote, the unexecuted part of the buy (sell) limit order will be placed at the best ask (bid) quote. The inside-spread limit orders are uniformly placed between the best bid quote and the best ask quote. The placement of off-spread limit order follows a power-law distribution with exponent β_{off-p},[3] and various order sizes are allowed in the model, which are generated from log-normal distributions.[4]

6.2.2 Algorithm of Model

An algorithm describing the artificial market simulation process is shown in Algorithm 2.

[2]This assumption is made based on the observation of LSE market that crossing limit orders are rarely placed far away from the best quote on the opposite side of the market.

[3]Empirical work (Zovko & Farmer, 2002; Potters & Bouchaud, 2003) finds that the relative price of the off-spread limit order follows a power-law distribution. The relative price is the difference between the limit price of the order and the best quote.

[4]Empirical work (Maslov & Millis, 2001) finds that order size is roughly distributed like a log-normal distribution with a power law tail.

Algorithm 2 Artificial Limit Order Market

Generate a random number from a uniform distribution $(0, 1)$ and determine the agent type according to the probability 0.5.
switch *Agent* **do**
  case *Buyer*
    Generate a random number from a uniform distribution $(0, 1)$ and determine the action type according to the probabilities $\lambda_o, \lambda_m, \lambda_l$ and λ_c.
    switch *Action* **do**
      case *Do nothing*
        Do nothing at all.
      endsw
      case *Submit a market order*
        Execute a market order with the order size generated from a log-normal distribution with parameters μ_{mo} and σ_{mo}
      endsw
      case *Submit a limit order*
        Generate a random number from a uniform distribution $(0, 1)$ and determine the limit order type according to probabilities $\lambda_{crs}, \lambda_{inspr}, \lambda_{spr}$ and λ_{offspr}.
        switch *LimitOrder* **do**
          case *Crossing Limit Order*
            Submit a limit order at the best ask price with the order size generated from a log-normal distribution with parameters μ_{crs} and σ_{crs}.
            Execute the crossing limit order. If it is not fully executed, place the rest of the limit order at the best ask price on the order book.
          endsw
          case *Inside-spread Limit Order*
            Generate a random value denoted as P_{inspr} from a uniform distribution $(BestBid, BestAsk)$.
            Place a limit order at the price P_{inspr} on the order book with the order size generated from a log-normal distribution with parameters μ_{inspr} and σ_{inspr}.
          endsw
          case *Spread Limit Order*
            Place a limit order at the best bid price on the order book with the order size generated from a log-normal distribution with parameters μ_{spr} and σ_{spr}.
          endsw
          case *Off-spread Limit Order*
            Generate a random value denoted as RP_{offspr} from a power-law distribution with exponent β_{off-p}.
            Place a limit order at the price $(BestBid - RP_{offspr})$ on the order book with the order size generated from a log-normal distribution with parameters μ_{offspr} and σ_{offspr}.
          endsw
        endsw
      endsw
      case *Cancel an outstanding limit order*
        Cancels the oldest outstanding limit order that the buyer previously submitted.
      endsw
    endsw
  endsw
  case *Seller*
    (similar to the case of buyer)
  endsw
endsw

6.3 Experimental Setup

The setup of the experiment is explained in this section.

6.3.1 Parameters of Model

In order to make the model more realistic, the 14 parameters (shown in Table 6.2) of the model are estimated using real data from London Stock Exchange (LSE). One important reason why the LSE market is studied is that it provides data that contain details of every order and every trade which enables us to calculate the parameters of the model. This section provides a brief introduction to the LSE data and describes how to estimate the parameters of the model using the data.

6.3.2 Parameter Estimation

For this study, the stock Barclays Capital is chosen, which is one of the most frequently traded stocks in LSE. The sample data covers details of all orders and trades from 1th April to 30th September 2010. Only the continuous trading session is considered. The records before 8:00:00 and after 16:30:00 are ignored. Table 6.1 shows the descriptive statistics of the trades and orders data for Barclays Capital.

The data covers 126 trading days. During the continuous trading session, there are 30,600,000 milliseconds in each trading day. The numbers of market orders, limit orders and order cancelations[5] and the probabilities of these events on each day are calculated. As it is assumed that only one event occurs at each millisecond in the model, the events occurring at the rest of the trading period are 'do nothing' events. The probability parameters for the four events (do nothing λ_o, limit

[5]We take order deletion and order expiration in ROB data as the same, both are counted as order cancelation events.

Table 6.1: Descriptive Statistics of ROB Data at Sep. 6, 2010 for Barclays Capital

Descriptive Statistics	
Average midquote price:	323.04
Average bid-ask spread:	0.11
Average trade size (pounds):	2,110,606
Average trade size (shares):	6,537
Number of trades:	4,329
Number of order cancelations:	68,783
Number of market orders:	3,972
Number of limit orders:	142,368

order λ_l, cancelation λ_c and market order λ_m,) of the model are estimated as the average daily probabilities from the real data. Similar calculations are performed for the probabilities of the limit order types on each trading day using LSE data. The probability parameters for limit order types (λ_{crs}, λ_{inspr}, λ_{spr} and λ_{offspr}) in the model are estimated as the average daily probabilities over the whole period weighted by the total number of limit orders on each trading day.

6.3.3 Simulation Setup

In the model, each order size is drawn from a log-normal distribution of the functional form:

$$exp(\mu + \sigma * r_{norm})$$

where μ and σ are parameters, r_{norm} is a random number drawn from N(0,1), and exp is an exponential function of the natural number. Each relative limit price is drawn from the power-law function as:

$$xmin \times (1 - r)^{-\frac{1}{1-\beta}}$$

where r is a random number uniformly generated from (0,1), $xmin$ and β are the parameters which need to be estimated. The data from the whole period is used to estimate these parameters for all kinds of orders using the method of maximum likelihood.[6] In order to better fit the power-law distribution, extreme values x_t whose probabilities satisfy $P(x > x_t) < 0.01$ or $P(x < x_t) < 0.01$ are excluded. All the estimated parameters are shown in Table 6.2.

The model is simulated using Matlab. In the model, there are always at least three levels on each side of the order book in order to prevent the order book from being empty. The artificial market runs for 34,200,000 hypothetical milliseconds (corresponding to 9 and a half hours) in each simulation. In some cases, the artificial market may be unstable initially due to the stochastic order flows. Thus, the first 3,600,000 milliseconds (corresponding to 1 hour) are used to warm up the market. Only the later 30,600,000 milliseconds which correspond to the continuous trading session in one trading day are used. The artificial market is run for 30 artificial trading days. Like the LSE, the system has a data recording function which records the details of every trade and every order.

[6]The matlab code for estimating the power-law distributions is developed by the researchers at Santa Fe Institute, which can be downloaded from the website http://tuvalu.santafe.edu/aaronc/powerlaws/.

Table 6.2: Parameters of Artificial Limit Order Market Simulation

Market Settings	**Values**
Initial mid-quote price	300
Initial bid-ask spread	0.5
Tick size	0.01
Agent Type	**Probabilities**
Buy or sell	0.5
Event Type	**Probabilities**
Do nothing	$\lambda_o = 0.9847$
Submit a market order	$\lambda_m = 0.0003$
Submit a limit order	$\lambda_l = 0.0077$
Cancel a limit order	$\lambda_c = 0.0073$
Limit Order Type	**Probabilities**
Crossing limit order	$\lambda_{crs} = 0.0032$
Inside-spread limit order	$\lambda_{inspr} = 0.0978$
Spread limit order	$\lambda_{spr} = 0.1726$
Off-spread limit order	$\lambda_{offspr} = 0.7264$
Order Size Type	**Parameters of Log-normal Distribution**
Market order size	$\mu_{mo} = 7.5663 \quad \sigma_{mo} = 1.3355$
Crossing limit order size	$\mu_{crs} = 8.4701 \quad \sigma_{crs} = 1.1982$
Inside-spread limit order size	$\mu_{inspr} = 7.8709 \quad \sigma_{inspr} = 0.9799$
Spread limit order size	$\mu_{spr} = 7.8929 \quad \sigma_{spr} = 0.8571$
Off-spread limit order size	$\mu_{offspr} = 8.2166 \quad \sigma_{offspr} = 0.9545$
Limit Price Type	**Parameters of Power-law Distribution**
Off-spread relative limit price	$xmin_{off-p} = 0.05 \quad \beta_{off-p} = 1.7248$

6.4 Results and Discussions

The results of this chapter are presented and discussed in this section.

6.4.1 Results of Probabilities of Events and Limit Order Types

After 30 simulation runs of the artificial market, the numbers of events, orders and probabilities were recorded. Table 6.3 shows the means and standard deviations of these over the 30 simulated trading days. An observation from the table is that the events and orders probabilities in the simulated market are very close to those in the LSE market (shown in Table 6.2).

Table 6.3: Events and Limit Order Types in Artificial Market

	Probabilities	**Daily Occurrences**
Order Sign		
buy	0.5000 ± 0.0001	$15,300,854 \pm 2,517$
sell	0.5000 ± 0.0001	$15,299,145 \pm 2,517$
Event Type		
do nothing	0.9847 ± 0.0000	$30,132,401 \pm 722$
submit a market order	0.0003 ± 0.0000	$9,399 \pm 104$
submit a limit order	0.0076 ± 0.0001	$233,785 \pm 483$
cancel a limit order	0.0073 ± 0.0000	$224,413 \pm 472$
Limit Order Type		
crossing limit order	0.0032 ± 0.0001	736 ± 29
inside-spread limit order	0.0977 ± 0.0006	$22,829 \pm 152$
spread limit order	0.1728 ± 0.0010	$40,408 \pm 232$
off-spread limit order	0.7262 ± 0.0011	$169,810 \pm 444$

6.4.2 Summary Statistics of Order Sizes and Relative Limit Prices

The order sizes of different order types in the LSE market and those from the simulated market were compared. Table 6.4 shows some summary statistics of order sizes and relative limit prices for the six-month period (126 trading days) in the LSE data and for the 30-artificial-day period in the simulated data respectively. One can observe that the average order sizes for all order types and the average relative limit prices in the simulated market are very close to those in the LSE.

Table 6.4: Summary Statistics of Order Sizes and Relative Limit Prices

		Min	**Max**	**Mean**	**S.D.**	Observations
Market Order Size	LSE Data	1	789,367	3,916	6,558	1,184,045
	Simulated Data	2	786,637	4,721	10,481	281,970
Crossing Limit Order Size	LSE Data	3	12,500,000	10,562	46,111	93,200
	Simulated Data	36	543,378	9,676	16,716	22,080
Inside-spread Limit Order Size	LSE Data	1	2,476,197	3,902	4,761	2,881,792
	Simulated Data	16	351,554	4,237	5,384	684,870
Spread Limit Order Size	LSE Data	1	3,398,413	3,723	4,782	5,087,395
	Simulated Data	33	183,278	3,867	4,028	1,212,240
Off-spread Limit Order Size	LSE Data	1	3,666,101	5,966	11,910	21,412,078
	Simulated Data	16	564,201	5,839	7,123	5,094,300
Off-spread Limit Order's Relative Price	LSE Data	0.05	100.00	1.29	2.76	21,412,078
	Simulated Data	0.05	99.99	0.93	4.72	5,107,980

6.4.3 Results of Price Impact vs. Trade Size

This study adopts the method used in Lillo et al. (2002) and Lillo et al. (2003) to measure the price impact of a market order. Letting the logarithm of midquote price be p, the price impact caused by a market order is calculated as

$$\Delta p = p_{after} - p_{before}$$

where p_{before} is the price just before the market order arrived at the market and p_{after} is the price immediately after the order is executed. The methods used to measure the trade size vary in previous literature. Lillo et al. (2003) measure it as traded value in dollars divided by the stock value while Hopman (2007) measure it as the number of shares in the order divided by the number of shares outstanding. In this study, the trade size is measured as the shares of the market order divided by the total trading volume in each trading day.

The average behavior of price impact is investigated by dividing the data based on trade size into 10 bins and compute the average price impact for the data in each bin. Figures 6.1 and 6.2 depict the relationship between the price impact and trade sizes in the LSE and the simulated market respectively.

From Figures 6.1, one finds that the relationship between the price impact and trade size is nonlinear and concave. Larger trades do not always have higher price impact than smaller trades. This finding is consistent with the conclusion in previous literature. There are three explanations for this concavity in the literature. The first explanation is that large-order trades have private information about the market price. The second explanation is that large-order traders are patient traders who wait for periods of high liquidity (Farmer et al., 2004). The third explanation is that the order book has deeper liquidities away from the best bid or ask (Bouchaud et al., 2002).

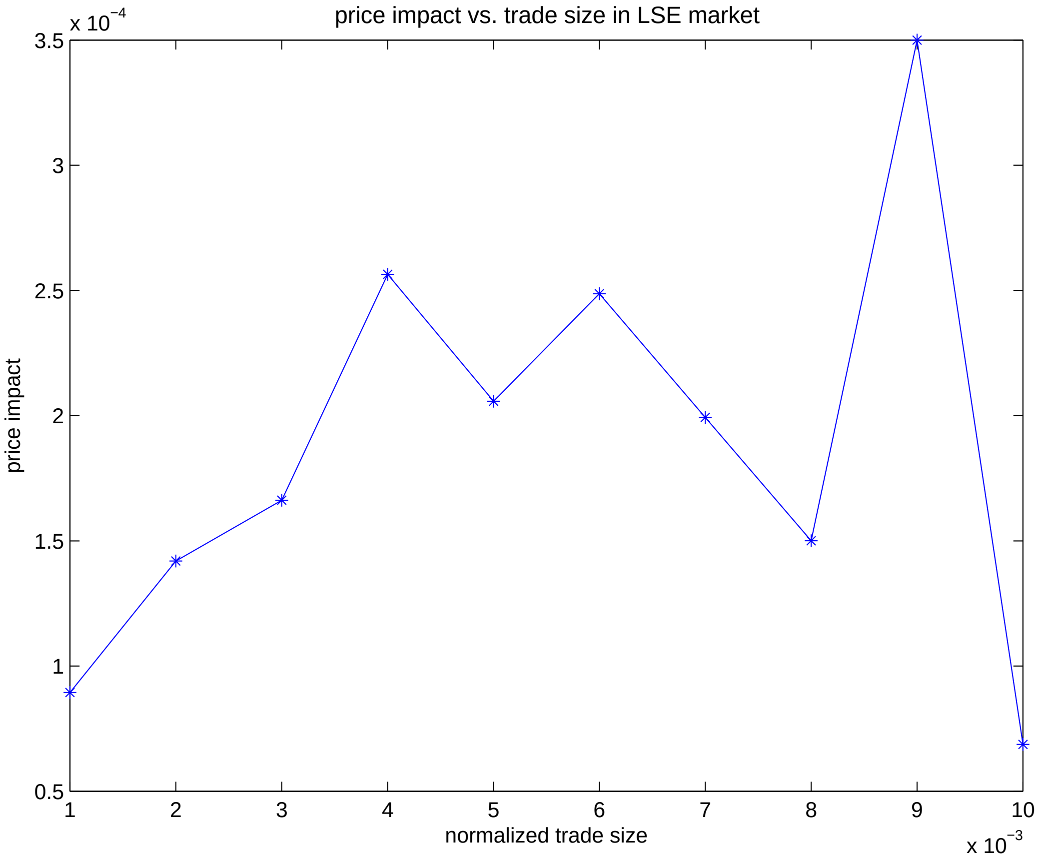


Figure 6.1: Price Impact vs. Trade Size in the LSE market

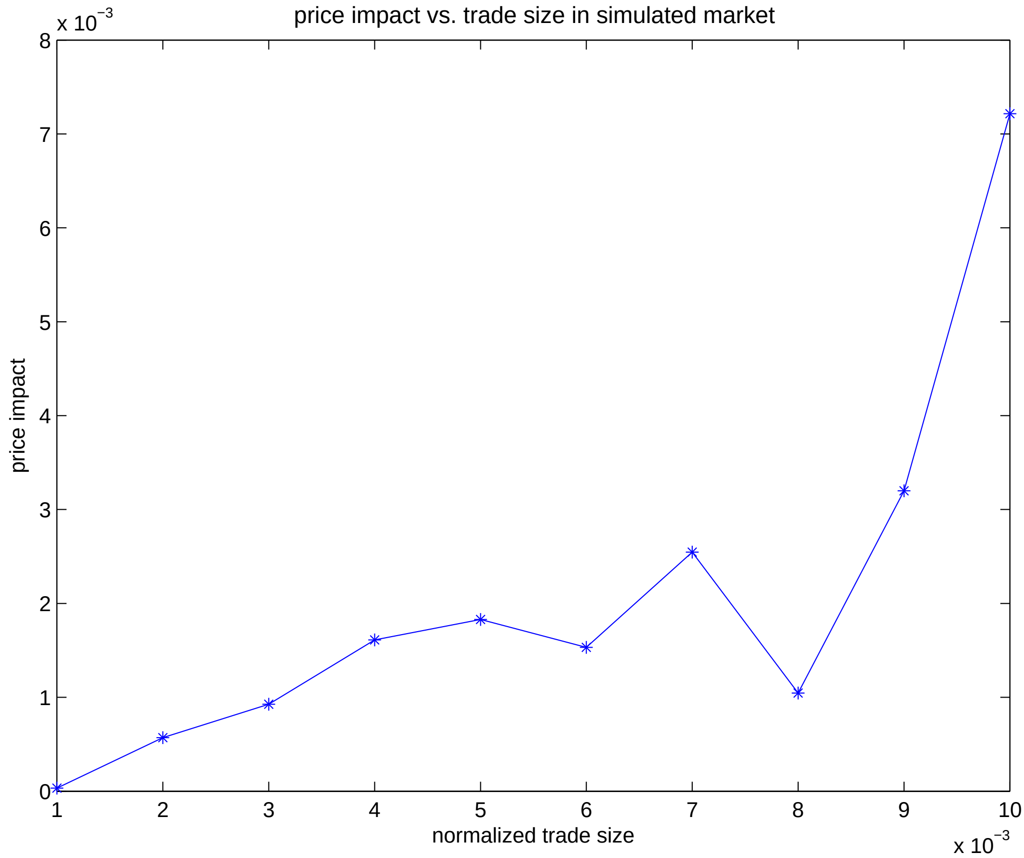


Figure 6.2: Price Impact vs. Trade Size in the simulated market

By comparing the price impacts in the two markets, one can find that generally the price impact in simulated market is larger than that in the LSE market. The price impacts for the smallest trade sizes are very close in the two markets, but the price impacts for larger trade sizes (ranging from $2 * 10^{-3}$ to $10 * 10^{-3}$) in simulated market are bigger than those in the LSE market.

There are two possible reasons to explain the difference on price impact between the simulated market and the real market. One possible reason is due to the difference between the price gaps of limit order book in these two markets. If the limit price gaps in the simulated market are larger than those in the real market, it is possible that the price impact in the simulated market is larger than that in real market. However, this reason can be excluded by the fact that the average relative limit price in simulated market is smaller than that in the LSE market (shown in Table 6.4).

The other possible reason is that traders in the LSE market execute their orders intelligently, for example, by observing the market conditions. The intelligent agents in real market condition their order sizes on market liquidity. Farmer et al. (2004) and Hopman (2007) both find that large orders are traded when the market liquidity is deep and small orders are traded when the market liquidity is shallow. However, the trader agents in the simulated market are 'zero intelligent' who do not care about the market liquidity when trading their orders. Therefore, the price impacts caused by trading large orders in the simulated market are higher than the price impacts in the LSE market, which is shown by the results in Figures 6.1 & 6.2. This indicates that agent intelligence plays an important role in determining the magnitude of price impact, and suggests that agent intelligence is needed when simulating an artificial market which tries to replicate the relationship between price impact and market order size.

6.5 Summary

Many studies have shown that the price impact of a trade is a concave function of its trade size, which is often called the "concavity" of price impact. Recently, a number of studies (Farmer et al., 2004; Hopman, 2007; Weber & Rosenow, 2006) argue that the concavity is due to selective liquidity taking: traders submit large orders when the market is more liquidity, and small orders when the market is less liquid. In an earlier study, Admati & Pfleiderer (1988) also argue that both liquidity and informed traders prefer to trade when the market is thick, that is when their trading has little effect on prices. This chapter investigates whether agent intelligence plays an important role in determining the magnitude of price impact. An agent-based modeling approach is adopted to perform this task. The advantage of this approach is that it is easy to analyse one specific factor by isolating it in the simulation. An zero-intelligence based artificial market was simulated in order to examine the price impact caused by the zero-intelligence agents.

This chapter provides evidence supporting that agent intelligence is a critical factor in determining the magnitude of price impact. It contributes to the literature on price impact. The finding is of interest to academics and practitioners. It suggests that agent intelligence is an important factor when modelling price impact. This chapter also contributes to the literature on agent-based modelling of financial markets. It indicates that agent intelligence is a necessary condition when simulating an artificial market where replicating realistic price impact is a concern.

Chapter 7

Dynamic Trade Execution with Limit and Market Orders

Much of the content of this chapter has been in various publications:

- Cui, W., Brabazon, A. and O'Neill, M. (2011). Adaptive Trade Execution: A Grammatical Evolution Approach. *International Journal of Financial Markets and Derivatives*. Vol. 2, No. 1/2, pp. 4-31. (Cui et al., 2011)

- Cui, W., Brabazon, A. and O'Neill, M. (2010). Evolutionary Computation and Trade Execution. *Natural Computing in Computational Finance (volume III)*, pp. 51-69, Springer-Verlag, Berlin. (Cui et al., 2010b)

- Cui, W., Brabazon, A. and O'Neill, M. (2010). Evolving Efficient Limit Order Strategy Using Grammatical Evolution. *Proceedings of the 2010 IEEE Congress on Evolutionary Computation*, pp. 1-6, Barcelona, Spain. (Cui et al., 2010d)

- Cui, W., Brabazon, A. and O'Neill, M. (2010). Evolving Dynamic Trade Execution Strategies Using Grammatical Evolution. *Proceedings of the*

2010 International Conference on Applications of Evolutionary Computation, Lecture Notes in Computer Science 6025, pp. 192-201. (nominated as best paper) (Cui et al., 2010c)

- Cui, W., Brabazon, A. and O'Neill, M. (2009). Efficient Trade Execution Using a Genetic Algorithm in an Order Book Based Artificial Stock Market. *Proceedings of the 11th Annual Conference Companion on Genetic and Evolutionary Computation Conference*, pp. 2023-2028. (Cui et al., 2009a)

Trading cost is a significant concern for financial investors as it directly impacts on investment returns. As shown by Perold (1988), a hypothetical portfolio without loss caused by execution costs outperforms the actual portfolio by almost 20% per year during the period from 1965 to 1986. In another study, Wermers (2002) find that mutual funds outperform the market by 1.3 percent per year before accounting for transaction costs, but underperform the market by 0.3 percent after accounting for expenses and transaction costs while investigating mutual fund performance. Trading cost is composed of broker commission, fee, bid-ask spread, price impact and opportunity cost. Empirical studies (Chan & Lakonishok, 1995, 1997; Collins, 1991; Kraus & Stoll, 1972) have identified that the dominant component is price impact. Price impact is variable and depends on market liquidity. Many studies have shown that a small-quantity market order can cause significant impact on an asset's price when volume at price levels on the order book is sparse at the time the order is submitted to the market.

Controlling price impact is an important issue in financial markets. It is especially important when trading large-quantity orders as their impact on the market can be significant. In electronic limit order markets, this trading process is implemented by computers automatically according to predefined trading algorithms. In recent years, there is a growing literature on trading algorithms that are designed to minimise the price impact of trading a large order. As reviewed in Section 2.5.1 of Chapter 2, one of the key assumptions in these models is that only market orders are adopted.

This chapter aims to shed light on the effect of order choice on the price impact of trading a large order. To the best of my knowledge, this is the first study which specifically looks at this issue in the context of designing optimal trade execution strategies. This chapter finds that order choice is an important determinant of the design of optimal trading strategies. Using limit orders can help to minimise the price impact of trading large quantities of orders. As the important task of providing liquidity in limit order markets is assigned to the complex trading interactions enabled by the emergence and disclosure of the limit order book, the state of the limit order book is extremely important for practitioners who use it to optimise their trading strategies. In this chapter, this consideration is also accounted in the design of execution strategies.

This chapter is organised as follows. A review of related work on trade execution and the genetic algorithm is given in Section 7.1; experimental design is described in Section 7.2; experimental setup is explained in Section 7.3; empirical results are documented in Section 7.4; a summary of this chapter is given in Section 7.5.

7.1 Related Work

Despite the importance of optimising trade execution, there has been relatively little attention paid in the literature to the application of evolutionary methods for this task. One notable exception is Lim & Coggins (2005b) who applied a Evolutionary Algorithm to evolve a trade execution strategy using order book data from a fully electronic limit order market, the Australian Stock Exchange (ASX).

In their study, a large order is to be completed within one trading day. The order is divided into ten child orders which are submitted to the market at regular intervals of half an hour. The relative sizes of these child orders are determined according to share volume trading patterns, which typically follow a U-shaped pattern with increased volumes trading at the open and close. The child orders are placed into the market as limit orders at the best available price and the Evolutionary Algorithm is used to find the optimal lifetime that a limit order would remain on the order book (if it had not already been executed) before it was automatically ticked over the spread to close out the trade. The fitness function was the Volume Weighted Average Price (VWAP) performance of that strategy relative to the benchmark daily VWAP. Each strategy was trained on three months' worth of transaction-level data using a market simulator. The results were tested out of sample on three highly liquid stocks and tested separately for sell side and buy side. The in sample and out of sample performances of the evolved strategies were better than those of pure market order strategies and pure limit order strategies. In Lim & Coggins (2005b), the strategies were tested using historical data. An agent-based artificial market is used instead in the present study.

7.2 Designs of Experiments

In order to investigate whether order choice is an important factor in determining price impact when trading a large order, three experiments are conducted in this chapter where three different combined strategies adopting different types of orders and benchmark strategies using one type of orders are designed. The three combined strategies are named Combined-I Strategy, Combined-II Strategy, and Combined-III Strategy.

In the first experiment, the aim is to examine whether order choice between various types of limit orders with different levels of aggressiveness plays an important role in determining the price impact when trading a large order. Empirical evidence suggests that the state of the order book affects order placement strategies. This effect is taken into account in the second and the third experiments, where the trader is allowed to use market information to determine his order choice. In the second experiment, the trader can place a limit order at three levels of aggressiveness, aggressive limit order, moderate limit order, and passive limit order, which are explained below. The aim of the second experiment is to examine whether the choice between these three types of limit orders influences the price impact of trading a large order. In the third experiment where the trader can use a combination of market order and moderate limit order, the objective is to investigate whether the choice between market order and limit order affects the price impact of trading a large order.

- Aggressive limit order is a limit order whose limit price is one tick better than the best price in the order book.
- Moderate limit order is a limit order which is placed at the best price in the order book.
- Passive limit order is a limit order whose limit price is one tick worse than the best price in the order book.

In these three experiments, the performances of the evolved strategies are compared with those of simple execution strategies which only use one type of order. In the first experiment, the evolved strategy is compared with a simple market order strategy which only adopts market orders. In the second experiment, the evolved strategy is compared with a simple aggressive limit order strategy, a simple moderate limit order strategy and a simple passive limit order strategy which only use aggressive limit orders, moderate limit orders and passive limit orders respectively. In the third experiment, the combined strategy is compared with simple market order strategy and simple limit order strategy which only employs market orders and moderate limit orders respectively.

Performance Evaluation

The standard industry metric for measuring trade execution performance is the *VWAP measure*, short for *Volume Weighted Average Price*. The VWAP price as a quality of execution measurement was first developed by Berkowitz et al. (1988), who argued that 'a market impact measurement system requires a benchmark price that is an un-biased estimate of prices that could be achieved in any relevant trading period by any randomly selected trader' and then defined VWAP as an appropriate benchmark that satisfied this criteria.

The VWAP is calculated as the ratio of the value traded and the volume traded within a specified time horizon (Berkowitz et al., 1988)

$$VWAP = \frac{\sum(Volume * Price)}{\sum(Volume)}$$

where $Volume$ represents the number of shares in each trade and $Price$ represents its corresponding traded price. An example is shown in Table 7.1.

In order to evaluate the performance of a trade execution strategy, its VWAP is compared against the VWAP of the overall market. The rationale here is that performance of a trade execution strategy is considered good if the VWAP of the strategy is more favorable than the VWAP of the market within the trading period and poor if the VWAP of the strategy is less favorable than the VWAP of the market within the trading period. For example, if the VWAP of a buy strategy ($VWAP_{strategy}$) is lower than the market VWAP ($VWAP_{market}$), it is considered a good trade execution strategy. Conversely, if the $VWAP_{strategy}$ is higher than the $VWAP_{market}$, it is considered a poor trade execution strategy. Although this

Table 7.1: VWAP Calculation of A Sample Buy Strategy

	Submission Time	Shares		Traded Price		Value
Child Order 1:	t_0	400	$*$	50.15	$=$	20,060
		600	$*$	50.16	$=$	30,096
Child Order 2:	$t_1(t_0 + \Delta t)$	1,000	$*$	50.40	$=$	50,400
Child Order 3:	$t_2(t_0 + 2\Delta t)$	200	$*$	50.34	$=$	10,068
		800	$*$	50.36	$=$	40,288
Child Order 4:	$t_3(t_0 + 3\Delta t)$	1,000	$*$	50.39	$=$	50,390
Child Order 5:	$t_4(t_0 + 4\Delta t)$	1,000	$*$	50.68	$=$	50,680
Child Order 6:	$t_5(t_0 + 5\Delta t)$	1,000	$*$	51.10	$=$	51,100
Child Order 7:	$t_6(t_0 + 6\Delta t)$	1,000	$*$	50.87	$=$	50,870
Child Order 8:	$t_7(t_0 + 7\Delta t)$	700	$*$	50.98	$=$	35,686
		300	$*$	51.00	$=$	15,300
Child Order 9:	$t_8(t_0 + 8\Delta t)$	1,000	$*$	50.39	$=$	50,390
Child Order 10:	$t_9(t_0 + 9\Delta t)$	1,000	$*$	50.26	$=$	50,260
Total:		10,000				505,588
VWAP $= 505,588/10,000 = 50.5588$						

is a simple metric, it largely filters out the effects of volatility, which composes market impact and price momentum during the trading period (Almgren, 2008).

The performance evaluation functions are as follows (which were used by Lim & Coggins (2005b)):

$$VWAP\ Ratio = \begin{cases} \frac{10^4 * (VWAP_{strategy} - VWAP_{market})}{VWAP_{market}} & (BuyStrategy) \\ \\ \frac{10^4 * (VWAP_{market} - VWAP_{strategy})}{VWAP_{market}} & (SellStrategy) \end{cases}$$

where $VWAP_{market}$ is the average execution price which takes into account all the trades over the day excluding the strategy's trades. This corrects for bias, especially if the order is a large fraction of the daily volume (Lim & Coggins, 2005b). For both buy and sell strategies, the smaller the VWAP Ratio, the better the strategy is. If the combined strategies outperform the simple execution strategies in terms of minimising price impact, it can be concluded that order choice is an important determinant of trade execution strategies.

The design of an execution strategy can be considered of consisting of two stages. The first stage is to divide a big block of shares into multiple small orders, and the second stage is to determine the parameters of each small order, including order type (limit/market order), submission time, limit price of limit order and lifetime (the time length when a limit order appears in the order book before it is cancelled or changed). How to divide a large trade depends on the order size and trading time. The following three sections describe the designs of execution strategies in these three experiments respectively.

7.2.1 Design of Experiment One

In this experiment, the large-quantity order is equally divided into 30 small orders, and submit each small order into the market at regular intervals over one trading day. In the design of the combined execution strategy, limit orders with varied levels of aggressiveness are allowed. A genetic algorithm is adopted as the learning method of the trader who learns where to place his limit order in the order book and when to cross his limit order over the bid-ask spread if it has not been fully executed. The representation of the trading strategy in the genetic algorithm method is shown in Figure 7.1.

The order aggressiveness is distinguished by the variable c_i which represents the relative price of order i. The following examples are used to illustrate how to use the variable c_i to determine the level of aggressiveness of each order. For example, at the time of order submission, the best bid and best ask in the market are 49.00 and 51.00 respectively and then the bid-ask spread is 2.00, assuming

b_1	……	b_{30}	c_1	……	c_{30}

b_1 - b_{30}: lifetimes of 30 orders
c_1 − c_{30}: relative limit prices of 30 orders

Figure 7.1: Representation of Order's Relative Price and Lifetime

that it is a buy strategy:

- $c_i < 2.00$ means that order i will be placed at $49.00 + c_i$, and its lifetime b_i means that at the end of its lifetime any unfilled part (if there is) of order i will get crossed over the bid-ask spread for trading completion;
- $c_i > 2.00$ or $c_i = 2.00$ means that order i is a marketable limit order which will be executed immediately at price 51.00.

7.2.2 Design of Experiment Two

In this experiment, a large-quantity order is equally divided into ten small orders and submitted to the market at regular intervals over the trading day. Three types of limit orders with varied levels of aggressiveness are considered: aggressive limit order, moderate limit order and passive limit order. Two execution strategies with two different settings of limit order's lifetime are devised. One takes a short lifetime, while the other takes a long lifetime. These two lifetime settings are explained as follows. If the limit order is not fully executed at the end of its lifetime, it will get crossed over the bid-ask spread for trading completion.

- In the short-lifetime strategy, each limit order can survive until the end of its interval period.
- In the long-lifetime strategy, each limit order can survive until the end of the trading day.

As prior literature indicates that the state of the limit order book affects order placement strategies, both this experiment and the next experiment take into account this effect. A range of possible explanatory variables, also known as information indicators, have been suggested in the literature. Six variables, which are commonly employed in recent literature of market microstructure, are adopted in the experiments. The explanations of these six variables are shown in Table 7.2.

Table 7.2: Definitions of Market Variables

Variables	**Definitions**
BidDepth	Number of shares at the best bid
AskDepth	Number of shares at the best ask
RelativeDepth	Total number of shares at the best five ask prices divided by total number of shares at the best five bid and ask prices
Spread	Difference between the best bid price and best ask price
Volatility	Standard deviation of the most recent 20 mid-quotes
PriceChange	Number of positive price changes within the past ten minutes divided by the total number of quotes submitted within the past ten minutes

In these experiments, a trader's order choice is determined by a number of rules which are composed of these six information indicators. A grammatical evolution algorithm is used as the learning method of the trader in the experiments in order to learn the rules which are used to determine his order choice. The grammar used in the grammatical evolution algorithm is defined in Figure 7.2.

In this grammar (Figure 7.2), two additional variables $PercOfTradedVolume$ and $PercOfPastTime$ (used in experiment two) are included, which represent the percentage of the traded volume accounting for the total volume V shares and the percentage of the past time accounting for the whole trading period at the observed time respectively. The six indicators, $BidDepth$, $AskDepth$, $RelativeDepth$, $Spread$, $Volatility$, and $PriceChange$ are observed at the time of order submission. In Figure 7.2, $AvgBidDepth$, $AvgVolatility$, and$AvgPriceChange$, represents the average bid depth of the market, $AvgAskDepth$ represents the average ask depth of the market, $AvgRelativeDepth$ represents the average relative depth of the market, $AvgSpread$ represents the average spread of the market, $AvgVolatility$ represents the average volatility of the market and $AvgPriceChange$ represents the average price change of the market.

```
<lc> ::= if (<stamt>)
            class = "AggressiveLimitOrder"
         else {
            if (<stamt>)
                class = "ModerateLimitOrder"
            else
                class = "PassiveLimitOrder"
                }
<stamt> ::= (<stamt><op><stamt>)|<cond1>|<cond2>|
            <cond3>|<cond4>|<cond5>|<cond6>|
            <cond7>|<cond8>
<op> ::= and
<cond1> ::= (BidDepth<comp>AvgBidDepth)
<cond2> ::= (AskDepth<comp>AvgAskDepth)
<cond3> ::= (RelativeDepth<comp>AvgRelativeDepth)
<cond4> ::= (Spread<comp>AvgSpread)
<cond5> ::= (Volatility<comp>AvgVolatility)
<cond6> ::= (PriceChange<comp>AvgPriceChange)
<cond7> ::= (PercOfTradedVolume<comp><threshold>)
<cond8> ::= (PercOfPastTime<comp><threshold>)
<comp> ::= <less>|<more>|<lessE>|<moreE>
<less> ::= "<"
<more> ::= ">"
<lessE> ::= "<="
<moreE> ::= ">="
<threshold> ::= 0.1|0.2|0.3|0.4|0.5|0.6|0.7|0.8|0.9
```

Figure 7.2: Grammars in Experiment Two

An example of an trade execution strategy evolved using the grammar in Figure 7.2 is illustrated in Figure 7.3. Using this example strategy, if $PercOfTradedVolume < 0.5 and PercOfPastTime >= 0.8$ is satisfied, an aggressive limit order is submitted. Otherwise, a moderate limit order is submitted if $Volatility <= AvgVolatility$, or a passive limit order is submitted if $Volatility > AvgVolatility$.

```
if (PercOfTradedVolume<0.5 and PercOfPastTime>=0.8)
    AggressiveLimitOrder
else{
    if (Volatility<=AvgVolatility)
        ModerateLimitOrder
    else
        PassiveLimitOrder
    }
```

Figure 7.3: An Example of Trade Execution Strategy in Experiment Two

7.2.3 Design of Experiment Three

In this experiment, a large-quantity order is equally divided into ten small orders and submitted to the market at regular intervals over the trading day. Both market orders and moderate limit orders are allowed in the design of execution strategy. Any unfilled orders at the end of trading day will get crossed over the bid-ask spread for trading completion. As in experiment two, a grammatical evolution algorithm is used as the learning method for the trader in order to learn the rules on how to choose between market orders and moderate limit orders based on the six information indicators of the limit order book. The grammar adopted in the grammatical evolution algorithm is shown in Figure 7.4.

An example of an trade execution strategy evolved using the grammar in Figure 7.4 is illustrated in Figure 7.5. Using this example strategy, if $(RelativeDepth > AvgRelativeDepth)$ is true, a market order is submitted, otherwise, a moderate limit order is submitted.

```
<lc> ::= if (<stamt>)
                class = "MarketOrder"
         else
                class = "ModerateLimitOrder"
<stamt> ::= <cond1><op><cond2><op><cond3><op>
            <cond4><op><cond5><op><cond6>
<op> ::= and | or
<cond1> ::= (BidDepth>AvgBidDepth) is <boolean>
<cond2> ::= (AskDepth>AvgAskDepth) is <boolean>
<cond3> ::= (RelativeDepth>AvgRelativeDepth) is
            <boolean>
<cond4> ::= (Spread>AvgSpread) is <boolean>
<cond5> ::= (Volatility>AvgVolatility) is <boolean>
<cond6> ::= (PriceChange>AvgPriceChange) is <boolean>
<boolean> ::= True | False
```

Figure 7.4: Grammars in Experiment Three

```
if (RelativeDepth>AvgRelativeDepth) is True
    MarketOrder
else
    ModerateLimitOrder
```

Figure 7.5: An Example of Trade Execution Strategy in Experiment Three

7.3 Artificial Market Simulators

The execution strategies are tested in an artificial market simulator. Critically, in the experiments the actions of the evolved execution strategies employed by the simulated trader impacts on the state of the order book facing all the other agents in the stock market and therefore impacts on their actions. In turn, the actions of those agents impact on the order book facing the trader and therefore on the utility of her execution strategy. In other words, the training and testing environment is *dynamic* and allows examination of the issues of market impact and opportunity cost. on trade execution. The use of an artificial stock market environment to test the utility of execution strategies is a novel contribution of this chapter, and opens up the door to a wide range of future work in this domain. Different market simulators are used to test those three experiments. The purpose of this is to examine whether the execution strategies keep robust in different market environments. Specifically, simulator I is adopted to test the strategies in experiment one, while simulator II is employed to test the strategies in experiments two and three.

7.3.1 Simulator I

In the artificial market simulator I, the agents are equally likely to generate a buy order or a sell order. The order can be a market order, a limit order, or a cancellation of a previous limit order. When an agent is active, she can try to issue a cancelation order with probability λ_c (oldest orders are canceled first), a market order with probability λ_m, a limit order with probability $\lambda_l = 1 - \lambda_c - \lambda_m$. In the model, the limit order price is uniformly distributed inside the spread with probability λ_{in}, and power-law distributed outside the spread with probability $1 - \lambda_{in}$. Limit order price ranges are illustrated in Figure 7.6.

In this model, order generation is modeled as a Poisson process, which means

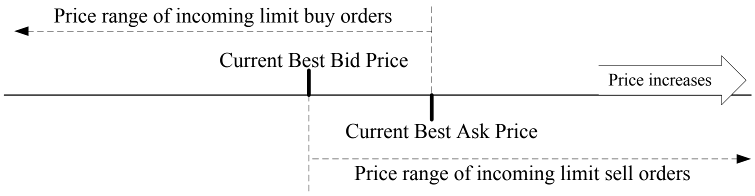


Figure 7.6: Place Limit Price

that the time between orders follows an exponential distribution. As each incoming buy (sell) market order arrives, the market agent will match it with the best ask (bid) limit order stored in the order book. If this market order is fully filled by the first limit order, the unfilled part will be matched to the next best ask (bid) limit order until it is fully filled. As each incoming limit order arrives, the market agent will store it in the order book according to price and time priority. As each incoming cancelation order arrives, the market agent will delete the relevant limit order in the order book. In the simulation, a Swarm platform in JAVA (Swarm, 2009) is adopted, which is one of the most popular agent-based modeling platforms. The algorithm used in the simulation I is described in Algorithm 3.

In order to ensure that the order flows generated by the artificial market are economically plausible, all the parameters in the model are derived from previous literature (Chakraborti et al., 2011; Farmer et al., 2005b,c,a, 2006; Mike & Farmer, 2008; Toth et al., 2009). The parameters used in the simulation are presented in Table 7.3.

Algorithm 3 Artificial Market Simulator I

Generate t from an exponential distribution with parameter τ;
$current_time = current_time + t$.
Generate P_{sign} from Bernoulli distribution, and determine the agent type with probability 0.5.
switch *Agent Type* **do**
 case *Buyer*
  Generate P_{type} from a uniform distribution, and determine the order type according to the probabilities λ_m, λ_l and λ_c.
  switch *Order Type* **do**
   case *Market order*
    Submit a market order.
   endsw
   case *Limit order*
    Generate P_{lo} from a uniform distribution, and determine the limit order type with probabilities $P_{inSpread}$ and $P_{offSpread}$.
    switch *Limit Order Type* **do**
     case *Spread limit order*
      Submit a limit order whose limit price is drawn from a uniform distribution of values between the best bid and best ask on the order book.
     endsw
     case *Off-spread limit order*
      Generate a random value denoted as RP_{offspr} from a power-law distribution with exponent α.
      Place a limit order at the price $(BestBid - RP_{offspr})$ on the order book.
     endsw
    endsw
   endsw
   case *Cancelation order*
    Cancel the oldest outstanding limit order.
   endsw
  endsw
 endsw
 case *Seller*
  (similar to the case of buyer)
 endsw
endsw

Table 7.3: Initial Parameters for Order Book based ASM I

Explanation	**Value**
Initial Price	$price^0 = 100$
Tick Price	$\delta = 0.01$
Probability of Cancellation Order	$\lambda_c = 0.07$
Probability of Market Order	$\lambda_m = 0.33$
Probability of Limit Order	$\lambda_l = 0.60$
Probability of Limit Order in Spread	$\lambda_{in} = 0.35$
Probability of Limit Order Out of Spread	$\lambda_{out} = 0.65$
Limit Price Tail Index	$1 + \alpha = 1.3$
Order Size	$(\mu, \sigma) \sim (4.5, 0.8) * 100$ shares
Waiting Time	$\tau = 6, 90$

7.3.2 Simulator II

As in simulator I, an agent in simulator II is equally likely to generate a buy order or a sell order at each time step, and this order can be a market order, or a limit order, or a cancellation of a previous limit order, with probabilities λ_m, λ_l, and λ_c respectively. The sum of these probabilities is one ($\lambda_m + \lambda_l + \lambda_c = 1$). In simulator II, for a limit buy (sell) order, it has a probability of $\lambda_{inSpread}$ falling inside the bid-ask spread, a probability of λ_{atBest} falling at the best bid (ask) price, and a probability of λ_{inBook} falling off the best bid (ask) price in the book, ($\lambda_{inSpread} + \lambda_{atBest} + \lambda_{inBook} = 1$). The limit price inside the spread follows a uniform distribution. The limit price off the best bid (ask) price follows a power law distribution with the exponent of $(1 + \mu_1)$. The log order size of a market order follows a power law distribution with the exponent of $(1 + \mu_2)$, while the log order size of a limit order follows a power law distribution with the exponent of $(1 + \mu_3)$. The algorithm used in the simulation II is described in Algorithm 4.

In order to ensure that the order flows generated by the artificial market are plausible, all the parameters in the model are derived from previous literature (Chakraborti et al., 2011; Farmer et al., 2005a,b; Mike & Farmer, 2008; Toth et al., 2009). The parameters used in the simulator II are presented in Table 7.4.

Algorithm 4 Artificial Market Simulator II

Generate a random number from a uniform distribution $(0, 1)$ and determine the agent type according to the probability 0.5.
switch *Agent* **do**
- **case** *Buyer*
  - Generate a random number from a uniform distribution $(0, 1)$ and determine the order type according to the probabilities λ_m, λ_l and λ_c.
  - **switch** *Order Type* **do**
    - **case** *Market order*
      - Execute a market order with the order size generated from a power-law distribution with exponent μ_2
    - **endsw**
    - **case** *Limit order*
      - Generate a random number from a uniform distribution $(0, 1)$ and determine the limit order type according to probabilities $\lambda_{inSpread}, \lambda_{atBest}, and \lambda_{inBook}$.
      - **switch** *Limit Order Type* **do**
        - **case** *Inside-spread Limit Order*
          - Generate a random value denoted as P_{inspr} from a uniform distribution $(BestBid, BestAsk)$.
          - Place a limit order at the price P_{inspr} on the order book with the order size generated from a power-law distribution with exponent μ_3.
        - **endsw**
        - **case** *Spread Limit Order*
          - Place a limit order at the best bid price on the order book with the order size generated from a power-law distribution with exponent μ_3.
        - **endsw**
        - **case** *Off-spread Limit Order*
          - Generate a random value denoted as RP_{offspr} from a power-law distribution with exponent μ_1.
          - Place a limit order at the price $(BestBid - RP_{offspr})$ on the order book with the order size generated from a power-law distribution with exponent μ_3.
        - **endsw**
      - **endsw**
    - **endsw**
    - **case** *Cancelation order*
      - Cancels the oldest outstanding limit order that the buyer previously submitted.
    - **endsw**
  - **endsw**
- **endsw**
- **case** *Seller*
  - (similar to the case of buyer)
- **endsw**

endsw

Table 7.4: Initial Parameters for Artificial Limit Order Market II

Explanation	**Value**
Initial Price	$price^0 = 50$
Tick Price	$\delta = 0.01$
Probability of Order Cancellation	$\lambda_c = 0.34$
Probability of Market Order	$\lambda_m = 0.16$
Probability of Limit Order	$\lambda_l = 0.50$
Probability of Limit Order in Spread	$\lambda_{inSpread} = 0.32$
Probability of Limit Order at Best Quote	$\lambda_{atBest} = 0.33$
Probability of Limit Order off the Best Quote	$\lambda_{inBook} = 0.35$
Limit Price Power Law Exponent	$1 + \mu_1 = 2.5$
Market Order Size Power Law Exponent	$1 + \mu_2 = 2.7$
Limit Order Size Power Law Exponent	$1 + \mu_3 = 2.1$

7.4 Results and Discussions

Results of the three experiments are presented and discussed as below.

7.4.1 Results of Experiment One

The Combined-I strategy was evolved using the parameters presented in Table 7.5, and then was evaluated in the artificial market for 30 simulation runs. Its VWAP ratio is shown in Table 7.6. The VWAP Ratio reveals the difference between the volume weighted execution price of Combined-I orders and the average traded price of all orders during the whole simulation time. The better strategies have smaller VWAP ratios.

Table 7.5: Parameters for Genetic Algorithm in Experiment One

Parameter	**Value**
Population size	30
Maximum number of generation	100
Generation gap	0.8
Crossover rate	0.75
Mutation rate	0.05
Selection method	Stochastic Universal Sampling
Crossover method	Single-Point

The VWAP ratio of simple market order strategy (SM), which was evaluated in the artificial market for 30 simulation runs, is also presented in Table 7.6. In order to analyze the Combined-I strategy, the execution types of the 30 orders are also recorded in the experiment. The three execution types are explained as follows.

1. MLO. It represents marketable limit order. If the order is placed at or beyond the best price on the opposite side of the order book, and gets executed immediately at its limit price, it is denoted as MLO.

2. LO. It represents limit order. If the order is executed against market orders with opposite order sign some time after its submission, it is denoted as LO.

3. Crossed. If the order is not executed in the required time period but gets executed at the end of the period for trading completeness, it is denoted as Crossed.

Table 7.6: Results of Experiment One.

	SM Strategy	**Combined-I Strategy**			
	VWAP Ratio	VWAP Ratio	TradedOrderType		
	(10^{-3})	(10^{-3})	MLO	LO	Crossed
Buy Order	5.9748	0.5146	11	8	11
Sell Order	3.2378	-1.8244	13	7	10

Note: SM Strategy means Simple Market Order Strategy.
MLO is short for Marketable Limit Order.
LO is short for Limit Order.
Crossed represents Crossed Limit Order.

From Table 7.6, the Combined-I strategy outperforms the SM strategy significantly, both in-sample and out-of-sample, which is consistent with the results in Lim & Coggins (2005b). Table 7.6 shows that the Combined-I strategy, which has more orders executed via LOs, has a smaller VWAP ratio, meaning better performance. All the Combined-I strategies with negative VWAP ratios have more orders executed in the way of LO than those executed in the two other ways, except the best out-of-sample strategy in Table 7.6. Also, Combined-I strategies have achieved better VWAP than that of the whole simulation time for buy and sell in in-sample test, which is showed by the negative values of VWAP ratios. This is more significant for the sell order.

7.4.2 Results of Experiment Two

The Combined-II strategy was evolved using the parameters presented in Table 7.7, and then was evaluated in the artificial market for 240 simulation runs. Its VWAP ratio is shown in Tables 7.8 & 7.9. The VWAP Ratio reveals the difference between the volume weighted execution price of Combined-II orders and the average traded price of all orders during the whole simulation time. The better strategies have smaller VWAP ratios.

Table 7.7: Parameters for Grammatical Evolution in Experiment Two

Parameter	**Value**
Population size	100
Maximum number of generation	40
Generation gap	0.8
Crossover rate	0.9
Mutation rate	0.01
Selection method	Roulette
Crossover method	Single-Point

The results (all out of sample) of buy strategies and sell strategies are provided in Tables 7.8 & 7.9. The "S-T" represents short-term lifetime and the "L-T" represents long-term lifetime. The "Mean" is the average VWAP ratio of each strategy over the 240 days, and "S.D." represents the standard deviation of the average (daily) VWAP ratio. P-values for the null hypothesis $H_1 : mean_{SA} \leq mean_{GE}$, $H_2 : mean_{SM} \leq mean_{GE}$, $H_3 : mean_{SP} \leq mean_{GE}$ are also shown in the table, to indicate the degree of statistical significance of the performance improvement of Combined-II strategies over the two simple strategies. The figures show that the null hypotheses are rejected at the ≤ 0.01 level.

Based on the results, Combined-II strategies notably outperform the three benchmark strategies, a simple aggressive limit order strategy (SA), a simple modest limit order strategy (SM) and a simple passive limit order strategy (SP). The performances of these four strategies can be described as follows.

$$\begin{aligned} &Performance_{Combined-II} > Performance_{SP} \\ &\quad > Performance_{SA} > Performance_{SM} \end{aligned}$$

Combined-II strategies perform the best, while SM strategies perform the worst. The negative VWAP ratios show that the Combined-II strategies achieve better execution prices than the average execution price of the market. The values of standard deviation show that the performance of Combined-II strategies is more stable than most of the other three kinds of strategies over the tested trading days, and the performance of SM strategies is more volatile than the other three kinds of strategies. Comparing the performance of the strategies for buy and sell orders, it can be observed that the performances of sell strategies are better than those of buy strategies in most cases, which means that trading costs of buys are higher than those of sells. This asymmetry is consistent with previous empirical findings, for example, see Kraus & Stoll (1972), Keim & Madhavan (1997), Bikker et al. (2007). And it is observed that L-T strategies all perform better than S-T strategies, which indicate that strategies with longer lifetime can achieve better execution prices than those with short lifetime.

Table 7.8: Results of Experiment Two (Buy Orders)

	SA Strategy		**SM Strategy**		**SP Strategy**		**Combined-II Strategy**				
	Mean	(S.D.)	Mean	(S.D.)	Mean	(S.D.)	Mean	(S.D.)	H_1	H_2	H_3
S-T	14.7	(1.74%)	66.25	(2.53%)	13.19	(1.76%)	-2.35	(1.52%)	0.01	0.00	0.01
L-T	5.14	(1.69%)	60.7	(2.03%)	9.37	(1.48%)	-5.86	(1.3%)	0.01	0.00	0.00

Note: SA Strategy represents Simple Aggressive limit order Strategy.
SM Strategy represents Simple Moderate limit order Strategy.
SP Strategy represents Simple Passive limit order Strategy.
S.D. is short for Standard Deviation.
S-T is short for Short-Term lifetime.
L-T is short for Long-Term lifetime.

Table 7.9: Results of Experiment Two (Sell Orders)

	SA Strategy		**SM Strategy**		**SP Strategy**		**Combined-II Strategy**				
	Mean	(S.D.)	Mean	(S.D.)	Mean	(S.D.)	Mean	(S.D.)	H_1	H_2	H_3
S-T	24.7	(1.86%)	43.6	(2.18%)	6.78	(2.03%)	-8.15	(1.87%)	0.00	0.00	0.01
L-T	5.28	(1.56%)	37.23	(2.27%)	4.91	(1.39%)	-12.07	(1.53%)	0.00	0.00	0.01

Note: SA Strategy represents Simple Aggressive limit order Strategy.
SM Strategy represents Simple Moderate limit order Strategy.
SP Strategy represents Simple Passive limit order Strategy.
S.D. is short for Standard Deviation.
S-T is short for Short-Term lifetime.
L-T is short for Long-Term lifetime.

7.4.3 Results of Experiment Three

The Combined-III strategy was evolved using the parameters presented in Table 7.10, and then was evaluated in the artificial market for 240 simulation runs. Its VWAP ratio is shown in Table 7.11. The VWAP Ratio reveals the difference between the volume weighted execution price of Combined-III orders and the average traded price of all orders during the whole simulation time. The better strategies have smaller VWAP ratios.

Table 7.10: Parameters for Grammatical Evolution in Experiment Three

Parameter	**Value**
Population size	100
Maximum number of generations	40
Generation gap	0.8
Crossover rate	0.9
Mutation rate	0.01
Selection method	Roulette
Crossover method	Single-Point

The results (all out of sample) of buy strategies and sell strategies are provided in Table 7.11. The "Mean" is the average VWAP ratio of each strategy over the 240 days, and "S.D." represents the standard deviation of the average (daily) VWAP ratio. P-values for the null hypothesis $H_1 : mean_{SM} \leq mean_{Combined-III}$ and $H_2 : mean_{SL} \leq mean_{Combined-III}$ are also shown in the table, to indicate the degree of statistical significance of the performance improvement of Combined-III strategies over the two simple strategies. The figures show that the null hypotheses are rejected at the ≤ 0.01 level.

Based on the results, Combined-III strategies notably outperform the two benchmark strategies, a simple market order strategy (SM) and a simple limit order strategy (SL). The negative VWAP ratios of -1.42 and -23.21 show that the GE evolved strategies achieve better execution prices than the average execution price of the

Table 7.11: Results of Experiment Three

	SM Strategy		SL Strategy		Combined-III Strategy			
	Mean	(S.D.)	Mean	(S.D.)	Mean	(S.D.)	H_1	H_2
Buy Order	69.64	(0.42%)	42.54	(1.45%)	-1.42	(0.49%)	0.00	0.01
Sell Order	68.73	(0.36%)	13.81	(1.59%)	-23.21	(0.48%)	0.00	0.01

Note: SM Strategy represents Simple Market order Strategy.
SL Strategy represents Simple Limit order Strategy.
S.D. is short for Standard Deviation.

market. The small standard deviations of 0.49 and 0.48 indicate that Combined-III strategy is robust over the tested trading days. The performance of SL strategies seems more volatile than those of SM strategies and Combined-III strategies, and the performance of SM strategies is more stable than those of the other two kinds of strategies. Comparing the performance of the strategies for buy and sell orders, it can be observed that the performances of sell strategies are better than those of buy strategies, which means that trading costs of buys are higher than those of sells. This asymmetry is consistent with previous empirical findings.

7.4.4 Discussions

The results in the three experiments all show that the combined strategies outperform the simple strategies. This indicates that order choice affects the performance of trade execution strategies, that is the price impact of trading large-orders, and suggests that order choice is an important decision in large-order trade execution.

In the above experiments, both buy and sell strategies are tested. The results of these experiments all show that sell strategies generally perform better than buy strategies, confirming the asymmetric price impacts of institutional trades found in previous literature (Bikker et al., 2007; Chan & Lakonishok, 1993, 1995;

Gemmill, 1996; Holthausen et al., 1987, 1990; Keim & Madhavan, 1995b, 1996, 1997; Kraus & Stoll, 1972).

Several explanations appear in the literature to account for this asymmetry phenomenon. Chan & Lakonishok (1993) and Keim & Madhavan (1996) argue that sells are more often liquidity-motivated rather than information-based than buys. Buys create new long-term positions and thus imply a preference to hold a particular stock. Saar (2001) provides a different explanation for the price impact asymmetry. He builds a theoretical model which demonstrates how the price impact asymmetry can arise. The main implication of the model is that the history of price performance of a stock affects the degree of asymmetry: the longer the run of price appreciations, the less positive the difference in permanent price impact between buys and sells. When the price run-up is long enough, sells may even have higher price impact than buys. Another explanation for the price impact asymmetry is given by Chiyachantana et al. (2004). They find that the asymmetry depends on particular market conditions: price effects of buyer-initiated trades are greater in bull markets (as in 1997-1998) whereas those of seller-initiated trades are larger in bear markets (as in 2001). Boscuk & Lasfer (2005) take a different view and show that the type of investor and the combination of the size of the trade and the trader's resulting level of ownership are the major determinants of price impact asymmetry at the London Stock Exchange.

7.5 Summary

Order choice problem and large-order trading problem have been studied separately in the market microstructure literature. However, large order trading with various order types has not been investigated. This chapter sheds light on large-order traders' use of limit orders. It examined whether order choice affects the price impact of trading a large order. Different combinations of order types were investigated. Agent-based artificial markets were simulated in order to evaluate the trading strategies. One advantage of this performance evaluation approach is that the simulated market can capture the dynamic effect of trading while historical data can not.

The results show that the combined-order-type strategies perform better than the simple-order-type strategies in terms of price impact. This confirms the finding in Keim & Madhavan (1997) that the investment style affects trading costs. This study also demonstrates that buys are more expensive than sells, which is consistent with the findings in Bikker et al. (2007); Chan & Lakonishok (1993, 1995); Gemmill (1996); Holthausen et al. (1987, 1990); Keim & Madhavan (1995b, 1996, 1997); Kraus & Stoll (1972).

This chapter compliments the prior literature on trade execution strategies, providing evidence that order choice plays an important role in determining the price impact when trading a large order. The findings in this chapter are of interest to academics. When modelling the behaviours of large-order traders, order choice as an import factor should be taken into account. Moreover, the results provide useful insights for institutional traders who are obligated to liquidate or acquire a big position, suggesting that large-order execution strategies adopting multiple order types can perform better than those using single order type.

In addition, this chapter contributes to the literature on agent-based modellings of financial markets. It makes the first attempt to use agent-based modelling to evaluate the performance of trade execution strategies. The artificial market model developed in this study can be extended to examine the effect of high-frequency trading on market quality.

Part III

Conclusions

Chapter 8

Conclusions and Future Work

This chapter begins with a brief summary of the thesis and the contributions it has made. The limitations of this thesis are highlighted in Section 8.2. Suggestions for future work complete the chapter in Section 8.3.

8.1 Thesis Summary

This thesis empirically investigated the effect of trading on price, that is price impact. Understanding price impact is a crucial task in financial markets. In order to get a better understanding of this effect, the following questions were asked:

1. *Does price impact on the LSE and the NYSE exhibit an intraday pattern?*

2. *Does agent intelligence affect the magnitude of price impact?*

3. *Does order choice affect the price impact of trading large orders?*

Empirical analysis on intraday behaviours of market liquidity and price impact were carried out in Chapter 5. The analysed data was drawn from the NYSE-Euronext TAQ database and the LSE ROB database. Before using the data for analysis, several steps were implemented to preprocess the data in Chapter 4. The entries in the TAQ data recorded during non-normal trading periods were deleted, as well as the entries with non-positive values. Identified outliers in the TAQ data were also removed. An algorithm was developed for inferring the hidden information contained in the ROB data. Then the limit order book for each trading day was rebuilt using the ROB data. With the clean data, six stocks selected from the US markets and six stocks selected from the UK markets were analysed. The results on intraday patterns generally confirm those found in the previous studies as reviewed in Chapter 2. In particular, intraday price impact is found to be reverse S-shaped for US and UK stocks.

The study in Chapter 6 addressed the second research question. As motivated by the argument in Farmer et al. (2004) that traders condition their order sizes based on market liquidity, this study investigated whether agent intelligence affects the magnitude of price impact. An agent-based modelling approach was adopted, as this method facilitates the isolation of a relevant factor in the market simulation. An artificial limit order market composed of zero-intelligence agents was simulated. It captured the salient features of electronic limit order markets including continuous trading, a visible book of orders, price-time priority rules, instantaneous trade reporting rules, order cancellation capabilities, and both limit order and market order functionality. In addition, the artificial market was calibrated using the LSE ROB data. The results showed that the price impact observed in the artificial market was generally larger than that observed in the real market, indicating that agent intelligence is an important determinant of the magnitude of price impact.

Price impact is an significant component of trading cost, and thus controlling price impact is important for improving execution performance. The study in Chapter 7 examined whether order choice affects the performance of large-order execution. Several combined strategies, which used various types of orders, were developed and their performances was then evaluated using price impact. The results were benchmarked against simple strategies which adopted a single order type. The results show that the combined strategies outperformed the simple strategies, indicating that order choice is important in controlling price impact for large-order execution.

Arising from this thesis, these series of experiments help our understanding of price impact. It identifies the important intraday behaviours of price impact, the crucial role of agent intelligence in determining the magnitude of price impact, and the significance of order choice in determining the price impact of trading large orders. These findings have further implications for research on market microstructure, for example, understanding how securities markets function and developing comprehensive models of market microstructure. Moreover, they are also of interest to institutional traders and trading desks.

8.1.1 Contributions

Through the gathering together of the literature on market microstructure and the literature on agent-based financial markets and the subsequent implementation of a series of experiments in the domain, a number of key insights are generated. This section reviews the contributions of this thesis as follows.

- **Up-to-date review of relevant literature**

  Chapters 2 & 3 provided an up-to-date survey of relevant literature on market microstructure, price impact, and agent-based financial markets. Chapter 2 initially gave an overview of the literature of market microstructure, and then specifically reviewed several streams of market microstructure research relevant to this thesis, including studies on the relationship between price impact and trade size, studies on large price impact, studies of intraday phenomena on price volatility, bid-ask spread, trading volume and market depth, studies on optimal trading strategies for a single asset, and studies on order submission strategies. This chapter uncovered several research gaps, which led to the experimental studies of this thesis. Chapter 3 reviewed various applications of agent-based modelling to financial markets, and revealed that there had been little attempt to use agent-based modelling for investigating price impact.

- **Analysis of intraday behaviours of price impact in UK and US markets**

  As some studies (Ahn & Cheung, 1999; Lee et al., 1993; Li et al., 2005; Vo, 2007) have documented an intraday pattern on market depth, price impact may exhibit an intraday pattern because price impact is affected by market liquidity. This thesis analysed intraday behaviours of price impact, using the recent data drawn from the NYSE-Euronext TAQ database and the LSE ROB database. It for the first time documented a reverse S-shaped intraday pattern on price impact in UK and US markets. Moreover, it found that price impact had a negative association with near market depth. The result suggests that the information conveyed by price impact is consistent with near market depth as indicators of intraday market liquidity.

- **Up-to-date evidence on intraday behaviours of market liquidity in UK and US markets**

  The previous studies on intraday phenomena are dated. This thesis analysed intraday behaviours of market liquidity, using recent data drawn from the NYSE-Euronext TAQ database and the LSE ROB database, and provided up-to-date evidence on intraday behaviours of price volatility, bid-ask spread, market depth, trading volume, trading frequency and trade size in

UK and US markets. Unlike previous studies on intraday phenomena which only examined the market depth at the best level of the limit order book, this thesis for the first time analyses the intraday behaviour of market depth for up to ten levels of the limit order book in the UK market.

- **Investigation of the role of agent intelligence in determining the magnitude of price impact**

  Prior studies on price impact typically found that price impact is a concave function of trade size. Some of them (e.g. Farmer et al. (2004); Hopman (2007); Weber & Rosenow (2006)) argued that this concavity was due to selective liquidity taking, namely agents conditioning order size on market liquidity. However, little study was conducted to investigate this. This thesis for the first time investigated whether agent intelligence plays an important role in determining the magnitude of price impact. The results support the hypothesis that agent intelligence plays an important role in determining the magnitude of price impact.

- **Examination of the effect of order choice on the price impact when trading a large order**

  Controlling price impact is an important issue in financial markets. Many studies have been devoted to examining this problem. However, these studies only adopted market orders, and no studies were conducted to investigate the effect of order choice on the price impact of trading a large order. This thesis for the first time analysed this. The results demonstrate that order choice plays an important role in determining the price impact of trading a large order.

- **Use of ABM for investigating price impact**

Agent-based modelling have been applied to economics and finance for a long time, but there has been little attempt to use ABM for investigating price impact. This thesis makes the first attempt to investigate price impact using ABM in its two experimental studies. It laid the framework for this use, and the model developed here can be extended for other applications, like investigation of the market effect of high-frequency trading.

8.2 Limitations of this Thesis

The primary focus of this thesis is on the investigation of price impact, and it makes useful contributions to our understanding of price impact. However, one thesis can not cover all aspects of this topic. There are a number of areas worthy of further research that fall outside the scope of this work.

Price impact can be caused by market orders, limit orders and upstairs orders. When examining the intraday behaviours of price impact, the focus of the present work is on the price impact caused by market orders. A few studies have investigated the price impact of other trading activities. For example, Eisler et al. (2012); Hautsch & Huang (2012) examine the price impact caused by limit orders, and Gabaix et al. (2006); Keim & Madhavan (1996) investigate the price impact of upstairs orders.

A number of studies have shown that price impact decays over time. In the current work, concentration is placed on the immediate effect of trading on prices. Some studies have been devoted to modelling the resilience of price impact. For example, Gatheral et al. (2011) model that price impact decays exponentially over time.

Large-order execution can occur at several trading venues. In this thesis, attention is limited to trading at one trading venue. Many traders are willing to route their orders to the cheapest trading place in order to reduce trading costs. A number of studies are focusing on this issue. For example, Foucault & Menkveld (2008) empirically examines smart order routing systems, and Rawal (2010) discusses the importance of intelligent decision-making in order routings.

8.3 Future Work

Building upon this research, there are numerous avenues for future work. A few of them are outlined as follows.

The study in Chapter 5 shows that market volume has a U-shaped intraday pattern and near market depth has an S-shaped intraday pattern. It would be interesting to determine who contributes to trading volume and market depths, that is the constituents of the market participants at different times of the trading day. Understanding whether institutional investors or individual investors dominate the market when trading volume is high or when depth is high would help us better understand the underlying drivers of the intraday patterns.

The price impact studied in Chapter 5 is limited to the price effect of market orders. The price impact of limit orders is not explored here. A recent study by Hautsch & Huang (2012) quantifies the price effect of submitting a limit order in a limit order market and find that limit orders have significant impact. However, the intraday behaviour of this impact is unknown. It would be interesting to analyse the intraday patterns of price impact of limit orders using the method in Hautsch & Huang (2012).

The study in Chapter 7 can be extended to an investigation of large-order trading in multiple markets. A number of artificial markets can be simultaneously simulated with an agent-based modelling approach. The choice of trading venue will be decided on for each order at the same time with other decisions like choice of order type and order size. This investigation would shed light on how liquidity is supplied and demanded across multiple linked markets.

Appendix A

Explanation of the Chosen Time Period in Chapter 5

One of the aims of this thesis is to investigate whether price impact exhibits an intraday pattern and whether this intraday pattern is consistent across different stocks in US and UK markets. Therefore, the stocks are chosen from different industries and with different market capitalisations. The chosen time period is a normal period. During these time periods, there were no disturbances in the markets. The chosen length of the time period is enough to generate intraday patterns. Figures A.1 & A.2 illustrate the intraday curves of price volatility using one-day data and one-month day separately.

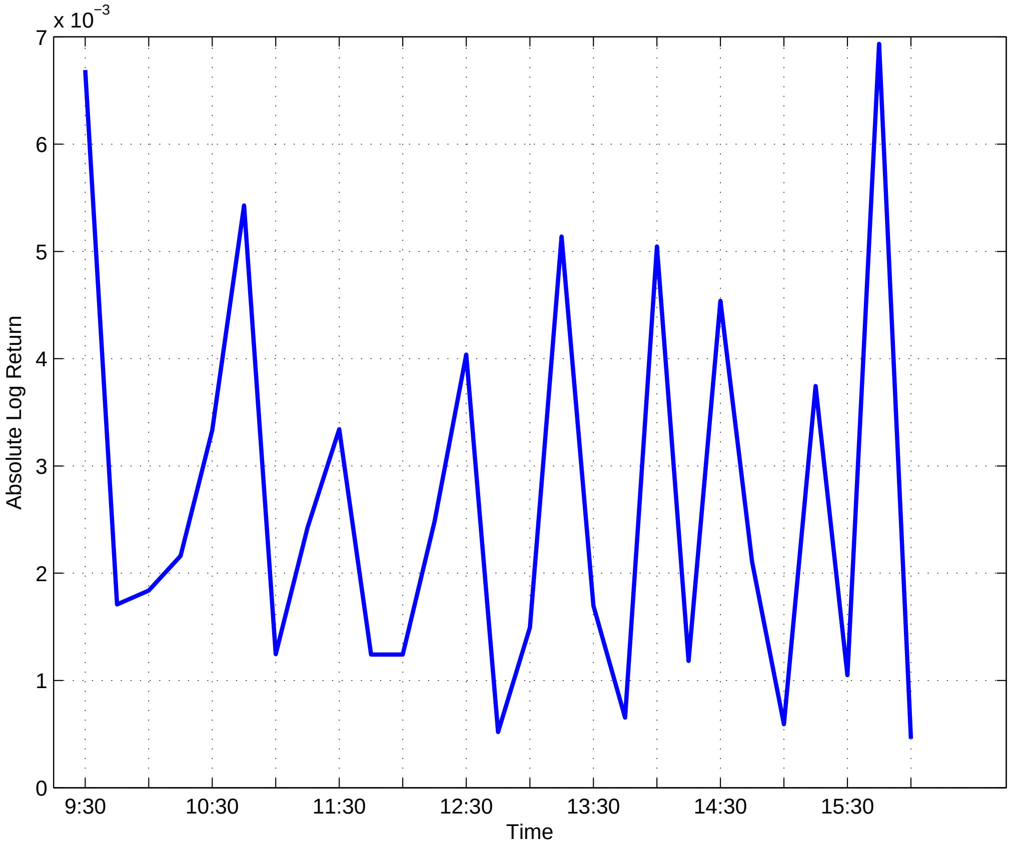


Figure A.1: Intraday Price Volatility (One-day data)

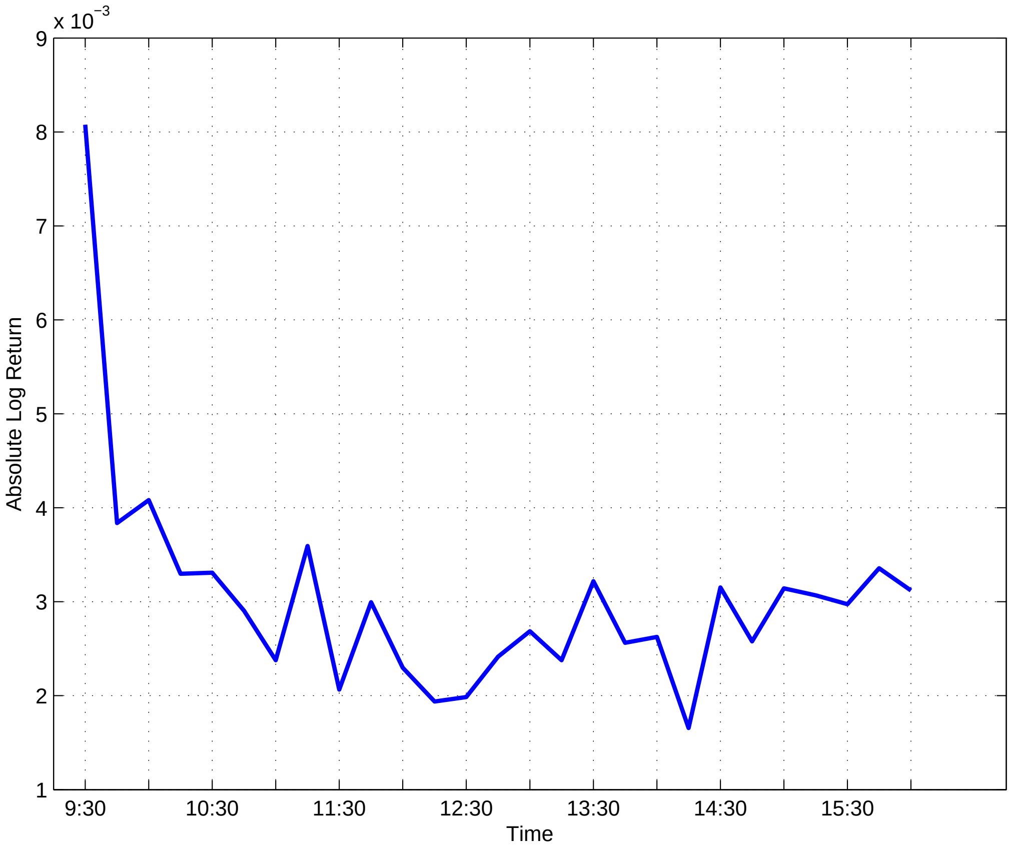
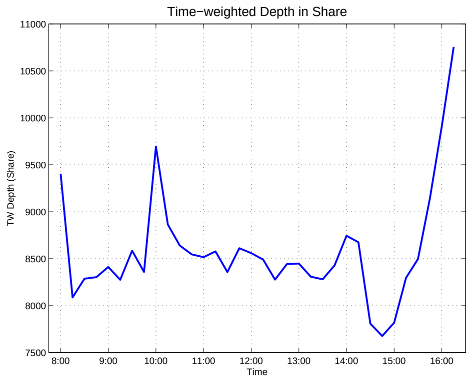


Figure A.2: Intraday Price Volatility (One-month data)

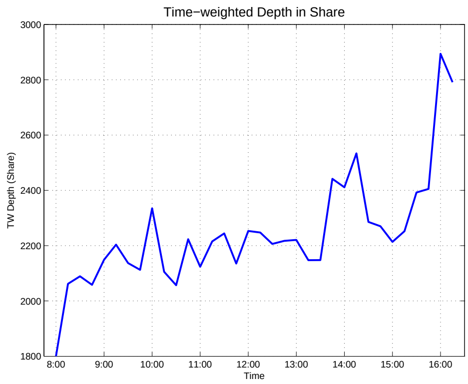
Appendix B

More Results of Chapter 5

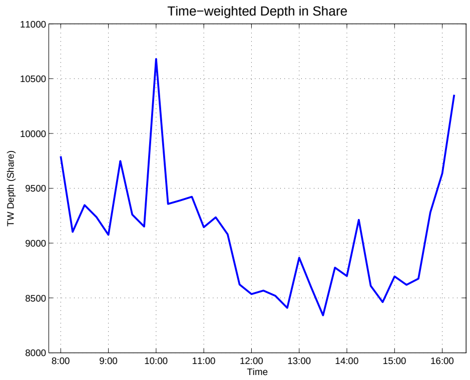
This appendix includes some results from experiment two in Chapter 5. This thesis measures market depth in share and in value as well as time-weighted market depth for UK stocks. Chapter 5 includes the figures showing intraday pattern on market depth in share. This appendix presents the intraday patterns on market depth in value and time-weighted market depth. From Figures B.1-B.10, it can be observed that the intraday patterns on market depth in value and on time-weighted market depths are almost identical to the intraday patterns on market depth in share as shown in Chapter 5.



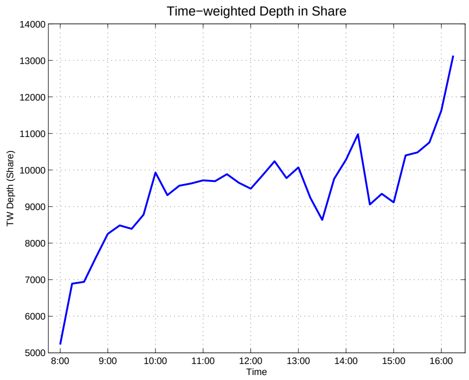
(a) BARC



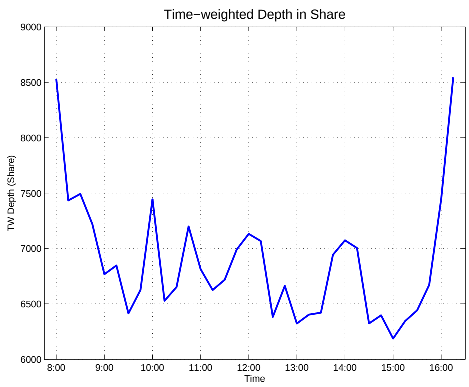
(b) BATS



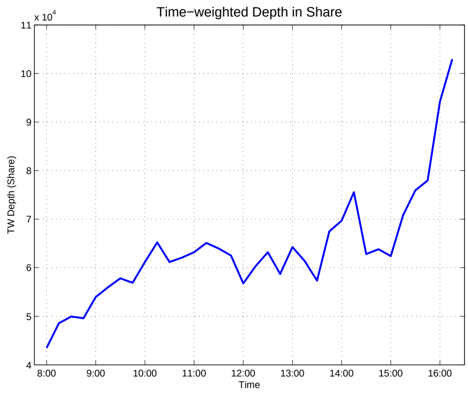
(c) BP



(d) GSK

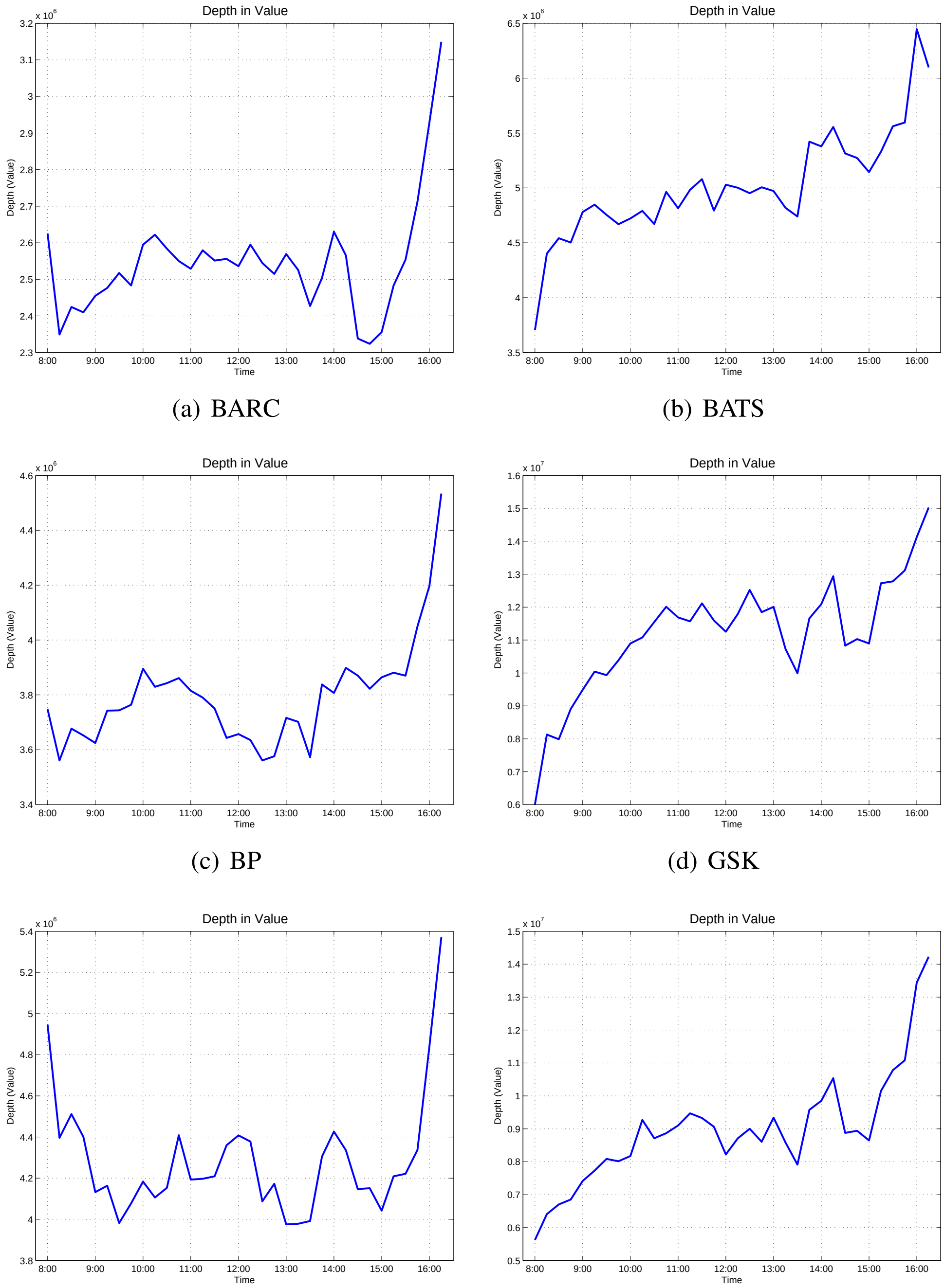


(e) HSBC



(f) VOD

Figure B.1: Intraday Time-weighted Depth in Share (UK stocks)

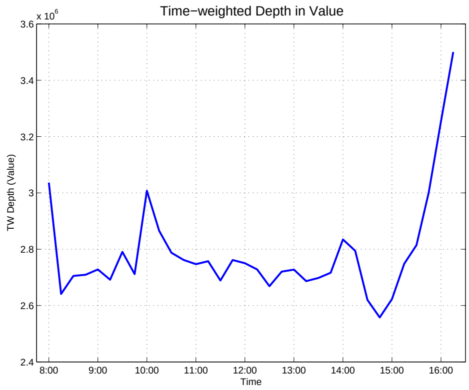


(a) BARC (b) BATS

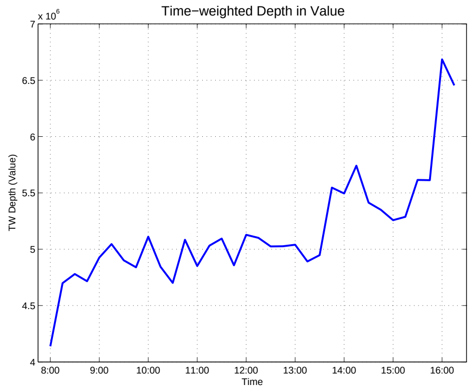
(c) BP (d) GSK

(e) HSBC (f) VOD

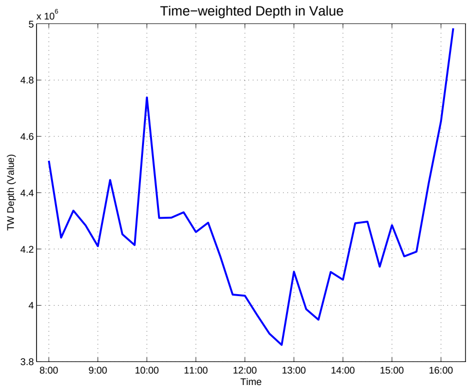
Figure B.2: Intraday Depth in Value (UK stocks)



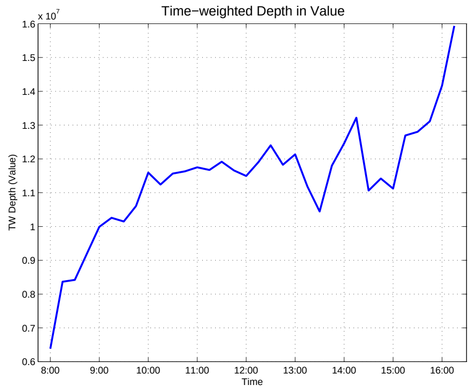
(a) BARC



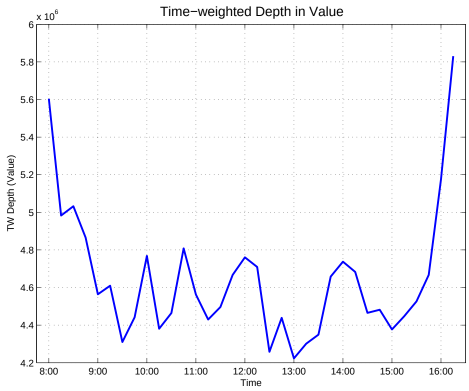
(b) BATS



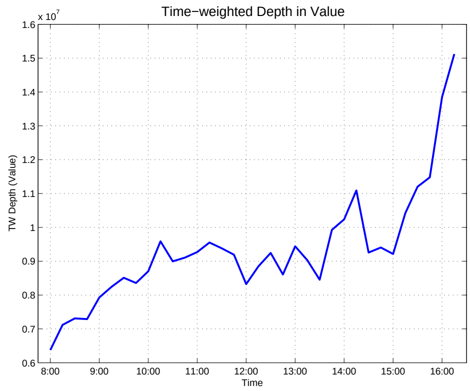
(c) BP



(d) GSK

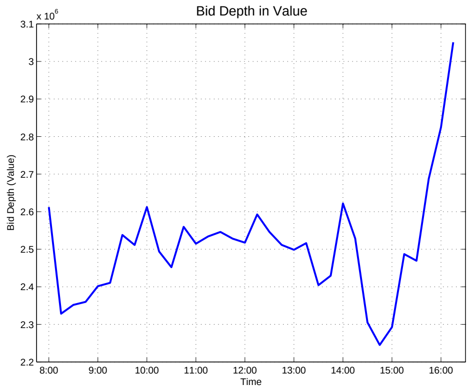


(e) HSBC

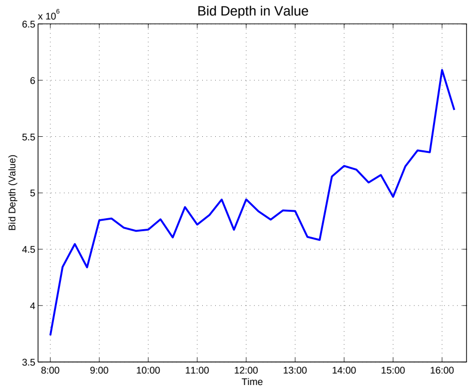


(f) VOD

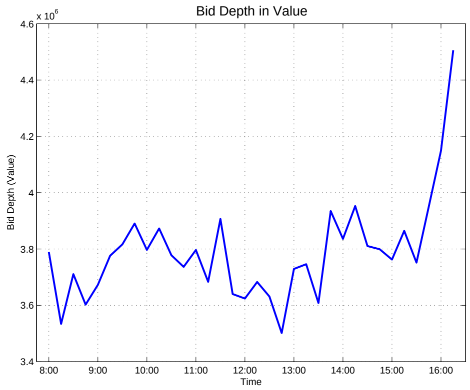
Figure B.3: Intraday Time-weighted Depth in Value (UK stocks)



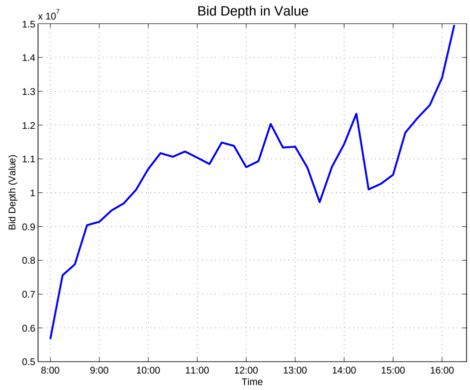
(a) BARC



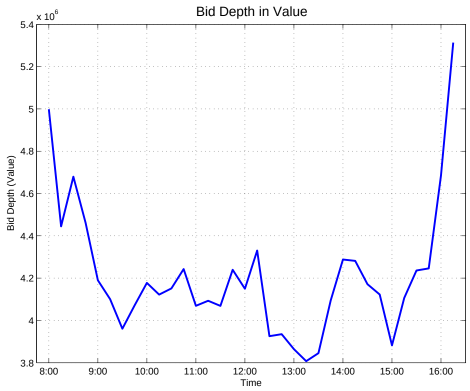
(b) BATS



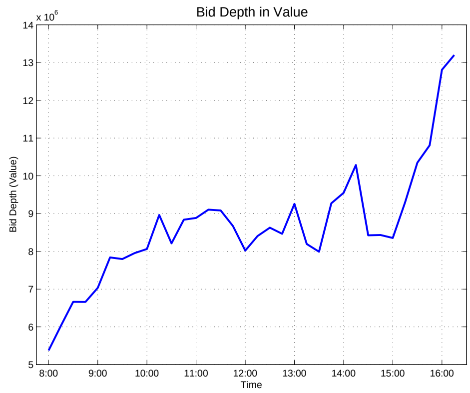
(c) BP



(d) GSK

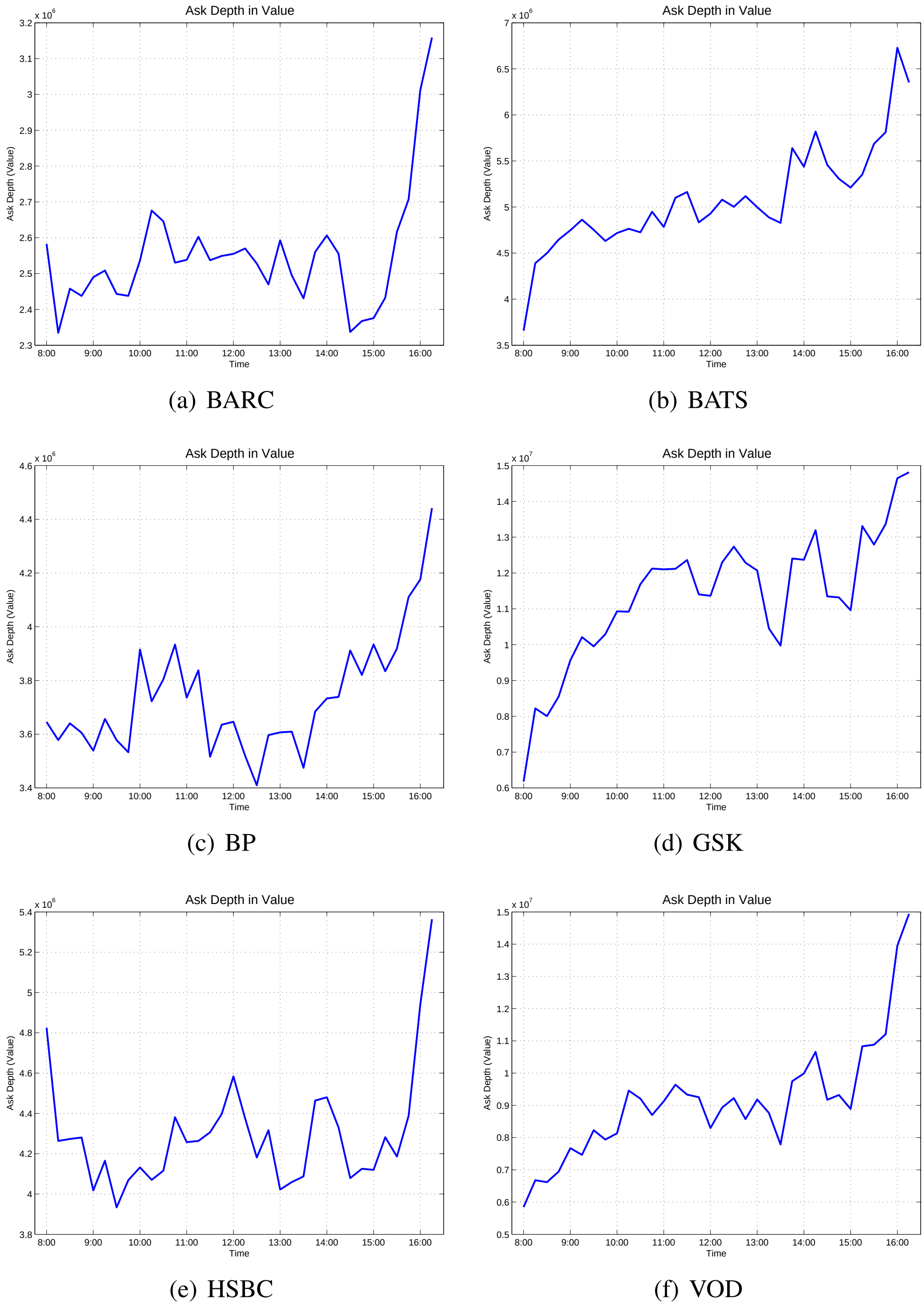


(e) HSBC



(f) VOD

Figure B.4: Intraday Bid Depth in Value (UK stocks)

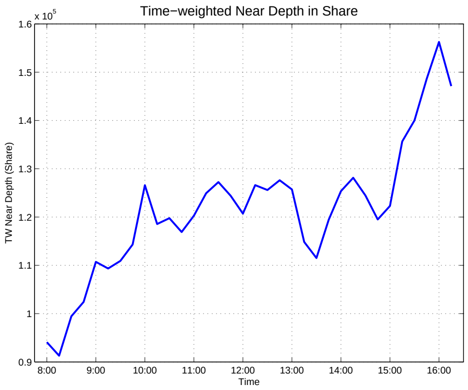


(a) BARC (b) BATS

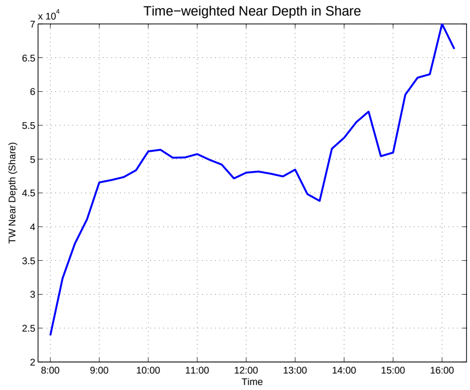
(c) BP (d) GSK

(e) HSBC (f) VOD

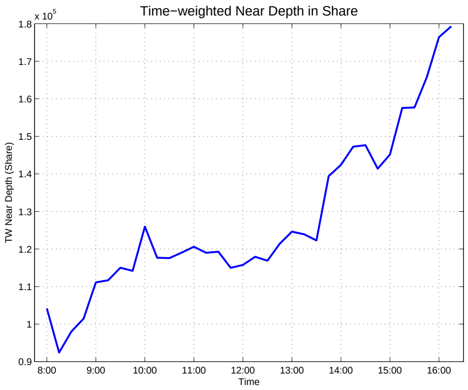
Figure B.5: Intraday Ask Depth in Value (UK stocks)



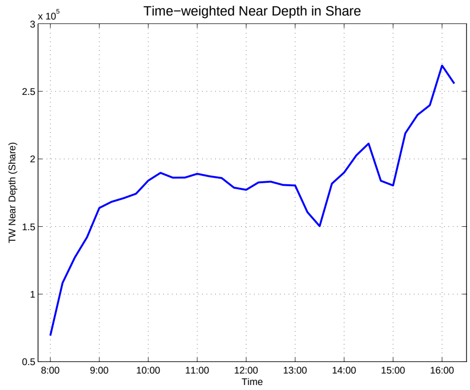
(a) BARC



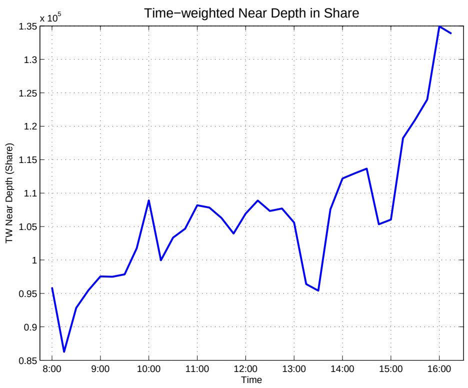
(b) BATS



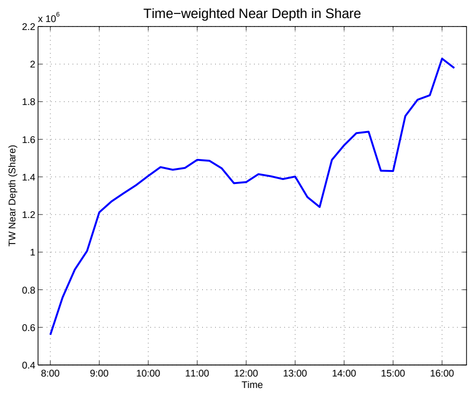
(c) BP



(d) GSK

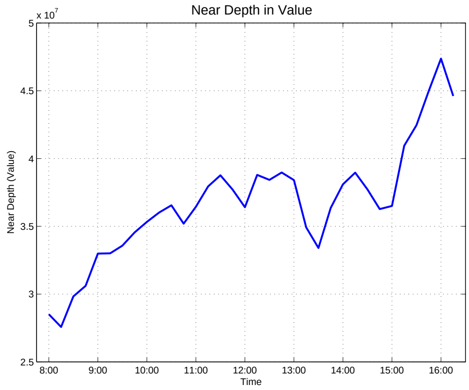


(e) HSBC

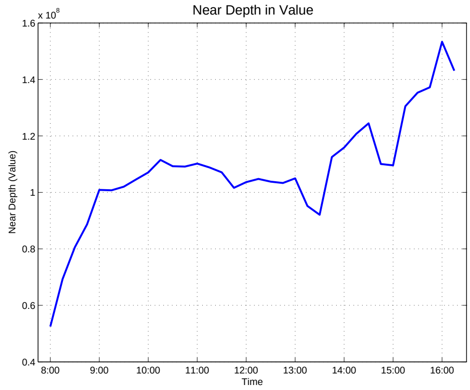


(f) VOD

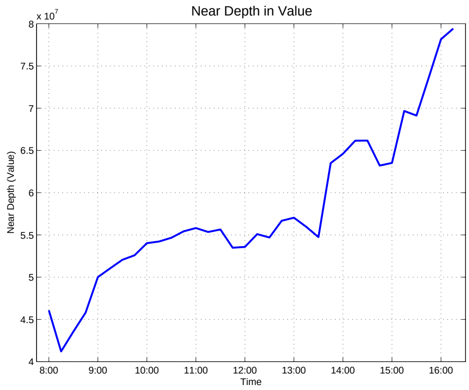
Figure B.6: Intraday Time-weighted Near Depth in Share (UK stocks)



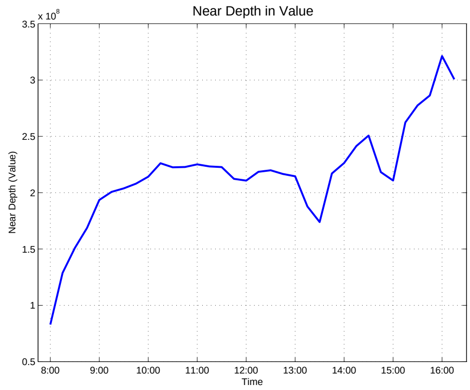
(a) BARC



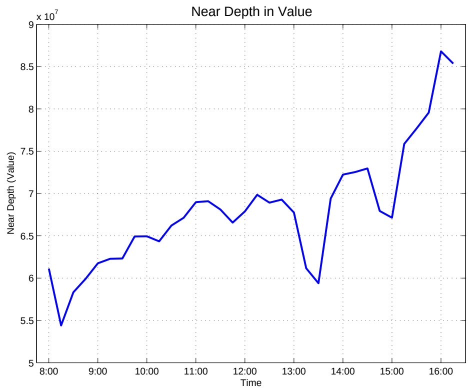
(b) BATS



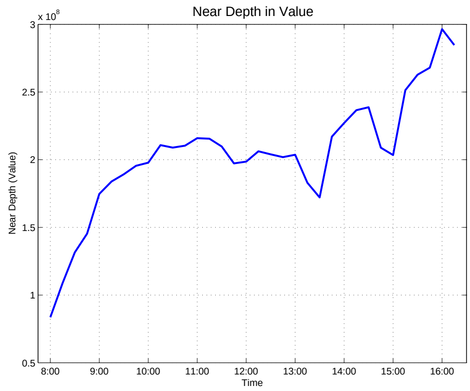
(c) BP



(d) GSK

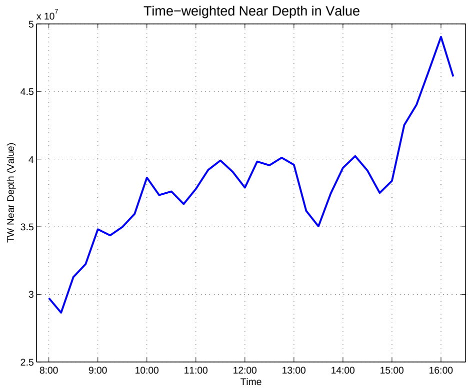


(e) HSBC

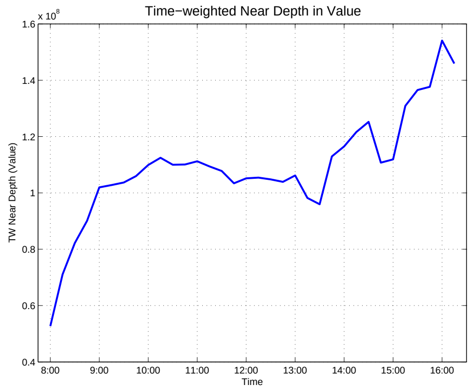


(f) VOD

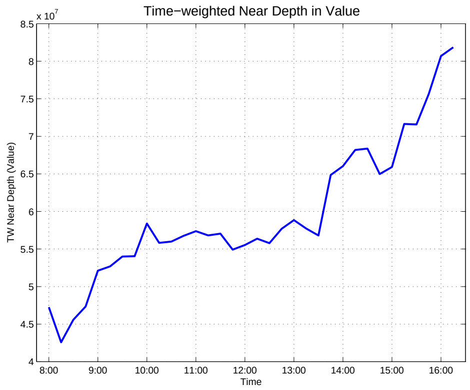
Figure B.7: Intraday Near Depth in Value (UK stocks)



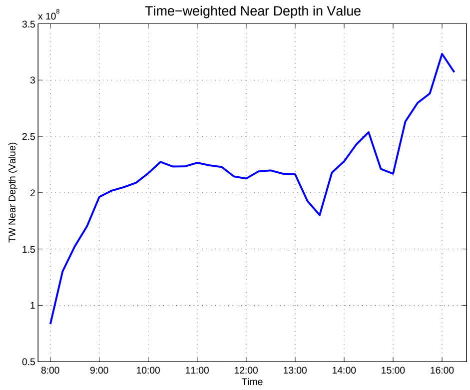
(a) BARC



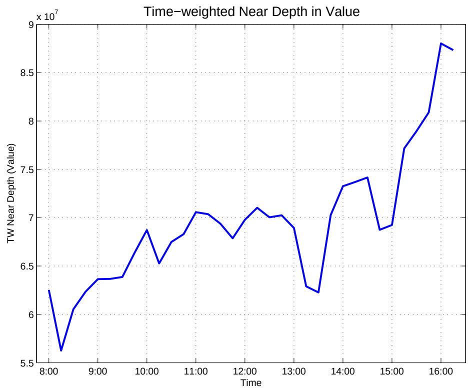
(b) BATS



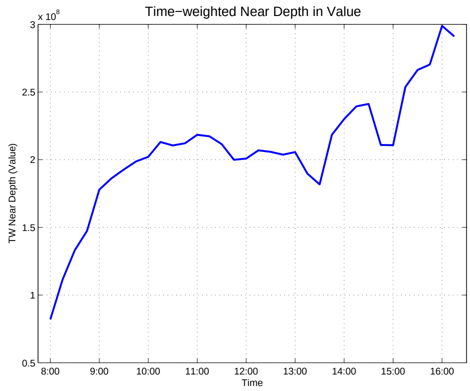
(c) BP



(d) GSK

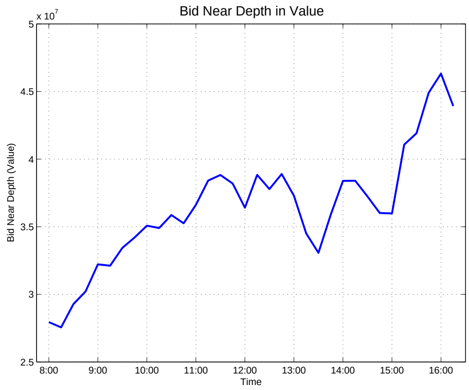


(e) HSBC

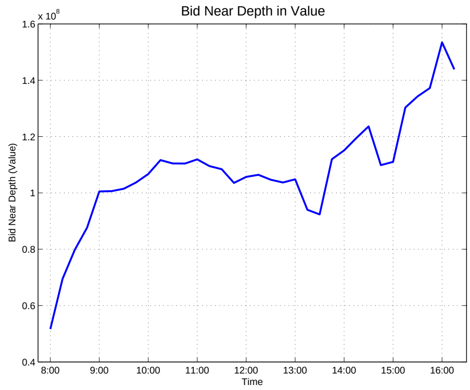


(f) VOD

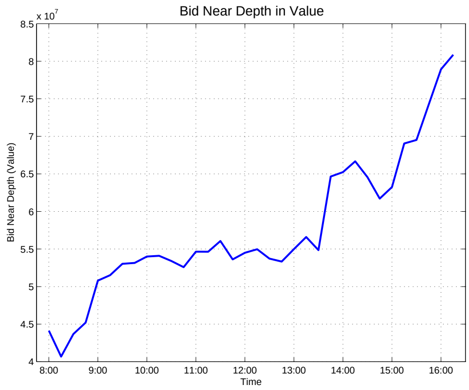
Figure B.8: Intraday Time-weighted Near Depth in Value (UK stocks)



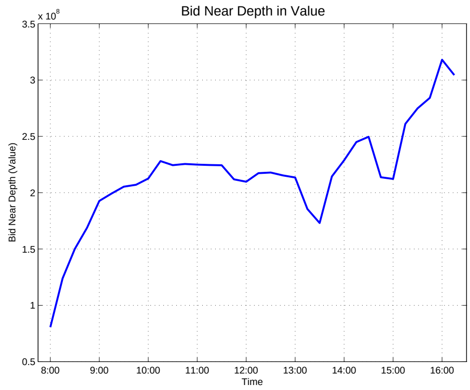
(a) BARC



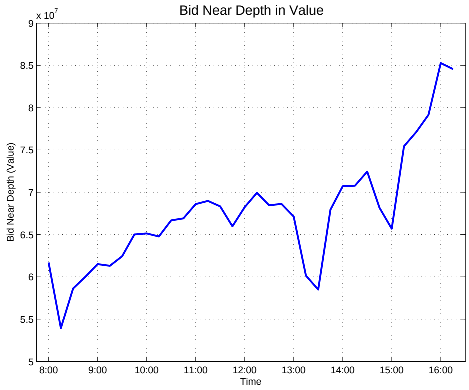
(b) BATS



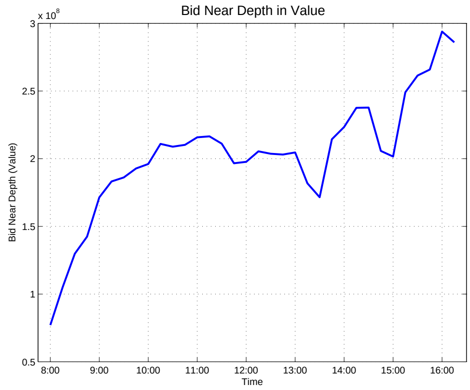
(c) BP



(d) GSK

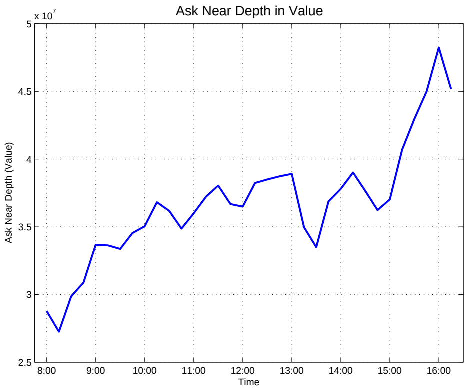


(e) HSBC

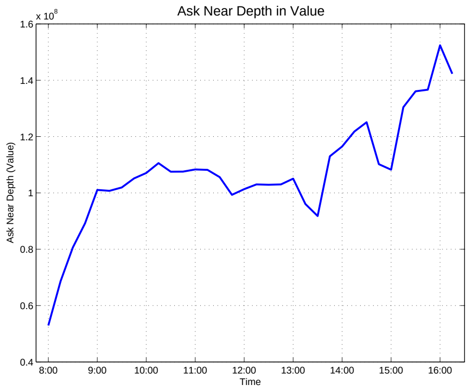


(f) VOD

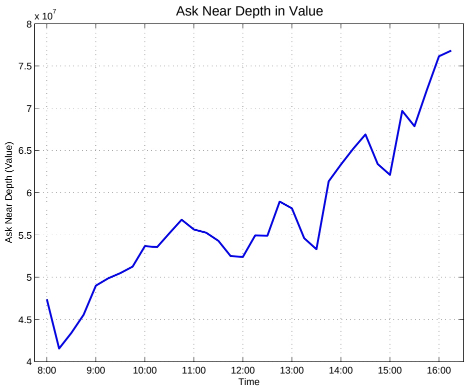
Figure B.9: Intraday Near Bid Depth in Value (UK stocks)



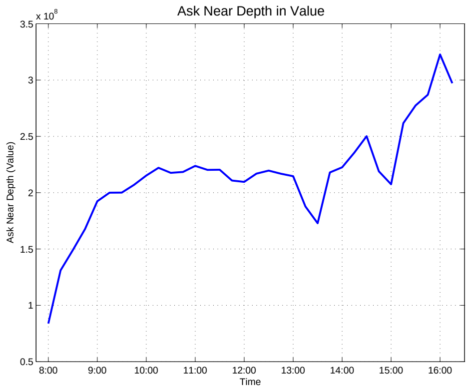
(a) BARC



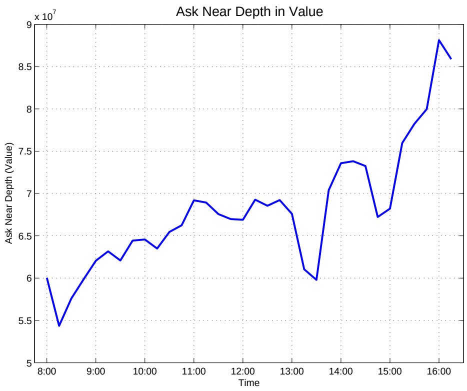
(b) BATS



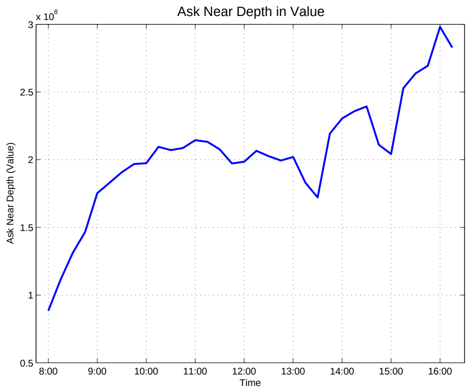
(c) BP



(d) GSK



(e) HSBC



(f) VOD

Figure B.10: Intraday Near Ask Depth in Value (UK stocks)

Bibliography

Abhyankar, A., Ghosh, D., Levin, E. and Limmack, R. (1997). 'Bid-ask spreads, trading volume and volatility: Intra-day evidence from the London Stock Exchange', *Journal of Business Finance and Accounting*, Vol. 24, No. 3, pp. 343–362.

Admati, A. and Pfleiderer, P. (1988). 'A theory of intraday patterns: volume and price variability', *Review of Financial Studies*, Vol. 1, No. 1, pp. 3–40.

Ahn, H., Bae, K. and Chan, K. (2001). 'Limit Orders, Depth, and Volatility: Evidence from the Stock Exchange of Hong Kong', *Journal of Finance*, Vol. 56, No. 2, pp. 767–788.

Ahn, H. and Cheung, Y. (1999). 'The intraday patterns of the spread and depth in a market without market makers: The Stock Exchange of Hong Kong', *Pacific-Basin Finance Journal*, Vol. 7, No. 5, pp. 539–556.

Aitken, M., Almeida, N., Harris, F. and McInish, T. (2007). 'Liquidity supply in electronic markets', *Journal of Financial Markets*, Vol. 10, No. 2, pp. 144–168.

Aitken, M., Berkman, H. and Mak, D. (2001). 'The use of undisclosed limit orders on the Australian Stock Exchange', *Journal of Banking and Finance*, Vol. 25, No. 8, pp. 1589–1603.

Aitken, M., Kua, A., Brown, P., Watter, T. and Izan, H. (1995). 'An Intraday Analysis of the Probability of Trading on the ASX at the Asking Price', *Australian Journal of Management*, Vol. 20, No. 2, pp. 115–154.

Al-Suhaibani, M. and Kryzanowski, L. (2000). 'An exploratory analysis of the order book, and order flow and execution on the Saudi stock market', *Journal of Banking and Finance*, Vol. 24, No. 5, pp. 1323–1357.

Alfonsi, A., Fruth, A. and Schied, A. (2010). 'Optimal execution strategies in limit order books with general shape functions', *Quantitative Finance*, Vol. 10, No. 2, pp. 143–157.

Alfonsi, A. and Schied, A. (2010). 'Optimal Trade Execution and Absence of Price Manipulations in Limit Order Book Models', *SIAM J. on Financial Mathematics*, Vol. 1, No. 3, pp. 490–522.

Allen, F. and Karjalainen, R. (1999). 'Using genetic algorithms to find technical trading rules', *Journal of Financial Economics*, Vol. 51, No. 3, pp. 245–271.

Almgren, R. (2003). 'Optimal execution with nonlinear impact functions and trading-enhanced risk', *Applied Mathematical Finance*, Vol. 10, No. 1, pp. 1–18.

Almgren, R. (2008). 'Execution costs', in R. Cont (ed.), *Encyclopedia of Quantitative Finance*, pp. 1–5, Wiley, UK.

Almgren, R. and Chriss, N. (1999). 'Value under liquidation', *Journal of Risk*, Vol. 12, No. 12, pp. 61–63.

Almgren, R. and Chriss, N. (2000). 'Optimal execution of portfolio transactions', *Journal of Risk*, Vol. 3, No. 2, pp. 5–39.

Aloud, M. and Tsang, E. (2011). 'Modelling the trading behaviour in high-frequency markets', in *Proceedings of the 3rd Computer Science and Electronic Engineering Conference*, pp. 7–12.

Anand, A., Chakravarty, S. and Martell, T. (2005). 'Empirical evidence on the evolution of liquidity: Choice of market versus limit orders by informed and uninformed traders', *Journal of Financial Markets*, Vol. 8, No. 3, pp. 289–309.

Arifovic, J. (1995). 'Genetic algorithms and inflationary economies', *Journal of Monetary Economics*, Vol. 36, No. 1, pp. 219–243.

Arifovic, J. (1996). 'The Behavior of the Exchange Rate in the Genetic Algorithm and Experimental Economies', *Journal of Political Economy*, Vol. 104, No. 3, pp. 510–541.

Arthur, W., Holland, J., LeBaron, B., Palmer, R. and Tayler, P. (1997). 'Asset pricing under endogenous expectations in an artificial stock market', in W. Arthur, S. Durlauf, & D. Lane (eds.), *The Economy as an Evolving Complex System II*, Addison-Wesley.

Atkins, A. and Basu, S. (1995). *Journal of Business Finance and Accounting*, Vol. 22, No. 6, pp. 789–809.

Bae, K., Jang, H. and Park, K. (2003). 'Traders' choice between limit and market orders: evidence from NYSE stocks', *Journal of Financial Markets*, Vol. 6, No. 4, pp. 517–538.

Bagehot, W. (1971). 'The only game in town', *Financial Analysts Journal*, Vol. 27, No. 2, pp. 12–14.

Bak, P., Paczuski, M. and Shubik, M. (1997). *Physica A: Statistical Mechanics and its Applications*, Vol. 246, No. 3-4, pp. 430–453.

Barclay, M. and Warner, J. (1993). 'Stealth trading and volatility : Which trades move prices?', *Journal of Financial Economics*, Vol. 34, No. 3, pp. 281–305.

Barndorff-Nielsen, O., Hansen, P., Lunde, A. and Shephard, N. (2009). 'Realized kernels in practice: trades and quotes', *Econometrics Journal*, Vol. 12, No. 3, pp. 1–32.

Bartolozzi, M. (2010). 'A Multi Agent Model for the Limit Order Book Dynamics', *European Physical Journal B*, Vol. 78, No. 2, pp. 265–273.

Beber, A. and Caglio, C. (2005). 'Order Submission Strategies and Information: Empirical Evidence from the NYSE'. Available at: http://ideas.repec.org/p/fam/rpseri/rp146.html.

Becker, G. (1962). 'Irrational Behavior and Economic Theory', *Journal of Political Economy*, Vol. 70, No. 1, pp. 1–13.

Berkowitz, S., Logue, D. and Noser, E. (1988). 'The total cost of transactions on the NYSE', *Journal of Finance*, Vol. 43, No. 1, pp. 97–112.

Bertsimas, D. and Lo, A. (1998). 'Optimal control of execution costs', *Journal of Financial Markets*, Vol. 1, No. 1, pp. 1–50.

Bessembinder, H. (2003). 'Issues in assessing trade execution costs', *Journal of Financial Markets*, Vol. 6, No. 3, pp. 233–257.

Bessembinder, H., Panayides, M. and Venkataraman, K. (2009). 'Hidden liquidity: An analysis of order exposure strategies in electronic stock markets', *Journal of Financial Economics*, Vol. 94, No. 3, pp. 361–383.

Biais, B., Hillion, P. and Spatt, C. (1995). 'An empirical analysis of the limit order book and the order flow in the Paris Bourse', *Journal of Finance*, Vol. 50, No. 3, pp. 1655–1689.

Bikker, J., Spierdijk, L. and Sluis, P. (2007). 'Market Impact Costs of Institutional Equity Trades', *Journal of International Money and Finance*, Vol. 1, No. 27, pp. 974–1000.

Black, F. (1971). 'Toward a Fully Automated Stock Exchange, Part I', *Financial Analysts Journal*, Vol. 27, No. 4, pp. 28–35.

Bloomfield, R., O'Hara, M. and Saar, G. (2005). 'The "make or take" decision in an electronic market: Evidence on the evolution of liquidity', *Journal of Financial Economics*, Vol. 75, No. 2, pp. 165–199.

Blume, M. and Goldstein, M. (1997). 'Quotes, Order Flow, and Price Discovery', *Journal of Finance*, Vol. 52, No. 1, pp. 221–244.

Bollerslev, T. and Domowitz, I. (1993). 'Some effects of restricting the electronic order book in an automated trade execution system', in D. Friedman & J. Rust (eds.), *The Double Auction Market: Institutions, Theories, and Evidence*, pp. 221–252, Perseus Publishing.

Bongiovanni, S., Borkovec, M. and Sinclair, R. (2006). 'Let's Play Hide-and-Seek: The Location and Size of Undisclosed Limit Order Volume', *Journal of Trading*, Vol. 1, No. 3, pp. 34–46.

Boscuk, A. and Lasfer, M. (2005). 'The information content of institutional trades on the London stock exchange', *Journal of Financial and Quantitative Analysis*, Vol. 40, No. 3, pp. 621–644.

Bouchaud, J., Farmer, J. and Lillo, F. (2009). 'How markets slowly digest changes in supply and demand', in T. Hens & K. Schenk-Hoppe (eds.), *Handbook of Financial Markets: Dynamics and Evolution*, pp. 57–160, North Holland, US.

Bouchaud, J., Gefen, Y., Potters, M. and Wyart, M. (2004). 'Fluctuations and response in financial markets: The subtle nature of 'random' price changes', *Quantitative Finance*, Vol. 4, No. 2, pp. 57–62.

Bouchaud, J., Mezard, M. and Potters, M. (2002). 'Statistical properties of stock order books: empirical results and models', *Quantitative Finance*, Vol. 2, No. 4, pp. 251–256.

Brabazon, A. and O'Neill, M. (2006). *Biologically Inspired Algorithms for Financial Modeling*, Spring-Verlag, Berlin.

Brabazon, A., O'Neill, M. and Dempsey, I. (2008). 'An Introduction to Evolutionary Computation in Finance', *IEEE Computational Intelligence Magazine*, Vol. 3, No. 4, pp. 42–55.

Brock, W. and Kleidon, A. (1992). 'Periodic market closure and trading volume: A model of intraday bids and asks', *Journal of Economic Dynamics and Control*, Vol. 16, No. 3, pp. 451–489.

Brownlees, C. and Gallo, G. (2006). 'Financial Econometric Analysis at Ultra-high Frequency: Data Handling Concerns', *Computational Statistics and Data Analysis*, Vol. 51, No. 4, pp. 2232–2245.

Cai, C., Hudson, R. and Keasey, K. (2004). 'Intraday bid-ask spreads, trading volume and volatility: recent empirical evidence from the London Stock Exchange', *Journal of Business Finance and Accounting*, Vol. 31, No. 5, pp. 647–676.

Cao, C., Hansch, O. and Wang, X. (2008). 'Order Placement Strategies In A Pure Limit Order Book Market', *Journal of Financial Research*, Vol. 26, No. 2, pp. 113–140.

Cao, C., Hansch, O. and Wang, X. (2009). 'The information content of an open limit order book', *Journal of Futures Markets*, Vol. 29, No. 1, pp. 16–41.

CFA (2009). 'Market Microstructure: The Impact of Fragmentation under the Markets in Financial Instruments Directive', *CFA Institute Publications*, Vol. 2009, No. 13, pp. 1–60.

Chaboud, A., Chiquoine, B., Hjalmarsson, E. and Vega, C. (2009). 'Rise of the machines: algorithmic trading in the foreign exchange market'. Working paper, Board of Governors of the Federal Reserve System (US).

Chakraborti, A., Toke, I., Patriarca, M. and Abergel, F. (2011). 'Econophysics review: I. Empirical facts', *Quantitative Finance*, Vol. 11, No. 7, pp. 991–1012.

Chakravarty, S. and Holden, C. (1995). 'An integrated model of market and limit orders', *Journal of Financial Intermediation*, Vol. 4, No. 3, pp. 213–241.

Challet, D. and Stinchcombe, R. (2001). 'Analyzing and modeling 1+ 1d markets', *Physica A*, Vol. 300, No. 2, pp. 285–299.

Chan, K., Christie, W. and Schultz, P. (1995). 'Market structure and the intraday pattern of bid-ask spreads for NASDAQ securities', *Journal of Business*, Vol. 68, No. 1, pp. 35–60.

Chan, L. and Lakonishok, J. (1995). 'The behavior of stock prices around institutional trades', *Journal of Finance*, Vol. 50, No. 5, pp. 1147–1174.

Chan, L. and Lakonishok, J. (1997). 'Institutional Equity Trading Costs: NYSE versus NASDAQ', *Journal of Finance*, Vol. 52, No. 2, pp. 713–735.

Chan, L. K. and Lakonishok, J. (1993). 'Institutional trades and intraday stock price behavior', *Journal of Financial Economics*, Vol. 33, No. 3, pp. 173–199.

Chan, N., LeBaron, B., Lo, A. and Poggio, T. (2001). 'Agent-based models of financial markets: A comparison with experimental markets'. Working paper, MIT.

Chan, Y. (2000). 'The price impact of trading on the stock exchange of Hong Kong', *Journal of Financial Markets*, Vol. 3, No. 1, pp. 1–16.

Chavez-Demoulin, V. and McGill, J. (2012). 'High-frequency financial data modeling using Hawkes processes', *Journal of Banking & Finance*, Vol. 36, No. 12, pp. 3415 – 3426.

Chen, S. (2007). 'Computationally intelligent agents in economics and finance', *Information Sciences*, Vol. 177, No. 5, pp. 1153–1168.

Chen, S. (2012). 'Varieties of agents in agent-based computational economics: A historical and an interdisciplinary perspective', *Journal of Economic Dynamics and Control*, Vol. 36, No. 1, pp. 1–25.

Chen, S., Chang, C. and Du, Y. (2012). 'Agent-based economic models and econometrics', *Knowledge Engineering Review*, Vol. 27, No. 2, pp. 187–219.

Chen, S. and Yeh, C. (1997). 'Toward a computable approach to the efficient market hypothesis: An application of genetic programming', *Journal of Economic Dynamics and Control*, Vol. 21, No. 5, pp. 1043–1063.

Chen, S. and Yeh, C. (2002). 'On the emergent properties of artificial stock markets: the efficient market hypothesis and the rational expectations hypothesis', *Journal of Economic Behavior & Organization*, Vol. 49, No. 2, pp. 217–239.

Cheung, Y., Ho., R., Pope, P. and Draper, P. (1994). 'Intraday stock return volatility: The Hong Kong evidence', *Pacific-Basin Finance Journal*, Vol. 2, No. 1, pp. 261–276.

Chiarella, C., Dieci, R. and He, X. (2009a). 'Heterogeneity, Market Mechanisms, and Asset Price Dynamics', in *Handbook of Financial Markets: Dynamics and Evolution*, pp. 277–344, North Holland, Amsterdam.

Chiarella, C. and Iori, G. (2002). 'A simulation analysis of the microstructure of double auction markets', *Quantitative Finance*, Vol. 2, No. 2, pp. 246–253.

Chiarella, C., Iori, G. and Perello, J. (2009b). 'The impact of heterogeneous trading rules on the limit order book and order flows', *Journal of Economic Dynamics and Control*, Vol. 33, No. 3, pp. 525–537.

Chiyachantana, C., Jain, P., Jiang, C. and Wood, R. (2004). 'International Evidence on Institutional Trading Behavior and Price Impact', *Journal of Finance*, Vol. 59, No. 2, pp. 869–898.

Chung, K., Van Ness, B. and Van Ness, R. (1999). 'Limit orders and the bid-ask spread', *Journal of Financial Economics*, Vol. 53, No. 3, pp. 255–287.

Collins, B.M.and Fabozzi, F. (1991). 'A Methodology for Measuring Transaction Costs', *Financial Analysts Journal*, Vol. 47, No. 2, pp. 27–36.

Consiglio, A. (ed.) (2007). *Artificial Markets Modeling: Methods and Applications*, Springer.

Cont, R. (2001). 'Empirical properties of asset returns: stylized facts and statistical issues', *Quantitative Finance*, Vol. 1, No. 2, pp. 223–236.

Copeland, T. and Galai, D. (1983). 'Information Effects on the Bid-Ask Spread', *Journal of Finance*, Vol. 38, No. 5, pp. 1457–1469.

Coppejans, M., Domowitz, I. and Madhavan, A. (2003). 'Dynamics of Liquidity in an Electronic Limit Order Book Market'. Working paper, Duke University.

Cristelli, M., Pietronero, L. and Zaccaria, A. (2011). 'Critical Overview of Agent-Based Models for Economics'. Working paper, arXiv. org.

Cui, W. and Brabazon, A. (2011). 'Ultra High-frequency Financial Data: Analysing NYSE Euronext TAQ Data', in *Proceedings of the 24nd Irish Accounting and Finance Association Conference*, pp. 1–38.

Cui, W. and Brabazon, A. (2012a). 'An Agent-based Modelling Approach to Study Price Impact', in *Proceedings of the 2012 IEEE Conference on Computational Intelligence for Financial Engineering & Economics*, pp. 1–8, IEEE Press.

Cui, W. and Brabazon, A. (2012b). 'An Analysis of Liquidity in London Stock Exchange', in *Proceedings of the 25th Irish Accounting and Finance Association Conference*, pp. 1–45.

Cui, W., Brabazon, A. and O'Neill, M. (2009a). 'Efficient Trade Execution Using a Genetic Algorithm in an Order Book Based Artificial Stock Market', in *Proceedings of the 11th Annual Conference Companion on Genetic and Evolutionary Computation Conference*, pp. 2023–2028, ACM Press, New York, NY, USA.

Cui, W., Brabazon, A. and O'Neill, M. (2009b). 'Evolutionary Computation in Trade Execution', in *Proceedings of the 22nd Irish Accounting and Finance Association Conference*, pp. 1–10, Dublin, Ireland.

Cui, W., Brabazon, A. and O'Neill, M. (2010a). 'An Analysis of Price Impact', in *Proceedings of the 3rd International Conference of European Research Consortium on Informatics and Mathematics*, London, UK.

Cui, W., Brabazon, A. and O'Neill, M. (2010b). 'Evolutionary Computation and

Trade Execution', in A. Brabazon & M. O'Neill (eds.), *Natural Computing in Computational Finance (volume III)*, pp. 51–69, Springer-Verlag, Berlin.

Cui, W., Brabazon, A. and O'Neill, M. (2010c). 'Evolving Dynamic Trade Execution Strategies Using Grammatical Evolution', in *Proceedings of the 2010 International Conference on Applications of Evolutionary Computation, Lecture Notes in Computer Science 6025*, pp. 191–200, Springer-Verlag, Berlin.

Cui, W., Brabazon, A. and O'Neill, M. (2010d). 'Evolving Efficient Limit Order Strategy Using Grammatical Evolution', in *Proceedings of the 2010 IEEE Congress On Evolutionary Computation*, pp. 1–6, IEEE Press.

Cui, W., Brabazon, A. and O'Neill, M. (2011). 'Dynamic Trade Execution: A Grammatical Evolution Approach', *International Journal of Financial Markets and Derivatives*, Vol. 2, No. 1/2, pp. 4–31.

Daniel, G. (2006). *Asynchronous simulations of a limit order book*, PhD thesis, University of Manchester, UK.

Daniels, M., Farmer, D., Gillemot, L., Iori, G. and Smith, E. (2003). 'Quantitative Model of Price Diffusion and Market Friction Based on Trading as a Mechanistic Random Process', *Physical Review Letters*, Vol. 90, No. 10, pp. 1–4.

Darley, V. and Outkin, A. (2007). *A Nasdaq Market Simulation: Insights on a Major Market from the Science of Complex Adaptive Systems*, World Scientific Publishing Company.

Darwin, C. (1859). *On The Origin of Species: By Means of Natural Selection or the Preservation of Favoured Races in the Struggle for Life (reprinted 1985)*, Penguin Books, London.

Decovny, S. (2008). 'Asian-Pacific gears up for algorithmic trading', *Market View*, Vol. 1, No. 1, p. 13.

Dempsey, I., O'Neill, M. and Brabazon, A. (2009). *Foundations in Grammatical Evolution for Dynamic Environment*, Springer-Verlag, Berlin.

Dempster, M. and Jones, C. (2001). 'A real-time adaptive trading system using Genetic Programming', *Quantitative Finance*, Vol. 1, No. 2, pp. 397–413.

Duffy, J. (2006). 'Agent-Based Models and Human Subject Experiments', in L. Tesfatsion & K. L. Judd (eds.), *Handbook of Computational Economics*, Vol. 2, Elsevier.

Dufour, A. and Engle, R. (2000). 'Time and the price impact of a trade', *Journal of Finance*, Vol. 55, No. 6, pp. 2467–2498.

Duong, H., Kalev, P. and Krishnamurti (2009). 'Order aggressiveness of institutional and individual investors', *Pacific-Basin Finance Journal*, Vol. 1, No. 4, pp. 1–14.

Durbin, M. (2010). *All about high-frequency trading*, McGraw-Hill.

Dworman, G., Kimbrough, S. and Laing, J. (1996). 'On automated discovery of models using genetic programming:bargaining in a three-agent coalitions game', *Journal of Management Information Systems*, Vol. 12, No. 3, pp. 97–125.

Easley, D. and O'Hara, M. (1992). 'Time and the process of security price adjustment', *Journal of Finance*, Vol. 47, No. 2, pp. 577–605.

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

Ellul, A., Holden, C., Jain, P. and Jennings, R. (2007). 'Order dynamics: recent evidence from the NYSE', *Journal of Empirical Finance*, Vol. 14, No. 3, pp. 636–661.

Engle, R. and Patton, A. (2004). 'Impacts of trades in an error-correction model of quote prices', *Journal of Financial Markets*, Vol. 7, No. 1, pp. 1–25.

Esser, A. and Monch, B. (2007). 'The navigation of an iceberg: The optimal use of hidden orders', *Finance Research Letters*, Vol. 4, No. 1, pp. 68–81.

Evans, M. and Lyons, R. (2002). 'Order Flow and Exchange Rate Dynamics', *Journal of Political Economy*, Vol. 110, No. 1, pp. 170–180.

Farmer, D., Gillemot, L., Iori, G., Krishnamurthy, S., Smith, E. and Daniels, M. (2005a). 'A random order placement model of price formation in the continuous double auction', in L. Blume & S. Durlauf (eds.), *The Economy as an Evolving Complex System III*, Oxford University Press.

Farmer, J. and Foley, D. (2009). 'The economy needs agent-based modelling', *Nature*, Vol. 460, pp. 685–686.

Farmer, J., Gerig, A., Lillo, F. and Mike, S. (2006). 'Market efficiency and the long-memory of supply and demand: Is price impact variable and permanent or fixed and temporary?', *Quantitative Finance*, Vol. 6, No. 2, pp. 107–112.

Farmer, J. and Lillo, F. (2004). 'On the origin of power-law tails in price fluctuations', *Quantitative Finance*, Vol. 4, No. 1, pp. 7–10.

Farmer, J. and Zamani, N. (2007). 'Mechanical vs. informational components of price impact', *European Physical Journal B.*, Vol. 55, No. 1, pp. 1899–2000.

Farmer, J. D., Gillemot, L., Lillo, F., Mike, S. and Sen, A. (2004). 'What really causes large price changes?', *Quantitative Finance*, Vol. 4, No. 4, pp. 383–397.

Farmer, J. D., Patelli, P. and Zovko, I. (2005b). 'The predictive power of zero intelligence models in financial markets', in *Proceedings of the National Academy of Sciences of the United States of America*, Vol. 102, pp. 2254–2259.

Farmer, J. D., Patelli, P. and Zovko, I. (2005c). 'Supplementary Material for 'The Predictive Power of Zero Intelligence in Financial Markets''. Available at: www.santafe.edu/jdf/papers/zerosuppl.pdf.

Feldman, T. and Friedman, D. (2010). 'Human and Artificial Agents in a Crash-Prone Financial Market', *Computational Economics*, Vol. 36, No. 3, pp. 201–229.

FinancialTimes (2011a). 'Daily Market Share: equities (source Thompson Reuters Equity Market Share Reporter)'. Available at: http://www.ft.com/intl/tradingroom (accessed 3 February 2011).

FinancialTimes (2011b). 'Nasdaq OMX lures trades to platform'. Available at: http://www.ft.com/tradingroom (accessed 30 January 2011).

Foster, F. and Viswanathan, S. (1993). 'Variations in Trading Volume, Return Volatility, and Trading Costs: Evidence on Recent Price Formation Models', *Journal of Finance*, Vol. 48, No. 1, pp. 187–211.

Foucault, T. (1999). 'Order flow composition and trading costs in a dynamic limit order market', *Journal of Financial Markets*, Vol. 2, No. 2, pp. 99–134.

Foucault, T., Kaden, O. and Kandel, E. (2005). 'The limit order book as a market for liquidity', *Review of Financial Studies*, Vol. 18, No. 1, pp. 1171–1217.

Foucault, T. and Menkveld, A. (2008). 'Competition for Order Flow and Smart Order Routing Systems', *Journal of Finance*, Vol. 63, No. 1, pp. 119–158.

French, K. and Roll, R. (1986). 'Stock return variances: the arrival of information and the reaction of traders', *Journal of Financial Economics*, Vol. 17, No. 1, pp. 5–26.

Fruth, A., Schoeneborn, T. and Urusov, M. (2011). 'Optimal trade execution and price manipulation in order books with time-varying liquidity'. Available at: http://arxiv.org/abs/1109.2631.

Gabaix, X., Gopikrishnan, P., Plerou, V. and Stanley, H. (2003). 'A theory of power law distributions in financial market fluctuations', *Nature*, Vol. 423, pp. 267–270.

Gabaix, X., Gopikrishnan, P., Plerou, V. and Stanley, H. (2006). 'Institutional Investors and Stock Market Volatility', *Quarterly Journal of Economics*, Vol. 121, No. 2, pp. 461–504.

Gatheral, J. (2010). 'No-Dynamic-Arbitrage and Market Impact', *Quantitative Finance*, Vol. 10, No. 7, pp. 749–759.

Gatheral, J., Schied, A. and Slynko, A. (2011). 'Exponential Resilience and Decay of Market Impact', in F. Abergel, B. Chakrabarti, A. Chakraborti, & M. Mitra (eds.), *Econophysics of Order-driven Markets*, Spring-Verlag, Italy.

Gemmill, G. (1996). 'Transparency and Liquidity: A Study of Block Trades on the London Stock Exchange under Different Publication Rules', *Journal of Finance*, Vol. 51, No. 5, pp. 1765–1790.

Gillemot, L., Farmer, J. and Lillo, F. (2005). 'There's more to volatility than volume'. Available at: http://arxiv.org/abs/physics/0510007.

Glosten, L. (1994). 'Is the Electronic Open Limit Order Book Inevitable?', *Journal of Finance*, Vol. 49, No. 4, pp. 1127–1161.

Glosten, L. and Milgrom, P. (1985). 'Bid, ask and transaction prices in a specialist market with heterogeneously informed traders', *Journal of Financial Economics*, Vol. 14, No. 1, pp. 71–100.

Gode, D. and Sunder, S. (1993). 'Allocative efficiency of markets with zero-intelligence traders', *Journal of Political Economy*, Vol. 101, No. 1, pp. 119–137.

Goettler, R., Parlour, C. and Rajan, U. (2005). 'Equilibrium in a dynamic limit order book', *Journal of Finance*, Vol. 60, No. 2, pp. 2149–2192.

Goldstein, M., Irvine, P., Kandel, E. and Wiener, Z. (2009). 'Brokerage Commissions and Institutional Trading Patterns', *Review of Financial Studies*, Vol. 22, No. 12, pp. 5175–5212.

Gould, M., Porter, M., Williams, S., McDonald, M., Fenn, D. and Howison, S. (2011). 'Limit order books'. Working paper, arXiv. org.

Grant, J. (2012). 'China's Citic in algorithmic trading deal'. Available at: http://www.ft.com/intl (accessed 14 May 2012).

Griffiths, M., Smith, B., Turnbull, D. and White, R. (2000). 'The costs and determinants of order aggressiveness', *Journal of Financial Economics*, Vol. 56, No. 1, pp. 65–88.

Hall, A. and Hautsch, N. (2006). 'Order aggressiveness and order book dynamics', *Empirical Economics*, Vol. 30, No. 1, pp. 973–1005.

Hall, A. and Hautsch, N. (2007). 'Modelling the buy and sell intensity in a limit order book market', *Journal of Financial Markets*, Vol. 10, No. 3, pp. 249–286.

Handa, P. and Schwartz, R. (1996). 'Limit Order Trading', *Journal of Finance*, Vol. 51, No. 5, pp. 1835–1861.

Handa, P., Schwartz, R. and Tiwari, A. (2003). 'Quote setting and price formation in an order-driven market', *Journal of Financial Markets*, Vol. 6, No. 1, pp. 461–489.

Harris, L. (1986). 'A transaction data study of weekly and intraday patterns in stock returns', *Journal of Financial Economics*, Vol. 16, No. 2, pp. 99–117.

Harris, L. (1998). 'Optimal Dynamic Order Submission Strategies in Some Stylized Trading Problems', *Financial Markets, Institutions and Instruments*, Vol. 7, No. 2, pp. 1–76.

Harris, L. and Hasbrouck, J. (1996). 'Market vs. Limit Orders - The SuperDOT Evidence On Order Submission Strategy', *Journal of Financial and Quantitative Analysis*, Vol. 31, No. 2, pp. 213–231.

Hasbrouck, J. (1991). 'Measuring the information content of stock trades', *Journal of Finance*, Vol. 46, No. 2, pp. 178–208.

Hasbrouck, J. (1999). 'The dynamics of discrete bid and ask quotes', *The Journal of finance*, Vol. 54, No. 6, pp. 2109–2142.

Hasbrouck, J. and Saar, G. (2002). 'Limit Orders and Volatility in a Hybrid Market: The Island ECN'. Available at: http://ssrn.com/abstract310940.

Hasbrouck, J., Sofianos, G. and Sosebee, D. (1993). 'New York Stock Exchange System and Trading Procedures'. Available at: http://citeseerx.ist.psu.edu/viewdoc/summarydoi10.1.1.132.5638.

Hausman, J., Lo, A. and MacKinlay, A. (1992). 'An ordered probit analysis of transaction stock prices', *Journal of Financial Economics*, Vol. 31, No. 3, pp. 319–379.

Hautsch, N. and Huang, R. (2012). 'The market impact of a limit order', *Journal of Economic Dynamics and Control*, Vol. 36, No. 4, pp. 501 – 522.

Hendershott, T., Jones, C. and Menkveld, A. (2011). 'Does Algorithmic Trading Improve Liquidity?', *Journal of Finance*, Vol. 66, No. 1, pp. 1–33.

Hendershott, T. and Riordan, R. (2012). 'Algorithmic Trading and the Market for Liquidity', *Journal of Financial and Quantitative Analysis (forthcoming)*.

Henker, T. and Wang, J. (2006). 'On the importance of timing specifications in market microstructure research', *Journal of Financial Markets*, Vol. 9, No. 2, pp. 162–179.

Ho, Y. and Cheung, Y. (1991). 'Behaviour of intra-daily stock return on an Asian emerging market - Hong Kong', *Applied Economics*, Vol. 23, No. 5, pp. 957–966.

Holthausen, R., Leftwich, R. and Mayers, D. (1987). 'The effect of large block transactions on security prices: A cross-sectional analysis', *Journal of Financial Economics*, Vol. 19, No. 2, pp. 237–267.

Holthausen, R., Leftwich, R. and Mayers, D. (1990). 'Large-block transactions, the speed of response, and temporary and permanent stock-price effects', *Journal of Financial Economics*, Vol. 26, No. 1, pp. 71–95.

Hommes, C. and Wagener, F. (2009). 'Complex Evolutionary Systems in Behavioral Finance', in T. Hens & K. Schenk (eds.), *Handbook of Financial Markets: Dynamics and Evolution*, North Holland.

Hopman, C. (2007). 'Do supply and demand drive stock prices?', *Quantitative Finance*, Vol. 7, No. 1, pp. 37–53.

Hora, M. (2006). ‘Tactical Liquidity Trading and Intraday Volume’. Available at: http://ssrn.com/abstract=931667.

Huang, Y., Chen, S., Hung, M. and Yu, T. (2012). ‘Liquidity cost of market orders in the Taiwan Stock Market: A study based on an order-driven agent-based artificial stock market’, *International Review of Financial Analysis*, Vol. 23, pp. 72–80.

Hussain, S. (2011). ‘The Intraday Behaviour of Bid-Ask Spreads, Trading Volume and Return Volatility: Evidence from DAX30’, *International Journal of Economics and Finance*, Vol. 3, No. 1, pp. 23–34.

Iori, G., Daniels, M., Farmer, J., Gillemot, L., Krishnamurthy, S. and Smith, E. (2003). ‘An analysis of price impact function in order-driven markets’, *Physica A*, Vol. 324, No. 1-2, pp. 146–151.

Izumi, K., Toriumi, F. and Matsui, H. (2009). ‘Evaluation of automated-trading strategies using an artificial market’, *Neurocomputing*, Vol. 72, No. 1, pp. 3469–3476.

Jain, P. (2003). ‘Institutional design and liquidity at stock exchanges around the world’.

Jain, P. and Gun-Ho, J. (1988). ‘The Dependence between Hourly Prices and Trading Volume’, *Journal of Financial and Quantitative Analysis*, Vol. 23, No. 3, pp. 269–283.

Joulin, A., Lefevre, A., Grunberg, D. and Bouchaud, J. (2008). ‘Stock price jumps: news and volume play a minor role’. Available at: http://arxiv.org/abs/0803.1769.

Kaniel, R. and Liu, H. (2006). ‘So What Orders Do Informed Traders Use?’, *Journal of Business*, Vol. 79, No. 4, pp. 1867–1913.

Keim, D. and Madhavan, A. (1995a). ‘Anatomy of the trading process: Empirical evidence on the behavior of institutional traders’, *Journal of Financial Economics*, Vol. 37, No. 1, pp. 371–398.

Keim, D. and Madhavan, A. (1995b). ‘Execution Costs and Investment Performance: An Empirical Analysis of Institutional Equity Trades’. Available at: http://ideas.repec.org/p/fth/pennfi/09-95.html.

Keim, D. and Madhavan, A. (1996). ‘The upstairs market for large-block transactions: analysis and measurement of price effects’, *Review of Financial Studies*, Vol. 9, No. 1, pp. 1–36.

Keim, D. and Madhavan, A. (1997). ‘Transactions costs and investment style: an inter-exchange analysis of institutional equity trades’, *Journal of Financial Economics*, Vol. 46, No. 3, pp. 265–292.

Kharroubi, I. and Pham, H. (2010). ‘Optimal Portfolio Liquidation with Execution Cost and Risk’, *SIAM Journal on Financial Mathematics*, Vol. 1, No. 1, pp. 897–931.

Kim, G. and Markowitz, H. (1989). ‘Investment rules, margin, and market volatility’, *Journal of Portfolio Management*, Vol. 16, No. 1, pp. 45–52.

King, L. (2011). ‘London Stock Exchange price data failures emerged immediately at Millenium launch’. Available at: http://www.computerworlduk.com/news/itbusiness/3261816/.

Koksal, B. (2012). ‘An analysis of intraday patterns and liquidity on the Istanbul stock exchange’. Available at: http://mpra.ub.unimuenchen.de/35968/.

Koza, J. (2010). ‘Human-competitive results produced by genetic programming’, *Genetic Programming and Evolvable Machines*, Vol. 11, No. 3, pp. 251–284.

Kozhan, R. and Salmon, M. (2011). ‘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.

Kraus, A. and Stoll, H. (1972). ‘Price impacts of block trading on the New York Stock Exchange’, *Journal of Finance*, Vol. 27, No. 3, pp. 569–588.

Kyle, A. (1985). ‘Continuous auctions and insider trading’, *Econometrica*, Vol. 53, No. 3, pp. 1315–1335.

Ladley, D. and Schenk, K. (2009). ‘Do stylised facts of order book markets need strategic behaviour’, *Journal of Economics Dynamics & Control*, Vol. 33, No. 3, pp. 817–831.

Large, J. (2007). ‘Measuring the resiliency of an electronic limit order book’, *Journal of Financial Markets*, Vol. 10, No. 1, pp. 1–25.

LeBaron, B. (2000). ‘Agent-based computational finance: suggested readings and early research’, *Journal of Economic Dynamics and Control*, Vol. 24, No. 1, pp. 679–702.

LeBaron, B. (2001). ‘A Builder’s Guide to Agent Based Financial Markets’, *Quantitative Finance*, Vol. 1, No. 2, pp. 254–261.

LeBaron, B. (2006). ‘Agent-based Computational Finance’, in L. Tesfatsion & K. Judd (eds.), *Handbook of Computational Economics, Volume 2: Agent-based Computational Economics*, pp. 134–151, Elsevier.

Lee, C., Mucklow, B. and Ready, M. (1993). ‘Spreads, depths, and the impact of earnings information: An intraday analysis’, *Review of Financial Studies*, Vol. 6, No. 2, pp. 345–375.

Lee, C. and Ready, M. (1991). ‘Inferring Trade Direction from Intraday Data’, *Journal of Finance*, Vol. 46, No. 2, pp. 733–746.

Lee, Y., Fok, R. and Liu, Y. (2001). ‘Explaining intraday pattern of trading volume from the order flow data’, *Journal of Business Finance and Accounting*, Vol. 28, No. 3, pp. 199–230.

Lehmann, B. and Modest, D. (1994). ‘Trading and Liquidity on the Tokyo Stock Exchange: A Bird’s Eye View’, *Journal of Finance*, Vol. 49, No. 3, pp. 951–984.

Li, Y., Van Ness, B. and Van Ness, R. (2005). ‘Daily and Intraday Patterns in Spread and Depth for Limit Orders and Specialists’, *Quaterly journal of business and economics*, Vol. 44, No. 3, pp. 3–14.

LiCalzi, M. and Pellizzari, P. (2003). ‘Fundamentalists clashing over the book: a study of order-driven stock markets’, *Quantitative Finance*, Vol. 3, No. 6, pp. 470–480.

LiCalzi, M. and Pellizzari, P. (2007). ‘Simple market protocols for efficient risk sharing’, *Journal of Economic Dynamics and Control*, Vol. 31, No. 11, pp. 3568–3590.

Lillo, F., Farmer, J. and Mantegna, R. (2002). ‘Single Curve Collapse of the Price Impact Function for the New York Stock Exchange’. Available at: http://arxiv.org/abs/cond-mat/0207428.

Lillo, F., Farmer, J. and Mantegna, R. (2003). 'Master curve for price impact function', *Nature*, Vol. 421, No. 3, pp. 129–130.

Lillo, F., Mike, S. and Farmer, J. (2005). 'Theory for long memory in supply and demand', *Physical Review E*, Vol. 71, No. 6, pp. 1–11.

Lim, M. and Coggins, R. (2005a). 'The immediate price impact of trades on the Australian Stock Exchange', *Quantitative Finance*, Vol. 5, No. 4, pp. 365–377.

Lim, M. and Coggins, R. (2005b). 'Optimal trade execution: an evolutionary approach', in *Proceedings of the IEEE Congress on Evolutionary Computation*, Vol. 2, pp. 1045–1052.

Liu, W. (2009). 'Monitoring and limit order submission risks', *Journal of Financial Markets*, Vol. 12, No. 1, pp. 107–141.

Lo, A., MacKinlay, A. and Zhang, J. (2002). 'Econometric models of limit-order executions', *Journal of Financial Economics*, Vol. 65, No. 1, pp. 31–71.

Lo, I. and Sapp, S. (2010). 'Order Aggressiveness and Quantity: How Are They Determined in a Limit Order Market?', *Journal of International Financial Markets, Institutions and Money*, Vol. 20, No. 3, pp. 223–243.

Lockwood, L. and Linn, S. (1990). 'An Examination of Stock Market Return Volatility During Overnight and Intraday Periods', *Journal of Finance*, Vol. 45, No. 2, pp. 591–601.

Lux, T. (1998). 'The socio-economic dynamics of speculative markets: interacting agents, chaos, and the fat tails of return distributions', *Journal of Economic Behavior & Organization*, Vol. 33, No. 2, pp. 143–165.

Lux, T. (2009). 'Stochastic Behavioral Asset Pricing Models and the Stylized Facts', in *Handbook of Financial Markets: Dynamics and Evolution*, North Holland.

Madhavan, A. (1992). 'Trading Mechanisms in Securities Markets', *Journal of Finance*, Vol. 47, No. 2, pp. 607–641.

Madhavan, A. (2000). 'Market Microstructure: A Survey', *Journal of Financial Markets*, Vol. 3, No. 1, pp. 205–258.

Madhavan, A., Richardson, M. and Roomans, M. (1997). 'Why do security prices change? A transaction-level analysis of NYSE stocks', *Review of Financial Studies*, Vol. 10, No. 4, pp. 1035–1064.

Maslov, S. (2000). 'Simple model of a limit order-driven market', *Physica A*, Vol. 278, No. 3, pp. 571–578.

Maslov, S. and Millis, M. (2001). 'Price fluctuations from the order book perspective - empirical facts and a simple model', *Physica A*, Vol. 299, No. 1, p. 234.

McInish, T. and Wood, R. (1990a). 'An analysis of transactions data for the Toronto Stock Exchange: Return patterns and end-of-the-day effect', *Journal of Banking & Finance*, Vol. 14, No. 2, pp. 441–458.

McInish, T. and Wood, R. (1990b). 'A transactions data analysis of the variability of common stock returns during 1980-1984', *Journal of Banking and Finance*, Vol. 14, No. 1, pp. 99–112.

McInish, T. and Wood, R. (1992). 'An Analysis of Intraday Patterns in Bid-Ask Spreads for NYSE Stocks', *Journal of Finance*, Vol. 47, No. 2, pp. 753–764.

McKay, R., Hoai, N., Whigham, P., Shan, Y. and O'Neill, M. (2010). 'Grammar-based Genetic Programming: a survey', *Genetic Programming and Evolvable Machines*, Vol. 11, No. 4, pp. 365–396.

Menkhoff, L., Osler, C. and Schmeling, M. (2010). 'Limit-order submission strategies under asymmetric information', *Journal of Banking and Finance*, Vol. 34, No. 11, pp. 2665–2677.

Mike, S. and Farmer, J. (2008). 'An empirical behavioral model of liquidity and volatility', *Journal of Economic Dynamics and Control*, Vol. 32, No. 1, pp. 200–234.

Miller, M. (2008). 'Don't let your robots grow up to be traders: Artificial intelligence, human intelligence, and asset-market bubbles', *Journal of Economic Behavior & Organization*, Vol. 68, No. 1, pp. 163–166.

Mishra, S., Daigler, R. and Holowczack, R. (2012). 'The Effect of Algorithmic Trading on Liquidity in the Options Market'. Available at: http://people.stern.nyu.edu/jhasbrou (accessed 12 Nov. 2012).

Moazeni, S., Coleman, T. and Li, Y. (2010). 'Optimal Portfolio Execution Strategies and Sensitivity to Price Impact Parameters', *SIAM Journal on Optimization*, Vol. 20, No. 3, pp. 1620–1654.

Neely, C. and Weller, P. (1999). 'Technical trading rules in the European monetary system', *Journal of International Money and Finance*, Vol. 18, No. 1, pp. 429–458.

Neely, C. and Weller, P. (2001). 'Technical analysis and central bank intervention', *Technical analysis and central bank intervention*, Vol. 20, No. 7, pp. 949–970.

Neely, C. and Weller, P. (2003). 'Intraday technical trading in the foreign exchange market', *Intraday technical trading in the foreign exchange market*, Vol. 22, No. 2, pp. 223–237.

Neely, C., Weller, P. and Dittmar, R. (1997). 'Is technical analysis in the foreign exchange market profitable? A genetic programming approach', *Journal of Financial and Quantitative Analysis*, Vol. 32, No. 4, pp. 405–428.

NYSE (2006). *TAQ 3 User's Guide (version 1.0.9)*, New York Stock Exchange Inc.

Obizhaeva, A. A. and Wang, J. (2012). 'Optimal trading strategy and supply/demand dynamics', *Journal of Financial Markets (forthcoming)*.

O'Hara, M. (1995). *Market Microstructure Theory*, Blackwell Publishers, Cambridge US.

Omura, K., Tanigawa, Y. and Uno, J. (2000). 'Execution Probability of Limit Orders on the Tokyo Stock Exchange'. Available at: http://ssrn.com/abstract=252588.

O'Neill, M. and Brabazon, A. (2009). 'Recent Patents on Genetic Programming', *Recent Patents on Computer Science*, Vol. 2, No. 1, pp. 43–49.

O'Neill, M. and Ryan, C. (2003). *Grammatical Evolution: Evolutionary Automatic Programming in an Arbitrary Language*, Kluwer Academic Publishers, Boston US.

Palit, I., Phelps, S. and Ng, W. (2012). 'Can a Zero-Intelligence Plus Model Explain the Stylized Facts of Financial Time Series Data?', in *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems*, pp. 653–660.

Palmer, R., Arthur, W., Holland, J., LeBaron, B. and Tayler, P. (1994). 'Artificial economic life: a simple model of a stockmarket', *Physica D: Nonlinear Phenomena*, Vol. 75, No. 1-3, pp. 264–274.

Parlour, C. (1998). 'Price dynamics in limit order markets', *Review of Financial Studies*, Vol. 11, No. 4, pp. 789–816.

Parlour, C. and Seppi, D. (2008). 'Limit Order Markets: A Survey', in A. V. Thakor & A. W. Boot (eds.), *Handbook of Financial Intermediation and Banking*, pp. 63–96, Elsevier, Amsterdam.

Pascual, R. and Verdas, D. (2009). 'What pieces of limit order book information matter in explaining order choice by patient and impatient traders?', *Quantitative Finance*, Vol. 9, No. 1, pp. 527–545.

Perold, A. (1988). 'The Implementation Shortfall: Paper vs. Reality', *Journal of Portfolio Management*, Vol. 14, No. 3, pp. 4–9.

Peterson, M. and Sirri, E. (2002). 'Order Submission Strategy And The Curious Case Of Marketable Limit Orders', *Journal of Financial and Quantitative Analysis*, Vol. 37, No. 2, pp. 221–241.

Peterson, M. and Sirri, E. (2003). 'Evaluation of the biases in execution cost estimation using trade and quote data', *Journal of Financial Markets*, Vol. 6, No. 3, pp. 259–280.

Potters, M. and Bouchaud, J. (2003). 'More statistical properties of order books and price impact', *Physica A: Statistical Mechanics and its Applications*, Vol. 324, No. 1, pp. 133–140.

Predoiu, S., Shaikhet, G. and Shreve, S. (2011). 'Optimal Execution in a General One-Sided Limit-Order Book', *SIAM Journal on Financial Mathematics*, Vol. 2, No. 1, pp. 183–212.

Raberto, M. and Cincotti, S. (2005). 'Modeling and simulation of a double auction artificial financial market', *Physica A: Statistical Mechanics and its applications*, Vol. 355, No. 1, pp. 34–45.

Ranaldo, A. (2004). 'Order Aggressiveness in Limit Order Book Markets', *Journal of Financial Markets*, Vol. 7, No. 1, pp. 53–74.

Rawal, D. (2010). 'Bringing intelligent decision-making to order routing', *Journal of Trading*, Vol. 5, No. 1, pp. 30–34.

Rosu, I. (2009). 'A Dynamic Model of the Limit Order Book', *Review of Financial Studies*, Vol. 22, No. 1, pp. 4601–4641.

Saar, G. (2001). 'Price Impact Asymmetry of Block Trades: an institutional trading explanation', *Review of Financial Studies*, Vol. 14, No. 4, pp. 1153–1181.

Samanidou, E., Zschischang, E., Stauffer, D. and Lux, T. (2007). 'Agent-based models of financial markets', *Reports on Progress in Physics*, Vol. 70, No. 3, pp. 409–450.

Sandas, P. (2001). 'Adverse Selection and Competitive Market Making: Empirical Evidence from a Limit Order Market', *Review of Financial Studies*, Vol. 14, No. 3, pp. 705–734.

Schelling, T. (1971). 'Dynamic models of segregation', *Journal of Mathematical Sociology*, Vol. 1, No. 2, pp. 143–186.

Schmidt, A. (2010). 'Optimal Execution in the Global FX Market', *Journal of Trading*, Vol. 5, No. 3, pp. 68–77.

Sefiane, S. and Benbouziane, M. (2012). ‘Portfolio Selection Using Genetic Algorithm’, *Journal of Applied Finance & Banking*, Vol. 2, No. 4, pp. 143–154.

Seppi, D. (1997). ‘Liquidity provision with limit orders and a strategic specialist’, *Review of Financial Studies*, Vol. 10, No. 2, pp. 103–150.

Simon, H. (1996). *The Sciences of the Artificial*, MIT Press, 3rd edn.

Smith, E., Farmer, J., Gillemot, L. and Krishnamurthy, S. (2003). ‘Statistical theory of the continous double auction’, *Quantitative Finance*, Vol. 3, No. 1, pp. 481–514.

Spencer, H. (1864). *The Principles of Biology*, Vol. 1, Williams and Norgate, London and Edinburgh.

Suzuki, K., Shimokawa, T. and Misawa, T. (2009). ‘An Agent-Based Approach to Option Pricing Anomalies’, *IEEE Transactions on Evolutionary Computation*, Vol. 13, No. 1, pp. 19–32.

Swarm (2009). ‘Swarm package can be obtained from’. Available at: http://www.swarm.org.

Tesfatsion, L. (2006). ‘Agent-based computational economics: A constructive approach to economic theory’, in L. Tesfatsion & K. Judd (eds.), *Handbook of computational economics: agent-based computational economics*, pp. 52–74, Elsevier, Amsterdam.

Torre, N. (1997). *BARRA Market Impact Model Handbook*, BARRA Inc., Berkeley.

Toth, B., Kertesz, J. and Farmer, J. (2009). ‘Studies of the limit order book around large price changes’, *European Physical Journal B.*, Vol. 71, No. 1, pp. 499–510.

Vergote, O. (2005). 'How to match trades and quotes for NYSE stocks'. Available at: http://ssrn.com/abstract808984.

Verhoeven, P., Ching, S. and Ng, H. (2004). 'Determinants of the decision to submit market or limit orders on the ASX', *Pacific-Basin Finance Journal*, Vol. 12, No. 3, pp. 1–18.

Vo, M. (2007). 'Limit orders and the intraday behavior of market liquidity: Evidence from the Toronto stock exchange', *Global Finance Journal*, Vol. 17, No. 3, pp. 379–396.

Wald, J. and Horriga, H. (2005). 'Optimal Limit Order Choice', *Journal of Business*, Vol. 78, No. 2, pp. 597–620.

Weber, P. and Rosenow, B. (2005). 'Order book approach to price impact', *Quantitative Finance*, Vol. 5, No. 1, pp. 357–364.

Weber, P. and Rosenow, B. (2006). 'Large stock price changes: volume or liquidity?', *Quantitative Finance*, Vol. 6, No. 1, pp. 7–14.

Wermers, R. (2002). 'Mutual fund performance: An empirical decomposition into stock-picking talent, style, transactions costs, and expenses', *The Journal of Finance*, Vol. 55, No. 4, pp. 1655–1703.

Werner, I. and Kleidon, A. (1996). 'U.K. and U.S. trading of British cross-listed stocks: an intraday analysis of market integration', *Review of Financial Studies*, Vol. 9, No. 2, pp. 619–664.

Wood, R., MiInish, T. and Ord, J. (1985). 'An Investigation of Transactions Data for NYSE Stocks', *Journal of Finance*, Vol. 40, No. 3, pp. 723–739.

Xu, Y. (2009). 'Order aggressiveness on the ASX market', *International Journal of Economics and Finance*, Vol. 1, No. 1, pp. 51–75.

Yang, J. (2002). 'The Efficiency of an Artificial Double Auction Stock Market with Neural Learning Agents', in S. Chen (ed.), *Evolutionary Computation in Economics and Finance*, pp. 85–106, Springer-Verlag.

Zovko, I. and Farmer, J. (2002). 'The power of patience: a behavioural regularity in limit-order placement', *Quantitative Finance*, Vol. 2, No. 5, pp. 387–392.